Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.
# Table of Contents

Welcome .......................................................................................................................... 1
Amazon SageMaker Service ............................................................................................... 1
Amazon SageMaker Runtime ............................................................................................. 1
Amazon Sagemaker Edge Manager .................................................................................. 1
Amazon SageMaker Feature Store Runtime ..................................................................... 1
Amazon SageMaker Service ............................................................................................. 1

**Actions** .......................................................................................................................... 2
Amazon SageMaker Service ............................................................................................... 8
AddAssociation .................................................................................................................. 14
AddTags ............................................................................................................................. 17
AssociateTrialComponent ................................................................................................. 19
BatchDescribeModelPackage ............................................................................................ 21
CreateAction ..................................................................................................................... 24
CreateAlgorithm ................................................................................................................. 27
CreateApp .......................................................................................................................... 32
CreateAppImageConfig ..................................................................................................... 35
CreateArtifact ................................................................................................................... 37
CreateAutoMLJob .............................................................................................................. 40
CreateCodeRepository .................................................................................................... 44
CreateCompilationJob ....................................................................................................... 46
CreateContext .................................................................................................................... 50
CreateDataQualityJobDefinition ....................................................................................... 53
CreateDeviceFleet .............................................................................................................. 57
CreateDomain .................................................................................................................... 60
CreateEdgePackagingJob ................................................................................................. 66
CreateEndpoint .................................................................................................................. 69
CreateEndpointConfig ...................................................................................................... 73
CreateExperiment .............................................................................................................. 77
CreateFeatureGroup ......................................................................................................... 80
CreateFlowDefinition ....................................................................................................... 85
CreateHumanTaskUi ............................................................................................................. 88
CreateHyperParameterTuningJob ..................................................................................... 90
CreateImage ...................................................................................................................... 97
CreateImageVersion ......................................................................................................... 100
CreateInferenceRecommendationsJob .............................................................................. 103
CreateLabelingJob .......................................................................................................... 107
CreateModel ...................................................................................................................... 113
CreateModelBiasJobDefinition ......................................................................................... 117
CreateModelExplainabilityJobDefinition ......................................................................... 121
CreateModelPackage ........................................................................................................ 125
CreateModelPackageGroup ............................................................................................... 134
CreateModelQualityJobDefinition .................................................................................... 136
CreateMonitoringSchedule ............................................................................................... 140
CreateNotebookInstance ................................................................................................. 144
CreateNotebookInstanceLifecycleConfig .......................................................................... 150
CreatePipeline .................................................................................................................. 153
CreatePresignedDomainUrl ............................................................................................. 157
CreatePresignedNotebookInstanceUrl ............................................................................. 160
CreateProcessingJob ........................................................................................................ 162
CreateProject .................................................................................................................... 167
CreateStudioLifecycleConfig ........................................................................................... 170
CreateTrainingJob ............................................................................................................ 173
CreateTransformJob ......................................................................................................... 182
CreateTrial ......................................................................................................................... 188
CreateTrialComponent ..................................................................................................... 191
<table>
<thead>
<tr>
<th>Method</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DescribeEndpointConfig</td>
<td>343</td>
</tr>
<tr>
<td>DescribeExperiment</td>
<td>346</td>
</tr>
<tr>
<td>DescribeFeatureGroup</td>
<td>349</td>
</tr>
<tr>
<td>DescribeFlowDefinition</td>
<td>354</td>
</tr>
<tr>
<td>DescribeHumanTaskUI</td>
<td>357</td>
</tr>
<tr>
<td>DescribeHyperParameterTuningJob</td>
<td>359</td>
</tr>
<tr>
<td>DescribeImage</td>
<td>367</td>
</tr>
<tr>
<td>DescribeImageVersion</td>
<td>370</td>
</tr>
<tr>
<td>DescribeInferenceRecommendationsJob</td>
<td>373</td>
</tr>
<tr>
<td>DescribeLabelingJob</td>
<td>377</td>
</tr>
<tr>
<td>DescribeLineageGroup</td>
<td>383</td>
</tr>
<tr>
<td>DescribeModel</td>
<td>386</td>
</tr>
<tr>
<td>DescribeModelBiasJobDefinition</td>
<td>389</td>
</tr>
<tr>
<td>DescribeModelExplainabilityJobDefinition</td>
<td>393</td>
</tr>
<tr>
<td>DescribeModelPackage</td>
<td>397</td>
</tr>
<tr>
<td>DescribeModelPackageGroup</td>
<td>405</td>
</tr>
<tr>
<td>DescribeModelQualityJobDefinition</td>
<td>408</td>
</tr>
<tr>
<td>DescribeMonitoringSchedule</td>
<td>412</td>
</tr>
<tr>
<td>DescribeNotebookInstance</td>
<td>416</td>
</tr>
<tr>
<td>DescribeNotebookInstanceLifecycleConfig</td>
<td>421</td>
</tr>
<tr>
<td>DescribePipeline</td>
<td>424</td>
</tr>
<tr>
<td>DescribePipelineDefinitionForExecution</td>
<td>428</td>
</tr>
<tr>
<td>DescribePipelineExecution</td>
<td>430</td>
</tr>
<tr>
<td>DescribeProcessingJob</td>
<td>434</td>
</tr>
<tr>
<td>DescribeProject</td>
<td>440</td>
</tr>
<tr>
<td>DescribeStudioLifecycleConfig</td>
<td>443</td>
</tr>
<tr>
<td>DescribeSubscribedWorkteam</td>
<td>446</td>
</tr>
<tr>
<td>DescribeTrainingJob</td>
<td>448</td>
</tr>
<tr>
<td>DescribeTransformJob</td>
<td>459</td>
</tr>
<tr>
<td>DescribeTrial</td>
<td>464</td>
</tr>
<tr>
<td>DescribeTrialComponent</td>
<td>467</td>
</tr>
<tr>
<td>DescribeUserProfile</td>
<td>472</td>
</tr>
<tr>
<td>DescribeWorkforce</td>
<td>476</td>
</tr>
<tr>
<td>DescribeWorkteam</td>
<td>478</td>
</tr>
<tr>
<td>DisableSagemakerServicecatalogPortfolio</td>
<td>480</td>
</tr>
<tr>
<td>DisassociateTrialComponent</td>
<td>481</td>
</tr>
<tr>
<td>EnableSagemakerServicecatalogPortfolio</td>
<td>483</td>
</tr>
<tr>
<td>GetDeviceFleetReport</td>
<td>484</td>
</tr>
<tr>
<td>GetLineageGroupPolicy</td>
<td>487</td>
</tr>
<tr>
<td>GetModelPackageGroupPolicy</td>
<td>489</td>
</tr>
<tr>
<td>GetSagemakerServicecatalogPortfolioStatus</td>
<td>491</td>
</tr>
<tr>
<td>GetSearchSuggestions</td>
<td>492</td>
</tr>
<tr>
<td>ListActions</td>
<td>494</td>
</tr>
<tr>
<td>ListAlgorithms</td>
<td>497</td>
</tr>
<tr>
<td>ListAppImageConfigs</td>
<td>500</td>
</tr>
<tr>
<td>ListApps</td>
<td>503</td>
</tr>
<tr>
<td>ListArtifacts</td>
<td>506</td>
</tr>
<tr>
<td>ListAssociations</td>
<td>509</td>
</tr>
<tr>
<td>ListAutoMLJobs</td>
<td>513</td>
</tr>
<tr>
<td>ListCandidatesForAutoMLJob</td>
<td>517</td>
</tr>
<tr>
<td>ListCodeRepositories</td>
<td>521</td>
</tr>
<tr>
<td>ListCompilationJobs</td>
<td>525</td>
</tr>
<tr>
<td>ListContexts</td>
<td>529</td>
</tr>
<tr>
<td>ListDataQualityJobDefinitions</td>
<td>532</td>
</tr>
<tr>
<td>ListDeviceFleets</td>
<td>535</td>
</tr>
<tr>
<td>ListDevices</td>
<td>538</td>
</tr>
<tr>
<td>ListDomains</td>
<td>541</td>
</tr>
<tr>
<td>Function</td>
<td>Page</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>StopHyperParameterTuningJob</td>
<td>729</td>
</tr>
<tr>
<td>StopEdgePackagingJob</td>
<td>728</td>
</tr>
<tr>
<td>StartNotebookInstance</td>
<td>719</td>
</tr>
<tr>
<td>SendPipelineExecutionStepFailure</td>
<td>713</td>
</tr>
<tr>
<td>RetryPipelineExecution</td>
<td>688</td>
</tr>
<tr>
<td>RenderUITemplate</td>
<td>685</td>
</tr>
<tr>
<td>QueryLineage</td>
<td>679</td>
</tr>
<tr>
<td>PutModelPackageGroupPolicy</td>
<td>677</td>
</tr>
<tr>
<td>ListTag</td>
<td>649</td>
</tr>
<tr>
<td>ListTrainingJobs</td>
<td>651</td>
</tr>
<tr>
<td>ListTrainingJobsForHyperParameterTuningJob</td>
<td>655</td>
</tr>
<tr>
<td>ListTransformJobs</td>
<td>658</td>
</tr>
<tr>
<td>ListTrialComponents</td>
<td>661</td>
</tr>
<tr>
<td>ListTrials</td>
<td>665</td>
</tr>
<tr>
<td>ListUserProfile</td>
<td>668</td>
</tr>
<tr>
<td>ListWorkforces</td>
<td>671</td>
</tr>
<tr>
<td>ListWorkteams</td>
<td>674</td>
</tr>
<tr>
<td>PutModelPackageGroupPolicy</td>
<td>677</td>
</tr>
<tr>
<td>QueryLineage</td>
<td>679</td>
</tr>
<tr>
<td>RegisterDevices</td>
<td>683</td>
</tr>
<tr>
<td>RenderUITemplate</td>
<td>685</td>
</tr>
<tr>
<td>RetryPipelineExecution</td>
<td>688</td>
</tr>
<tr>
<td>Search</td>
<td>691</td>
</tr>
<tr>
<td>SendPipelineExecutionStepFailure</td>
<td>713</td>
</tr>
<tr>
<td>SendPipelineExecutionStepSuccess</td>
<td>715</td>
</tr>
<tr>
<td>StartMonitoringSchedule</td>
<td>717</td>
</tr>
<tr>
<td>StartNotebookInstance</td>
<td>719</td>
</tr>
<tr>
<td>StartPipelineExecution</td>
<td>721</td>
</tr>
<tr>
<td>StopAutoMLJob</td>
<td>724</td>
</tr>
<tr>
<td>StopCompilationJob</td>
<td>726</td>
</tr>
<tr>
<td>StopEdgePackagingJob</td>
<td>728</td>
</tr>
<tr>
<td>StopHyperParameterTuningJob</td>
<td>729</td>
</tr>
<tr>
<td>Class</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>NestedFilters</td>
<td>1222</td>
</tr>
<tr>
<td>NetworkConfig</td>
<td>1223</td>
</tr>
<tr>
<td>NotebookInstanceLifecycleConfigSummary</td>
<td>1224</td>
</tr>
<tr>
<td>NotebookInstanceLifecycleHook</td>
<td>1225</td>
</tr>
<tr>
<td>NotebookInstanceSummary</td>
<td>1226</td>
</tr>
<tr>
<td>NotificationConfiguration</td>
<td>1229</td>
</tr>
<tr>
<td>ObjectiveStatusCounters</td>
<td>1230</td>
</tr>
<tr>
<td>OfflineStoreConfig</td>
<td>1231</td>
</tr>
<tr>
<td>OfflineStoreStatus</td>
<td>1232</td>
</tr>
<tr>
<td>OidcConfig</td>
<td>1233</td>
</tr>
<tr>
<td>OidcConfigForResponse</td>
<td>1235</td>
</tr>
<tr>
<td>OidcMemberDefinition</td>
<td>1237</td>
</tr>
<tr>
<td>OnlineStoreConfig</td>
<td>1238</td>
</tr>
<tr>
<td>OnlineStoreSecurityConfig</td>
<td>1239</td>
</tr>
<tr>
<td>OutputConfig</td>
<td>1240</td>
</tr>
<tr>
<td>OutputDataConfig</td>
<td>1244</td>
</tr>
<tr>
<td>OutputParameter</td>
<td>1246</td>
</tr>
<tr>
<td>ParallelismConfiguration</td>
<td>1247</td>
</tr>
<tr>
<td>Parameter</td>
<td>1248</td>
</tr>
<tr>
<td>ParameterRange</td>
<td>1249</td>
</tr>
<tr>
<td>ParameterRanges</td>
<td>1250</td>
</tr>
<tr>
<td>Parent</td>
<td>1251</td>
</tr>
<tr>
<td>ParentHyperParameterTuningJob</td>
<td>1252</td>
</tr>
<tr>
<td>PendingDeploymentSummary</td>
<td>1253</td>
</tr>
<tr>
<td>PendingProductionVariantSummary</td>
<td>1254</td>
</tr>
<tr>
<td>Phase</td>
<td>1257</td>
</tr>
<tr>
<td>Pipeline</td>
<td>1258</td>
</tr>
<tr>
<td>PipelineDefinitionS3Location</td>
<td>1261</td>
</tr>
<tr>
<td>PipelineExecution</td>
<td>1262</td>
</tr>
<tr>
<td>PipelineExecutionStep</td>
<td>1265</td>
</tr>
<tr>
<td>PipelineExecutionStepMetadata</td>
<td>1267</td>
</tr>
<tr>
<td>PipelineExecutionSummary</td>
<td>1270</td>
</tr>
<tr>
<td>PipelineExperimentConfig</td>
<td>1272</td>
</tr>
<tr>
<td>PipelineSummary</td>
<td>1273</td>
</tr>
<tr>
<td>ProcessingClusterConfig</td>
<td>1275</td>
</tr>
<tr>
<td>ProcessingFeatureStoreOutput</td>
<td>1277</td>
</tr>
<tr>
<td>ProcessingInput</td>
<td>1278</td>
</tr>
<tr>
<td>ProcessingJob</td>
<td>1279</td>
</tr>
<tr>
<td>ProcessingJobStepMetadata</td>
<td>1283</td>
</tr>
<tr>
<td>ProcessingJobSummary</td>
<td>1284</td>
</tr>
<tr>
<td>ProcessingOutput</td>
<td>1286</td>
</tr>
<tr>
<td>ProcessingOutputConfig</td>
<td>1287</td>
</tr>
<tr>
<td>ProcessingResources</td>
<td>1288</td>
</tr>
<tr>
<td>ProcessingS3Input</td>
<td>1289</td>
</tr>
<tr>
<td>ProcessingS3Output</td>
<td>1291</td>
</tr>
<tr>
<td>ProcessingStoppingCondition</td>
<td>1292</td>
</tr>
<tr>
<td>ProductionVariant</td>
<td>1293</td>
</tr>
<tr>
<td>ProductionVariantCoreDumpConfig</td>
<td>1295</td>
</tr>
<tr>
<td>ProductionVariantServerlessConfig</td>
<td>1297</td>
</tr>
<tr>
<td>ProductionVariantStatus</td>
<td>1298</td>
</tr>
<tr>
<td>ProductionVariantSummary</td>
<td>1299</td>
</tr>
<tr>
<td>ProfilerConfig</td>
<td>1301</td>
</tr>
<tr>
<td>ProfilerConfigForUpdate</td>
<td>1303</td>
</tr>
<tr>
<td>ProfilerRuleConfiguration</td>
<td>1305</td>
</tr>
<tr>
<td>ProfilerRuleEvaluationStatus</td>
<td>1307</td>
</tr>
<tr>
<td>Project</td>
<td>1309</td>
</tr>
<tr>
<td>ProjectSummary</td>
<td>1312</td>
</tr>
</tbody>
</table>
Welcome

Amazon SageMaker Service

Provides APIs for creating and managing SageMaker resources.

Other Resources:
- SageMaker Developer Guide
- Amazon Augmented AI Runtime API Reference

Amazon SageMaker Runtime

The Amazon SageMaker runtime API.

Amazon Sagemaker Edge Manager

SageMaker Edge Manager dataplane service for communicating with active agents.

Amazon SageMaker Feature Store Runtime

Contains all data plane API operations and data types for the Amazon SageMaker Feature Store. Use this API to put, delete, and retrieve (get) features from a feature store.

Use the following operations to configure your OnlineStore and OfflineStore features, and to create and manage feature groups:

- CreateFeatureGroup
- DeleteFeatureGroup
- DescribeFeatureGroup
- ListFeatureGroups
Actions

The following actions are supported by Amazon SageMaker Service:

- AddAssociation (p. 14)
- AddTags (p. 17)
- AssociateTrialComponent (p. 19)
- BatchDescribeModelPackage (p. 21)
- CreateAction (p. 24)
- CreateAlgorithm (p. 27)
- CreateApp (p. 32)
- CreateAppImageConfig (p. 35)
- CreateArtifact (p. 37)
- CreateAutoMLJob (p. 40)
- CreateCodeRepository (p. 44)
- CreateCompilationJob (p. 46)
- CreateContext (p. 50)
- CreateDataQualityJobDefinition (p. 53)
- CreateDeviceFleet (p. 57)
- CreateDomain (p. 60)
- CreateEdgePackagingJob (p. 66)
- CreateEndpoint (p. 69)
- CreateEndpointConfig (p. 73)
- CreateExperiment (p. 77)
- CreateFeatureGroup (p. 80)
- CreateFlowDefinition (p. 85)
- CreateHumanTaskUi (p. 88)
- CreateHyperParameterTuningJob (p. 90)
- CreateImage (p. 97)
- CreateImageVersion (p. 100)
- CreateInferenceRecommendationsJob (p. 103)
- CreateLabelingJob (p. 107)
- CreateModel (p. 113)
- CreateModelBiasJobDefinition (p. 117)
- CreateModelExplainabilityJobDefinition (p. 121)
- CreateModelPackage (p. 125)
- CreateModelPackageGroup (p. 134)
- CreateModelQualityJobDefinition (p. 136)
- CreateMonitoringSchedule (p. 140)
- CreateNotebookInstance (p. 144)
- CreateNotebookInstanceLifecycleConfig (p. 150)
- CreatePipeline (p. 153)
- CreatePresignedDomainUrl (p. 157)
• CreatePresignedNotebookInstanceUrl (p. 160)
• CreateProcessingJob (p. 162)
• CreateProject (p. 167)
• CreateStudioLifecycleConfig (p. 170)
• CreateTrainingJob (p. 173)
• CreateTransformJob (p. 182)
• CreateTrial (p. 188)
• CreateTrialComponent (p. 191)
• CreateUserProfile (p. 195)
• CreateWorkforce (p. 199)
• CreateWorkteam (p. 202)
• DeleteAction (p. 205)
• DeleteAlgorithm (p. 207)
• DeleteApp (p. 208)
• DeleteAppImageConfig (p. 210)
• DeleteArtifact (p. 212)
• DeleteAssociation (p. 214)
• DeleteCodeRepository (p. 216)
• DeleteContext (p. 217)
• DeleteDataQualityJobDefinition (p. 219)
• DeleteDeviceFleet (p. 221)
• DeleteDomain (p. 223)
• DeleteEndpoint (p. 225)
• DeleteEndpointConfig (p. 227)
• DeleteExperiment (p. 229)
• DeleteFeatureGroup (p. 231)
• DeleteFlowDefinition (p. 233)
• DeleteHumanTaskUi (p. 235)
• DeleteImage (p. 237)
• DeleteImageVersion (p. 239)
• DeleteModel (p. 241)
• DeleteModelBiasJobDefinition (p. 243)
• DeleteModelExplainabilityJobDefinition (p. 245)
• DeleteModelPackage (p. 247)
• DeleteModelPackageGroup (p. 249)
• DeleteModelPackageGroupPolicy (p. 251)
• DeleteModelQualityJobDefinition (p. 252)
• DeleteMonitoringSchedule (p. 254)
• DeleteNotebookInstance (p. 256)
• DeleteNotebookInstanceLifecycleConfig (p. 258)
• DeletePipeline (p. 259)
• DeleteProject (p. 261)
• DeleteStudioLifecycleConfig (p. 263)
• DeleteTags (p. 265)
• DeleteTrial (p. 267)
• DeleteTrialComponent (p. 269)
• DeleteUserProfile (p. 271)
• DeleteWorkforce (p. 273)
• DeleteWorkteam (p. 275)
• DeregisterDevices (p. 277)
• DescribeAction (p. 279)
• DescribeAlgorithm (p. 283)
• DescribeApp (p. 289)
• DescribeAppImageConfig (p. 293)
• DescribeArtifact (p. 296)
• DescribeAutoMLJob (p. 300)
• DescribeCodeRepository (p. 306)
• DescribeCompilationJob (p. 308)
• DescribeContext (p. 313)
• DescribeDataQualityJobDefinition (p. 317)
• DescribeDevice (p. 321)
• DescribeDeviceFleet (p. 325)
• DescribeDomain (p. 328)
• DescribeEdgePackagingJob (p. 334)
• DescribeEndpoint (p. 338)
• DescribeEndpointConfig (p. 343)
• DescribeExperiment (p. 346)
• DescribeFeatureGroup (p. 349)
• DescribeFlowDefinition (p. 354)
• DescribeHumanTaskUi (p. 357)
• DescribeHyperParameterTuningJob (p. 359)
• DescribeImage (p. 367)
• DescribeImageVersion (p. 370)
• DescribeInferenceRecommendationsJob (p. 373)
• DescribeLabelingJob (p. 377)
• DescribeLineageGroup (p. 383)
• DescribeModel (p. 386)
• DescribeModelBiasJobDefinition (p. 389)
• DescribeModelExplainabilityJobDefinition (p. 393)
• DescribeModelPackage (p. 397)
• DescribeModelPackageGroup (p. 405)
• DescribeModelQualityJobDefinition (p. 408)
• DescribeMonitoringSchedule (p. 412)
• DescribeNotebookInstance (p. 416)
• DescribeNotebookInstanceLifecycleConfig (p. 421)
• DescribePipeline (p. 424)
• DescribePipelineDefinitionForExecution (p. 428)
• DescribePipelineExecution (p. 430)
• DescribeProcessingJob (p. 434)
• DescribeProject (p. 440)
• DescribeStudioLifecycleConfig (p. 443)
• DescribeSubscribedWorkteam (p. 446)
• DescribeTrainingJob (p. 448)
• DescribeTransformJob (p. 459)
• DescribeTrial (p. 464)
• DescribeTrialComponent (p. 467)
• DescribeUserProfile (p. 472)
• DescribeWorkforce (p. 476)
• DescribeWorkteam (p. 478)
• DisableSagemakerServicecatalogPortfolio (p. 480)
• DisassociateTrialComponent (p. 481)
• EnableSagemakerServicecatalogPortfolio (p. 483)
• GetDeviceFleetReport (p. 484)
• GetLineageGroupPolicy (p. 487)
• GetModelPackageGroupPolicy (p. 489)
• GetSagemakerServicecatalogPortfolioStatus (p. 491)
• GetSearchSuggestions (p. 492)
• ListActions (p. 494)
• ListAlgorithms (p. 497)
• ListAppImageConfigs (p. 500)
• ListApps (p. 503)
• ListArtifacts (p. 506)
• ListAssociations (p. 509)
• ListAutoMLJobs (p. 513)
• ListCandidatesForAutoMLJob (p. 517)
• ListCodeRepositories (p. 521)
• ListCompilationJobs (p. 525)
• ListContexts (p. 529)
• ListDataQualityJobDefinitions (p. 532)
• ListDeviceFleets (p. 535)
• ListDevices (p. 538)
• ListDomains (p. 541)
• ListEdgePackagingJobs (p. 543)
• ListEndpointConfigs (p. 547)
• ListEndpoints (p. 550)
• ListExperiments (p. 553)
• ListFeatureGroups (p. 556)
• ListFlowDefinitions (p. 559)
• ListHumanTaskUis (p. 562)
• ListHyperParameterTuningJobs (p. 565)
• ListImages (p. 569)
• ListImageVersions (p. 572)
• ListInferenceRecommendationsJobs (p. 575)
• ListLabelingJobs (p. 578)
• ListLabelingJobsForWorkteam (p. 582)
• ListLineageGroups (p. 585)
• ListModelBiasJobDefinitions (p. 588)
• ListModelExplainabilityJobDefinitions (p. 591)
• ListModelMetadata (p. 594)
• ListModelPackageGroups (p. 596)
• ListModelPackages (p. 599)
• ListModelQualityJobDefinitions (p. 603)
• ListModels (p. 606)
• ListMonitoringExecutions (p. 609)
• ListMonitoringSchedules (p. 613)
• ListNotebookInstanceLifecycleConfigs (p. 617)
• ListNotebookInstances (p. 620)
• ListPipelineExecutions (p. 624)
• ListPipelineExecutionSteps (p. 627)
• ListPipelineParametersForExecution (p. 631)
• ListPipelines (p. 634)
• ListProcessingJobs (p. 637)
• ListProjects (p. 640)
• ListStudioLifecycleConfigs (p. 643)
• ListSubscribedWorkteams (p. 647)
• ListTags (p. 649)
• ListTrainingJobs (p. 651)
• ListTrainingJobsForHyperParameterTuningJob (p. 655)
• ListTransformJobs (p. 658)
• ListTrialComponents (p. 661)
• ListTrials (p. 665)
• ListUserProfiles (p. 668)
• ListWorkforces (p. 671)
• ListWorkteams (p. 674)
• PutModelPackageGroupPolicy (p. 677)
• QueryLineage (p. 679)
• RegisterDevices (p. 683)
• RenderUiTemplate (p. 685)
• RetryPipelineExecution (p. 688)
• Search (p. 691)
• SendPipelineExecutionStepFailure (p. 713)
• SendPipelineExecutionStepSuccess (p. 715)
• StartMonitoringSchedule (p. 717)
• StartNotebookInstance (p. 719)
• StartPipelineExecution (p. 721)
• StopAutoMLJob (p. 724)
• StopCompilationJob (p. 726)
• StopEdgePackagingJob (p. 728)
• StopHyperParameterTuningJob (p. 729)
• StopInferenceRecommendationsJob (p. 731)
• StopLabelingJob (p. 733)
• StopMonitoringSchedule (p. 735)
• StopNotebookInstance (p. 737)
• StopPipelineExecution (p. 739)
• StopProcessingJob (p. 741)
• StopTrainingJob (p. 743)
• StopTransformJob (p. 745)
• UpdateAction (p. 747)
• UpdateAppImageConfig (p. 750)
• UpdateArtifact (p. 752)
• UpdateCodeRepository (p. 755)
• UpdateContext (p. 757)
• UpdateDeviceFleet (p. 760)
• UpdateDevices (p. 762)
• UpdateDomain (p. 764)
• UpdateEndpoint (p. 767)
• UpdateEndpointWeightsAndCapacities (p. 771)
• UpdateExperiment (p. 773)
• UpdateImage (p. 775)
• UpdateModelPackage (p. 778)
• UpdateMonitoringSchedule (p. 781)
• UpdateNotebookInstance (p. 784)
• UpdateNotebookInstanceLifecycleConfig (p. 789)
• UpdatePipeline (p. 791)
• UpdatePipelineExecution (p. 794)
• UpdateProject (p. 797)
• UpdateTrainingJob (p. 800)
• UpdateTrial (p. 802)
• UpdateTrialComponent (p. 804)
• UpdateUserProfile (p. 808)
• UpdateWorkforce (p. 811)
• UpdateWorkteam (p. 814)

The following actions are supported by Amazon SageMaker Runtime:

• InvokeEndpoint (p. 817)
• InvokeEndpointAsync (p. 823)

The following actions are supported by Amazon Sagemaker Edge Manager:

• GetDeviceRegistration (p. 827)
• SendHeartbeat (p. 829)

The following actions are supported by Amazon SageMaker Feature Store Runtime:

• BatchGetRecord (p. 832)
• DeleteRecord (p. 835)
• GetRecord (p. 837)
• PutRecord (p. 840)
Amazon SageMaker Service

The following actions are supported by Amazon SageMaker Service:

- AddAssociation (p. 14)
- AddTags (p. 17)
- AssociateTrialComponent (p. 19)
- BatchDescribeModelPackage (p. 21)
- CreateAction (p. 24)
- CreateAlgorithm (p. 27)
- CreateApp (p. 32)
- CreateAppImageConfig (p. 35)
- CreateArtifact (p. 37)
- CreateAutoMLJob (p. 40)
- CreateCodeRepository (p. 44)
- CreateCompilationJob (p. 46)
- CreateContext (p. 50)
- CreateDataQualityJobDefinition (p. 53)
- CreateDeviceFleet (p. 57)
- CreateDomain (p. 60)
- CreateEdgePackagingJob (p. 66)
- CreateEndpoint (p. 69)
- CreateEndpointConfig (p. 73)
- CreateExperiment (p. 77)
- CreateFeatureGroup (p. 80)
- CreateFlowDefinition (p. 85)
- CreateHumanTaskUi (p. 88)
- CreateHyperParameterTuningJob (p. 90)
- CreateImage (p. 97)
- CreateImageVersion (p. 100)
- CreateInferenceRecommendationsJob (p. 103)
- CreateLabelingJob (p. 107)
- CreateModel (p. 113)
- CreateModelBiasJobDefinition (p. 117)
- CreateModelExplainabilityJobDefinition (p. 121)
- CreateModelPackage (p. 125)
- CreateModelPackageGroup (p. 134)
- CreateModelQualityJobDefinition (p. 136)
- CreateMonitoringSchedule (p. 140)
- CreateNotebookInstance (p. 144)
- CreateNotebookInstanceLifecycleConfig (p. 150)
- CreatePipeline (p. 153)
- CreatePresignedDomainUrl (p. 157)
- CreatePresignedNotebookInstanceUrl (p. 160)
- CreateProcessingJob (p. 162)
- CreateProject (p. 167)
• CreateStudioLifecycleConfig (p. 170)
• CreateTrainingJob (p. 173)
• CreateTransformJob (p. 182)
• CreateTrial (p. 188)
• CreateTrialComponent (p. 191)
• CreateUserProfile (p. 195)
• CreateWorkforce (p. 199)
• CreateWorkteam (p. 202)
• DeleteAction (p. 205)
• DeleteAlgorithm (p. 207)
• DeleteApp (p. 208)
• DeleteAppImageConfig (p. 210)
• DeleteArtifact (p. 212)
• DeleteAssociation (p. 214)
• DeleteCodeRepository (p. 216)
• DeleteContext (p. 217)
• DeleteDataQualityJobDefinition (p. 219)
• DeleteDeviceFleet (p. 221)
• DeleteDomain (p. 223)
• DeleteEndpoint (p. 225)
• DeleteEndpointConfig (p. 227)
• DeleteExperiment (p. 229)
• DeleteFeatureGroup (p. 231)
• DeleteFlowDefinition (p. 233)
• DeleteHumanTaskUi (p. 235)
• DeleteImage (p. 237)
• DeleteImageVersion (p. 239)
• DeleteModel (p. 241)
• DeleteModelBiasJobDefinition (p. 243)
• DeleteModelExplanationJobDefinition (p. 245)
• DeleteModelPackage (p. 247)
• DeleteModelPackageGroup (p. 249)
• DeleteModelPackageGroupPolicy (p. 251)
• DeleteModelQualityJobDefinition (p. 252)
• DeleteMonitoringSchedule (p. 254)
• DeleteNotebookInstance (p. 256)
• DeleteNotebookInstanceLifecycleConfig (p. 258)
• DeletePipeline (p. 259)
• DeleteProject (p. 261)
• DeleteStudioLifecycleConfig (p. 263)
• DeleteTags (p. 265)
• DeleteTrial (p. 267)
• DeleteTrialComponent (p. 269)
• DeleteUserProfile (p. 271)
• DeleteWorkforce (p. 273)
• DeleteWorkteam (p. 275)
• DeregisterDevices (p. 277)
• DescribeAction (p. 279)
• DescribeAlgorithm (p. 283)
• DescribeApp (p. 289)
• DescribeAppImageConfig (p. 293)
• DescribeArtifact (p. 296)
• DescribeAutoMLJob (p. 300)
• DescribeCodeRepository (p. 306)
• DescribeCompilationJob (p. 308)
• DescribeContext (p. 313)
• DescribeDataQualityJobDefinition (p. 317)
• DescribeDevice (p. 321)
• DescribeDeviceFleet (p. 325)
• DescribeDomain (p. 328)
• DescribeEdgePackagingJob (p. 334)
• DescribeEndpoint (p. 338)
• DescribeEndpointConfig (p. 343)
• DescribeExperiment (p. 346)
• DescribeFeatureGroup (p. 349)
• DescribeFlowDefinition (p. 354)
• DescribeHumanTaskUi (p. 357)
• DescribeHyperParameterTuningJob (p. 359)
• DescribeImage (p. 367)
• DescribeImageVersion (p. 370)
• DescribeInferenceRecommendationsJob (p. 373)
• DescribeLabelingJob (p. 377)
• DescribeLineageGroup (p. 383)
• DescribeModel (p. 386)
• DescribeModelBiasJobDefinition (p. 389)
• DescribeModelExplainabilityJobDefinition (p. 393)
• DescribeModelPackage (p. 397)
• DescribeModelPackageGroup (p. 405)
• DescribeModelQualityJobDefinition (p. 408)
• DescribeMonitoringSchedule (p. 412)
• DescribeNotebookInstance (p. 416)
• DescribeNotebookInstanceLifecycleConfig (p. 421)
• DescribePipeline (p. 424)
• DescribePipelineDefinitionForExecution (p. 428)
• DescribePipelineExecution (p. 430)
• DescribeProcessingJob (p. 434)
• DescribeProject (p. 440)
• DescribeStudioLifecycleConfig (p. 443)
• DescribeSubscribedWorkteam (p. 446)
• DescribeTrainingJob (p. 448)
• DescribeTransformJob (p. 459)
• DescribeTrial (p. 464)
• DescribeTrialComponent (p. 467)
• DescribeUserProfile (p. 472)
• DescribeWorkforce (p. 476)
• DescribeWorkteam (p. 478)
• DisableSagemakerServicecatalogPortfolio (p. 480)
• DisassociateTrialComponent (p. 481)
• EnableSagemakerServicecatalogPortfolio (p. 483)
• GetDeviceFleetReport (p. 484)
• GetLineageGroupPolicy (p. 487)
• GetModelPackageGroupPolicy (p. 489)
• GetSagemakerServicecatalogPortfolioStatus (p. 491)
• GetSearchSuggestions (p. 492)
• ListActions (p. 494)
• ListAlgorithms (p. 497)
• ListAppImageConfigs (p. 500)
• ListApps (p. 503)
• ListArtifacts (p. 506)
• ListAssociations (p. 509)
• ListAutoMLJobs (p. 513)
• ListCandidatesForAutoMLJob (p. 517)
• ListCodeRepositories (p. 521)
• ListCompilationJobs (p. 525)
• ListContexts (p. 529)
• ListDataQualityJobDefinitions (p. 532)
• ListDeviceFleets (p. 535)
• ListDevices (p. 538)
• ListDomains (p. 541)
• ListEdgePackagingJobs (p. 543)
• ListEndpointConfigs (p. 547)
• ListEndpoints (p. 550)
• ListExperiments (p. 553)
• ListFeatureGroups (p. 556)
• ListFlowDefinitions (p. 559)
• ListHumanTaskUis (p. 562)
• ListHyperParameterTuningJobs (p. 565)
• ListImages (p. 569)
• ListImageVersions (p. 572)
• ListInferenceRecommendationsJobs (p. 575)
• ListLabelingJobs (p. 578)
• ListLabelingJobsForWorkteam (p. 582)
• ListLineageGroups (p. 585)
• ListModelBiasJobDefinitions (p. 588)
• ListModelExplainabilityJobDefinitions (p. 591)
• ListModelMetadata (p. 594)
• ListModelPackageGroups (p. 596)
• ListModelPackages (p. 599)
• ListModelQualityJobDefinitions (p. 603)
• ListModels (p. 606)
• ListMonitoringExecutions (p. 609)
• ListMonitoringSchedules (p. 613)
• ListNotebookInstanceLifecycleConfigs (p. 617)
• ListNotebookInstances (p. 620)
• ListPipelineExecutions (p. 624)
• ListPipelineExecutionSteps (p. 627)
• ListPipelineParametersForExecution (p. 631)
• ListPipelines (p. 634)
• ListProcessingJobs (p. 637)
• ListProjects (p. 640)
• ListStudioLifecycleConfigs (p. 643)
• ListSubscribedWorkteams (p. 647)
• ListTags (p. 649)
• ListTrainingJobs (p. 651)
• ListTrainingJobsForHyperParameterTuningJob (p. 655)
• ListTransformJobs (p. 658)
• ListTrialComponents (p. 661)
• ListTrials (p. 665)
• ListUserProfiles (p. 668)
• ListWorkforces (p. 671)
• ListWorkteams (p. 674)
• PutModelPackageGroupPolicy (p. 677)
• QueryLineage (p. 679)
• RegisterDevices (p. 683)
• RenderUiTemplate (p. 685)
• RetryPipelineExecution (p. 688)
• Search (p. 691)
• SendPipelineExecutionStepFailure (p. 713)
• SendPipelineExecutionStepSuccess (p. 715)
• StartMonitoringSchedule (p. 717)
• StartNotebookInstance (p. 719)
• StartPipelineExecution (p. 721)
• StopAutoMLJob (p. 724)
• StopCompilationJob (p. 726)
• StopEdgePackagingJob (p. 728)
• StopHyperParameterTuningJob (p. 729)
• StopInferenceRecommendationsJob (p. 731)
• StopLabelingJob (p. 733)
• StopMonitoringSchedule (p. 735)
• StopNotebookInstance (p. 737)
• StopPipelineExecution (p. 739)
• StopProcessingJob (p. 741)
• StopTrainingJob (p. 743)
• StopTransformJob (p. 745)
- UpdateAction (p. 747)
- UpdateAppImageConfig (p. 750)
- UpdateArtifact (p. 752)
- UpdateCodeRepository (p. 755)
- UpdateContext (p. 757)
- UpdateDeviceFleet (p. 760)
- UpdateDevices (p. 762)
- UpdateDomain (p. 764)
- UpdateEndpoint (p. 767)
- UpdateEndpointWeightsAndCapacities (p. 771)
- UpdateExperiment (p. 773)
- UpdateImage (p. 775)
- UpdateModelPackage (p. 778)
- UpdateMonitoringSchedule (p. 781)
- UpdateNotebookInstance (p. 784)
- UpdateNotebookInstanceLifecycleConfig (p. 789)
- UpdatePipeline (p. 791)
- UpdatePipelineExecution (p. 794)
- UpdateProject (p. 797)
- UpdateTrainingJob (p. 800)
- UpdateTrial (p. 802)
- UpdateTrialComponent (p. 804)
- UpdateUserProfile (p. 808)
- UpdateWorkforce (p. 811)
- UpdateWorkteam (p. 814)
AddAssociation

Service: Amazon SageMaker Service

Creates an association between the source and the destination. A source can be associated with multiple destinations, and a destination can be associated with multiple sources. An association is a lineage tracking entity. For more information, see Amazon SageMaker ML Lineage Tracking.

Request Syntax

```json
{
    "AssociationType": "string",
    "DestinationArn": "string",
    "SourceArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AssociationType (p. 14)**

The type of association. The following are suggested uses for each type. Amazon SageMaker places no restrictions on their use.

- **ContributedTo** - The source contributed to the destination or had a part in enabling the destination. For example, the training data contributed to the training job.
- **AssociatedWith** - The source is connected to the destination. For example, an approval workflow is associated with a model deployment.
- **DerivedFrom** - The destination is a modification of the source. For example, a digest output of a channel input for a processing job is derived from the original inputs.
- **Produced** - The source generated the destination. For example, a training job produced a model artifact.

Type: String

Valid Values: ContributedTo | AssociatedWith | DerivedFrom | Produced

Required: No

**DestinationArn (p. 14)**

The Amazon Resource Name (ARN) of the destination.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

Required: Yes

**SourceArn (p. 14)**

The ARN of the source.

Type: String
AddAssociation

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

Required: Yes

Response Syntax

```json
{
    "DestinationArn": "string",
    "SourceArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**DestinationArn (p. 15)**

The Amazon Resource Name (ARN) of the destination.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

**SourceArn (p. 15)**

The ARN of the source.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
AddTags
Service: Amazon SageMaker Service

Adds or overwrites one or more tags for the specified SageMaker resource. You can add tags to notebook instances, training jobs, hyperparameter tuning jobs, batch transform jobs, models, labeling jobs, work teams, endpoint configurations, and endpoints.

Each tag consists of a key and an optional value. Tag keys must be unique per resource. For more information about tags, see For more information, see AWS Tagging Strategies.

Note
Tags that you add to a hyperparameter tuning job by calling this API are also added to any training jobs that the hyperparameter tuning job launches after you call this API, but not to training jobs that the hyperparameter tuning job launched before you called this API. To make sure that the tags associated with a hyperparameter tuning job are also added to all training jobs that the hyperparameter tuning job launches, add the tags when you first create the tuning job by specifying them in the Tags parameter of CreateHyperParameterTuningJob (p. 90)

Note
Tags that you add to a SageMaker Studio Domain or User Profile by calling this API are also added to any Apps that the Domain or User Profile launches after you call this API, but not to Apps that the Domain or User Profile launched before you called this API. To make sure that the tags associated with a Domain or User Profile are also added to all Apps that the Domain or User Profile launches, add the tags when you first create the Domain or User Profile by specifying them in the Tags parameter of CreateDomain (p. 60) or CreateUserProfile (p. 195).

Request Syntax

```
{
   "ResourceArn": "string",
   "Tags": [
   {
      "Key": "string",
      "Value": "string"
   }
   ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ResourceArn (p. 17)

The Amazon Resource Name (ARN) of the resource that you want to tag.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:.+

Required: Yes

Tags (p. 17)

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.
AddTags

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: Yes

Response Syntax

```json
{
   "Tags": [
      {
         "Key": "string",
         "Value": "string"
      }
   ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Tags (p. 18)

A list of tags associated with the SageMaker resource.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
**AssociateTrialComponent**

*Service: Amazon SageMaker Service*

Associates a trial component with a trial. A trial component can be associated with multiple trials. To disassociate a trial component from a trial, call the [DisassociateTrialComponent](#) API.

**Request Syntax**

```json
{
    "TrialComponentName": "string",
    "TrialName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**TrialComponentName (p. 19)**

The name of the component to associated with the trial.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}`

Required: Yes

**TrialName (p. 19)**

The name of the trial to associate with.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}`

Required: Yes

**Response Syntax**

```json
{
    "TrialArn": "string",
    "TrialComponentArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
**TrialArn (p. 19)**

The Amazon Resource Name (ARN) of the trial.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial/.*`

**TrialComponentArn (p. 19)**

The ARN of the trial component.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial-component/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFoundException**

Resource being accessed is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
BatchDescribeModelPackage
Service: Amazon SageMaker Service

This action batch describes a list of versioned model packages

Request Syntax

```json
{
    "ModelPackageArnList": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ModelPackageArnList (p. 21)**

The list of Amazon Resource Name (ARN) of the model package groups.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package/.*`

Required: Yes

Response Syntax

```json
{
    "BatchDescribeModelPackageErrorMap": {
        "string": {
            "ErrorCode": "string",
            "ErrorResponse": "string"
        }
    },
    "ModelPackageSummaries": {
        "string": {
            "CreationTime": number,
            "InferenceSpecification": {
                "Containers": [
                    {
                        "ContainerHostname": "string",
                        "Environment": {
                            "string": "string"
                        },
                        "Framework": "string",
                        "FrameworkVersion": "string",
                        "Image": "string",
                        "ImageDigest": "string",
                        "ModelDataUrl": "string",
                        "ModelInput": {
                            "DataInputConfig": "string"
                        }
                    }
                ]
            }
        }
    }
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**BatchDescribeModelPackageErrorMap (p. 21)**

A map of the resource and BatchDescribeModelPackageError objects reporting the error associated with describing the model package.

- **Type:** String to **BatchDescribeModelPackageError (p. 926)** object map
- **Key Length Constraints:** Minimum length of 1. Maximum length of 2048.
- **Key Pattern:** `arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:model-package/.*`

**ModelPackageSummaries (p. 21)**

The summaries for the model package versions

- **Type:** String to **BatchDescribeModelPackageSummary (p. 927)** object map
- **Key Length Constraints:** Minimum length of 1. Maximum length of 2048.
- **Key Pattern:** `arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:model-package/.*`

**Errors**

For information about the errors that are common to all actions, see **Common Errors (p. 1472).**

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- **AWS Command Line Interface**
- **AWS SDK for .NET**
- **AWS SDK for C++**
- **AWS SDK for Go**
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateAction
Service: Amazon SageMaker Service

Creates an action. An action is a lineage tracking entity that represents an action or activity. For example, a model deployment or an HPO job. Generally, an action involves at least one input or output artifact. For more information, see Amazon SageMaker ML Lineage Tracking.

Request Syntax

```json
{
    "ActionName": "string",
    "ActionType": "string",
    "Description": "string",
    "MetadataProperties": {
        "CommitId": "string",
        "GeneratedBy": "string",
        "ProjectId": "string",
        "Repository": "string"
    },
    "Properties": {
        "string": "string"
    },
    "Source": {
        "SourceId": "string",
        "SourceType": "string",
        "SourceUri": "string"
    },
    "Status": "string",
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ActionName (p. 24)**

The name of the action. Must be unique to your account in an AWS Region.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: Yes

**ActionType (p. 24)**

The action type.

Type: String
Length Constraints: Maximum length of 256.

Required: Yes

**Description (p. 24)**

The description of the action.

Type: String

Length Constraints: Maximum length of 3072.

Pattern: . *

Required: No

**MetadataProperties (p. 24)**

Metadata properties of the tracking entity, trial, or trial component.

Type: MetadataProperties (p. 1143) object

Required: No

**Properties (p. 24)**

A list of properties to add to the action.

Type: String to string map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: . *

Value Length Constraints: Maximum length of 256.

Value Pattern: . *

Required: No

**Source (p. 24)**

The source type, ID, and URI.

Type: ActionSource (p. 860) object

Required: Yes

**Status (p. 24)**

The status of the action.

Type: String

Valid Values: Unknown | InProgress | Completed | Failed | Stopping | Stopped

Required: No

**Tags (p. 24)**

A list of tags to apply to the action.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

Response Syntax

```
{
   "ActionArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

*ActionArn (p. 26)*

The Amazon Resource Name (ARN) of the action.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:action/.*`

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

*ResourceLimitExceeded*

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateAlgorithm
Service: Amazon SageMaker Service

Create a machine learning algorithm that you can use in SageMaker and list in the AWS Marketplace.

Request Syntax

```json
{
    "AlgorithmDescription": "string",
    "AlgorithmName": "string",
    "CertifyForMarketplace": boolean,
    "InferenceSpecification": {
        "Containers": [
            {
                "ContainerHostname": "string",
                "Environment": {
                    "string": "string"
                },
                "Framework": "string",
                "FrameworkVersion": "string",
                "Image": "string",
                "ImageDigest": "string",
                "ModelDataUrl": "string",
                "ModelInput": {
                    "DataInputConfig": "string"
                },
                "NearestModelName": "string",
                "ProductId": "string"
            }
        ],
        "SupportedContentTypes": [ "string" ],
        "SupportedRealtimeInferenceInstanceTypes": [ "string" ],
        "SupportedResponseMIMETypes": [ "string" ],
        "SupportedTransformInstanceTypes": [ "string" ]
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
    "TrainingSpecification": {
        "MetricDefinitions": [
            {
                "Name": "string",
                "Regex": "string"
            }
        ],
        "SupportedHyperParameters": [
            {
                "DefaultValue": "string",
                "Description": "string",
                "IsRequired": boolean,
                "IsTunable": boolean,
                "Name": "string",
                "Range": {
                    "CategoricalParameterRangeSpecification": {
                        "Values": [ "string" ]
                    },
                    "ContinuousParameterRangeSpecification": {
                        "MaxValue": "string",
                        "MinValue": "string"
                    }
                }
            }
        ]
    }
}
```
"IntegerParameterRangeSpecification": {
  "MaxValue": "string",
  "MinValue": "string"
},
"Type": "string"
},
"SupportedTrainingInstanceTypes": [ "string" ],
"SupportedTuningJobObjectiveMetrics": [ {
  "MetricName": "string",
  "Type": "string"
} ],
"SupportsDistributedTraining": boolean,
"TrainingChannels": [ {
  "Description": "string",
  "IsRequired": boolean,
  "Name": "string",
  "SupportedCompressionTypes": [ "string" ],
  "SupportedContentType": [ "string" ],
  "SupportedInputMode": [ "string" ]
} ],
"TrainingImage": "string",
"TrainingImageDigest": "string"
},
"ValidationSpecification": {
  "ValidationProfiles": [ {
    "ProfileName": "string",
    "TrainingJobDefinition": {
      "HyperParameters": {
        "string": "string"
      },
      "InputDataConfig": [ {
        "ChannelName": "string",
        "CompressionType": "string",
        "ContentType": "string",
        "DataSource": { "FileSystemDataSource": { "DirectoryPath": "string",
          "FileSystemAccessMode": "string",
          "FileSystemId": "string",
          "FileSystemType": "string" },
        "S3DataSource": { "AttributeNames": [ "string" ],
          "S3DataDistributionType": "string",
          "S3DataType": "string",
          "S3Uri": "string" } },
        "InputMode": "string",
        "RecordWrapperType": "string",
        "ShuffleConfig": { "Seed": number
      } }
    },
    "OutputDataConfig": { "KmsKeyId": "string",
      "S3OutputPath": "string"
    }
  } ]
}
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AlgorithmDescription (p. 27)

A description of the algorithm.

Type: String

Length Constraints: Maximum length of 1024.
CreateAlgorithm

Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*

Required: No

**AlgorithmName (p. 27)**

The name of the algorithm.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

**CertifyForMarketplace (p. 27)**

Whether to certify the algorithm so that it can be listed in AWS Marketplace.

Type: Boolean

Required: No

**InferenceSpecification (p. 27)**

Specifies details about inference jobs that the algorithm runs, including the following:
- The Amazon ECR paths of containers that contain the inference code and model artifacts.
- The instance types that the algorithm supports for transform jobs and real-time endpoints used for inference.
- The input and output content formats that the algorithm supports for inference.

Type: InferenceSpecification (p. 1105) object

Required: No

**Tags (p. 27)**

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**TrainingSpecification (p. 27)**

Specifies details about training jobs run by this algorithm, including the following:
- The Amazon ECR path of the container and the version digest of the algorithm.
- The hyperparameters that the algorithm supports.
- The instance types that the algorithm supports for training.
- Whether the algorithm supports distributed training.
- The metrics that the algorithm emits to Amazon CloudWatch.
- Which metrics that the algorithm emits can be used as the objective metric for hyperparameter tuning jobs.
- The input channels that the algorithm supports for training data. For example, an algorithm might support train, validation, and test channels.

Type: TrainingSpecification (p. 1400) object
Required: Yes

**ValidationSpecification (p. 27)**

Specifies configurations for one or more training jobs and that SageMaker runs to test the algorithm's training code and, optionally, one or more batch transform jobs that SageMaker runs to test the algorithm's inference code.

Type: `AlgorithmValidationSpecification (p. 875)` object

Required: No

### Response Syntax

```
{
   "AlgorithmArn": "string"
}
```

### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AlgorithmArn (p. 31)**

The Amazon Resource Name (ARN) of the new algorithm.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:algorithm/.*`

### Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateApp
Service: Amazon SageMaker Service

Creates a running app for the specified UserProfile. Supported apps are JupyterServer and KernelGateway. This operation is automatically invoked by Amazon SageMaker Studio upon access to the associated Domain, and when new kernel configurations are selected by the user. A user may have multiple Apps active simultaneously.

Request Syntax

```json
{
    "AppName": "string",
    "AppType": "string",
    "DomainId": "string",
    "ResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
    "UserProfileName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AppName (p. 32)

The name of the app.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9]{0,62}

Required: Yes

AppType (p. 32)

The type of app. Supported apps are JupyterServer and KernelGateway. TensorBoard is not supported.

Type: String

Valid Values: JupyterServer | KernelGateway

Required: Yes

DomainId (p. 32)

The domain ID.
Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**ResourceSpec (p. 32)**

The instance type and the Amazon Resource Name (ARN) of the SageMaker image created on the instance.

**Note**

The value of `InstanceType` passed as part of the `ResourceSpec` in the `CreateApp` call overrides the value passed as part of the `ResourceSpec` configured for the user profile or the domain. If `InstanceType` is not specified in any of those three `ResourceSpec` values for a `KernelGateway` app, the `CreateApp` call fails with a request validation error.

Type: `ResourceSpec (p. 1341)` object

Required: No

**Tags (p. 32)**

Each tag consists of a key and an optional value. Tag keys must be unique per resource.

Type: Array of `Tag (p. 1377)` objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**UserProfileName (p. 32)**

The user profile name.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

### Response Syntax

```
{
  "AppArn": "string"
}
```

### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AppArn (p. 33)**

The Amazon Resource Name (ARN) of the app.

Type: String

Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:app/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateAppImageConfig
Service: Amazon SageMaker Service

Creates a configuration for running a SageMaker image as a KernelGateway app. The configuration specifies the Amazon Elastic File System (EFS) storage volume on the image, and a list of the kernels in the image.

Request Syntax

```json
{
    "AppImageConfigName": "string",
    "KernelGatewayImageConfig": {
        "FileSystemConfig": {
            "DefaultGid": number,
            "DefaultUid": number,
            "MountPath": "string"
        },
        "KernelSpecs": [
            {
                "DisplayName": "string",
                "Name": "string"
            }
        ],
        "Tags": [
            {
                "Key": "string",
                "Value": "string"
            }
        ]
    }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AppImageConfigName (p. 35)**

The name of the AppImageConfig. Must be unique to your account.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9]){0,62}$

Required: Yes

**KernelGatewayImageConfig (p. 35)**

The KernelGatewayImageConfig.

Type: KernelGatewayImageConfig (p. 1117) object

Required: No

**Tags (p. 35)**

A list of tags to apply to the AppImageConfig.
Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```json
{
  "AppImageConfigArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AppImageConfigArn (p. 36)**

The Amazon Resource Name (ARN) of the AppImageConfig.

Type: String

Length Constraints: Maximum length of 256.

Pattern: Arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:app-image-config/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateArtifact
Service: Amazon SageMaker Service

Creates an *artifact*. An artifact is a lineage tracking entity that represents a URI addressable object or data. Some examples are the S3 URI of a dataset and the ECR registry path of an image. For more information, see Amazon SageMaker ML Lineage Tracking.

Request Syntax

```json
{
  "ArtifactName": "string",
  "ArtifactType": "string",
  "MetadataProperties": {
    "CommitId": "string",
    "GeneratedBy": "string",
    "ProjectId": "string",
    "Repository": "string"
  },
  "Properties": {
    "string": "string"
  },
  "Source": {
    "SourceTypes": [
      {
        "SourceIdType": "string",
        "Value": "string"
      }
    ],
    "SourceUri": "string"
  },
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ArtifactName (p. 37)**

The name of the artifact. Must be unique to your account in an AWS Region.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9]{0,119}$

Required: No

**ArtifactType (p. 37)**

The artifact type.
Type: String
Length Constraints: Maximum length of 256.
Required: Yes

**MetadataProperties (p. 37)**
Metadata properties of the tracking entity, trial, or trial component.
Type: `MetadataProperties (p. 1143)` object
Required: No

**Properties (p. 37)**
A list of properties to add to the artifact.
Type: String to string map
Map Entries: Maximum number of 30 items.
Key Length Constraints: Maximum length of 256.
Key Pattern: .*
Value Length Constraints: Maximum length of 256.
Value Pattern: .*
Required: No

**Source (p. 37)**
The ID, ID type, and URI of the source.
Type: `ArtifactSource (p. 892)` object
Required: Yes

**Tags (p. 37)**
A list of tags to apply to the artifact.
Type: Array of `Tag (p. 1377)` objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

**Response Syntax**
```json
{
   "ArtifactArn": "string"
}
```

**Response Elements**
If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.
ArtifactArn (p. 38)

The Amazon Resource Name (ARN) of the artifact.

Type: String

Length Constraints: Maximum length of 256.

Pattern: \texttt{arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:artifact/.*}

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceLimitExceeded

You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateAutoMLJob
Service: Amazon SageMaker Service

Creates an Autopilot job.

Find the best-performing model after you run an Autopilot job by calling DescribeAutoMLJob (p. 300).

For information about how to use Autopilot, see Automate Model Development with Amazon SageMaker Autopilot.

Request Syntax

```json
{
    "AutoMLJobConfig": {
        "CandidateGenerationConfig": {
            "FeatureSpecificationS3Uri": "string"
        },
        "CompletionCriteria": {
            "MaxAutoMLJobRuntimeInSeconds": number,
            "MaxCandidates": number,
            "MaxRuntimePerTrainingJobInSeconds": number
        },
        "DataSplitConfig": {
            "ValidationFraction": number
        },
        "SecurityConfig": {
            "EnableInterContainerTrafficEncryption": boolean,
            "VolumeKmsKeyId": "string",
            "VpcConfig": {
                "SecurityGroupIds": [ "string" ],
                "Subnets": [ "string" ]
            }
        },
        "AutoMLJobName": "string",
        "AutoMLJobObjective": {
            "MetricName": "string"
        },
        "GenerateCandidateDefinitionsOnly": boolean,
        "InputDataConfig": [ {
            "ChannelType": "string",
            "CompressionType": "string",
            "ContentType": "string",
            "DataSource": { "S3DataSource": { "S3DataType": "string",
            "S3Uri": "string" }
            },
            "TargetAttributeName": "string"
        } ],
        "ModelDeployConfig": {
            "AutoGenerateEndpointName": boolean,
            "EndpointName": "string"
        },
        "OutputDataConfig": {
            "KmsKeyId": "string",
            "S3OutputPath": "string"
        },
        "ProblemType": "string",
        "RoleArn": "string"
    }
}
```
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
]

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AutoMLJobConfig (p. 40)**

A collection of settings used to configure an AutoML job.

Type: AutoMLJobConfig (p. 916) object

Required: No

**AutoMLJobName (p. 40)**

Identifies an Autopilot job. The name must be unique to your account and is case-insensitive.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}$

Required: Yes

**AutoMLJobObjective (p. 40)**

Defines the objective metric used to measure the predictive quality of an AutoML job. You provide an AutoMLJobObjective:MetricName (p. 917) and Autopilot infers whether to minimize or maximize it.

Type: AutoMLJobObjective (p. 917) object

Required: No

**GenerateCandidateDefinitionsOnly (p. 40)**

Generates possible candidates without training the models. A candidate is a combination of data preprocessors, algorithms, and algorithm parameter settings.

Type: Boolean

Required: No

**InputDataConfig (p. 40)**

An array of channel objects that describes the input data and its location. Each channel is a named input source. Similar to InputDataConfig supported by HyperParameterTrainingJobDefinition (p. 1082). Format(s) supported: CSV, Parquet. A minimum of 500 rows is required for the training dataset. There is not a minimum number of rows required for the validation dataset.

Type: Array of AutoMLChannel (p. 908) objects

Array Members: Minimum number of 1 item. Maximum number of 2 items.
CreateAutoMLJob

Required: Yes

**ModelDeployConfig (p. 40)**

Specifies how to generate the endpoint name for an automatic one-click Autopilot model deployment.

Type: ModelDeployConfig (p. 1156) object

Required: No

**OutputDataConfig (p. 40)**

Provides information about encryption and the Amazon S3 output path needed to store artifacts from an AutoML job. Format(s) supported: CSV.

Type: AutoMLOutputDataConfig (p. 921) object

Required: Yes

**ProblemType (p. 40)**

Defines the type of supervised learning available for the candidates. For more information, see Amazon SageMaker Autopilot problem types and algorithm support.

Type: String

Valid Values: BinaryClassification | MulticlassClassification | Regression

Required: No

**RoleArn (p. 40)**

The ARN of the role that is used to access the data.

Type: String


Pattern: `^arn:aws[a-zA-Z-]*:iam::\d{12}:role/\?[a-zA-Z-0-9+=,.@-_/]+$`

Required: Yes

**Tags (p. 40)**

Each tag consists of a key and an optional value. Tag keys must be unique per resource.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```json
{
   "AutoMLJobArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**AutoMLJobArn (p. 42)**

The unique ARN assigned to the AutoML job when it is created.

- **Type:** String
- **Length Constraints:** Minimum length of 1. Maximum length of 256.
- **Pattern:** `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

- Resource being accessed is in use.
  - HTTP Status Code: 400

**ResourceLimitExceeded**

- You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.
  - HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateCodeRepository
Service: Amazon SageMaker Service

Creates a Git repository as a resource in your SageMaker account. You can associate the repository with notebook instances so that you can use Git source control for the notebooks you create. The Git repository is a resource in your SageMaker account, so it can be associated with more than one notebook instance, and it persists independently from the lifecycle of any notebook instances it is associated with.

The repository can be hosted either in AWS CodeCommit or in any other Git repository.

Request Syntax

```json
{
    "CodeRepositoryName": "string",
    "GitConfig": {
        "Branch": "string",
        "RepositoryUrl": "string",
        "SecretArn": "string"
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CodeRepositoryName (p. 44)**

The name of the Git repository. The name must have 1 to 63 characters. Valid characters are a-z, A-Z, 0-9, and - (hyphen).

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

**GitConfig (p. 44)**

Specifies details about the repository, including the URL where the repository is located, the default branch, and credentials to use to access the repository.

Type: `GitConfig (p. 1054)` object

Required: Yes

**Tags (p. 44)**

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.
Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```json
{
   "CodeRepositoryArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CodeRepositoryArn (p. 45)**

The Amazon Resource Name (ARN) of the new repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:code-repository/.*`

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateCompilationJob
Service: Amazon SageMaker Service

Starts a model compilation job. After the model has been compiled, Amazon SageMaker saves the resulting model artifacts to an Amazon Simple Storage Service (Amazon S3) bucket that you specify.

If you choose to host your model using Amazon SageMaker hosting services, you can use the resulting model artifacts as part of the model. You can also use the artifacts with AWS IoT Greengrass. In that case, deploy them as an ML resource.

In the request body, you provide the following:

- A name for the compilation job
- Information about the input model artifacts
- The output location for the compiled model and the device (target) that the model runs on
- The Amazon Resource Name (ARN) of the IAM role that Amazon SageMaker assumes to perform the model compilation job.

You can also provide a Tag to track the model compilation job's resource use and costs. The response body contains the CompilationJobArn for the compiled job.

To stop a model compilation job, use StopCompilationJob (p. 726). To get information about a particular model compilation job, use DescribeCompilationJob (p. 308). To get information about multiple model compilation jobs, use ListCompilationJobs (p. 525).

Request Syntax

```json
{
    "CompilationJobName": "string",
    "InputConfig": {
        "DataInputConfig": "string",
        "Framework": "string",
        "FrameworkVersion": "string",
        "S3Uri": "string"
    },
    "ModelPackageVersionArn": "string",
    "OutputConfig": {
        "CompilerOptions": "string",
        "KmsKeyId": "string",
        "S3OutputLocation": "string",
        "TargetDevice": "string",
        "TargetPlatform": {
            "Accelerator": "string",
            "Arch": "string",
            "Os": "string"
        }
    },
    "RoleArn": "string",
    "StoppingCondition": {
        "MaxRuntimeInSeconds": number,
        "MaxWaitTimeInSeconds": number
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
    "VpcConfig": {
```
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CompilationJobName (p. 46)**

A name for the model compilation job. The name must be unique within the AWS Region and within your AWS account.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

**InputConfig (p. 46)**

Provides information about the location of input model artifacts, the name and shape of the expected data inputs, and the framework in which the model was trained.

Type: InputConfig (p. 1107) object

Required: No

**ModelPackageVersionArn (p. 46)**

The Amazon Resource Name (ARN) of a versioned model package. Provide either a ModelPackageVersionArn or an InputConfig object in the request syntax. The presence of both objects in the CreateCompilationJob request will return an exception.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:model-package/.*

Required: No

**OutputConfig (p. 46)**

Provides information about the output location for the compiled model and the target device the model runs on.

Type: OutputConfig (p. 1240) object

Required: Yes

**RoleArn (p. 46)**

The Amazon Resource Name (ARN) of an IAM role that enables Amazon SageMaker to perform tasks on your behalf.

During model compilation, Amazon SageMaker needs your permission to:
CreateCompilationJob

- Read input data from an S3 bucket
- Write model artifacts to an S3 bucket
- Write logs to Amazon CloudWatch Logs
- Publish metrics to Amazon CloudWatch

You grant permissions for all of these tasks to an IAM role. To pass this role to Amazon SageMaker, the caller of this API must have the `iam:PassRole` permission. For more information, see Amazon SageMaker Roles.

**Type:** String

**Length Constraints:** Minimum length of 20. Maximum length of 2048.

**Pattern:** ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@-_\//]+$

**Required:** Yes

**StoppingCondition (p. 46)**

Specifies a limit to how long a model compilation job can run. When the job reaches the time limit, Amazon SageMaker ends the compilation job. Use this API to cap model training costs.

**Type:** StoppingCondition (p. 1370) object

**Required:** Yes

**Tags (p. 46)**

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

**Type:** Array of Tag (p. 1377) objects

**Array Members:** Minimum number of 0 items. Maximum number of 50 items.

**Required:** No

**VpcConfig (p. 46)**

A VpcConfig (p. 1456) object that specifies the VPC that you want your compilation job to connect to. Control access to your models by configuring the VPC. For more information, see Protect Compilation Jobs by Using an Amazon Virtual Private Cloud.

**Type:** NeoVpcConfig (p. 1221) object

**Required:** No

**Response Syntax**

```json
{
  "CompilationJobArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
CompilationJobArn (p. 48)

If the action is successful, the service sends back an HTTP 200 response. Amazon SageMaker returns the following data in JSON format:

• CompilationJobArn: The Amazon Resource Name (ARN) of the compiled job.

  Type: String
  Length Constraints: Maximum length of 256.

  Pattern: arn:aws[a-z\-*]:sagemaker:[a-z0-9\-*]:[0-9]{12}:compilation-job/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse

  Resource being accessed is in use.

  HTTP Status Code: 400

ResourceLimitExceeded

  You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

  HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreateContext

Service: Amazon SageMaker Service

Creates a context. A context is a lineage tracking entity that represents a logical grouping of other tracking or experiment entities. Some examples are an endpoint and a model package. For more information, see Amazon SageMaker ML Lineage Tracking.

Request Syntax

```
{
  "ContextName": "string",
  "ContextType": "string",
  "Description": "string",
  "Properties": {
    "string": "string"
  },
  "Source": {
    "SourceId": "string",
    "SourceType": "string",
    "SourceUri": "string"
  },
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ContextName (p. 50)**

The name of the context. Must be unique to your account in an AWS Region.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9-][a-zA-Z0-9-]{0,119}$

Required: Yes

**ContextType (p. 50)**

The context type.

Type: String

Length Constraints: Maximum length of 256.

Required: Yes

**Description (p. 50)**

The description of the context.
Type: String
Length Constraints: Maximum length of 3072.
Pattern: . *
Required: No
Properties (p. 50)
A list of properties to add to the context.
Type: String to string map
Map Entries: Maximum number of 30 items.
Key Length Constraints: Maximum length of 256.
Key Pattern: . *
Value Length Constraints: Maximum length of 256.
Value Pattern: . *
Required: No
Source (p. 50)
The source type, ID, and URI.
Type: ContextSource (p. 959) object
Required: Yes
Tags (p. 50)
A list of tags to apply to the context.
Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

Response Syntax

```
{
  "ContextArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

ContextArn (p. 51)

The Amazon Resource Name (ARN) of the context.
Type: String
Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:context/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateDataQualityJobDefinition
Service: Amazon SageMaker Service

Creates a definition for a job that monitors data quality and drift. For information about model monitor, see Amazon SageMaker Model Monitor.

Request Syntax

```json
{
    "DataQualityAppSpecification": {
        "ContainerArguments": ["string"],
        "ContainerEntrypoint": ["string"],
        "Environment": {
            "string": "string"
        },
        "ImageUri": "string",
        "PostAnalyticsProcessorSourceUri": "string",
        "RecordPreprocessorSourceUri": "string"
    },
    "DataQualityBaselineConfig": {
        "BaselineJobName": "string",
        "ConstraintsResource": {
            "S3Uri": "string"
        },
        "StatisticsResource": {
            "S3Uri": "string"
        }
    },
    "DataQualityJobInput": {
        "EndpointInput": {
            "EndpointName": "string",
            "EndTimeOffset": "string",
            "FeaturesAttribute": "string",
            "InferenceAttribute": "string",
            "LocalPath": "string",
            "ProbabilityAttribute": "string",
            "ProbabilityThresholdAttribute": number,
            "S3DataDistributionType": "string",
            "S3InputMode": "string",
            "StartTimeOffset": "string"
        }
    },
    "DataQualityJobOutputConfig": {
        "KmsKeyId": "string",
        "MonitoringOutputs": [
            {
                "S3Output": {
                    "LocalPath": "string",
                    "S3UploadMode": "string",
                    "S3Uri": "string"
                }
            }
        ],
        "JobDefinitionName": "string",
        "JobResources": {
            "ClusterConfig": {
                "InstanceCount": number,
                "InstanceType": "string",
                "VolumeKmsKeyId": "string",
                "VolumeSizeInGB": number
            }
        }
    }
}
```
```json
"NetworkConfig": {
  "EnableInterContainerTrafficEncryption": boolean,
  "EnableNetworkIsolation": boolean,
  "VpcConfig": {
    "SecurityGroupIds": [ "string" ],
    "Subnets": [ "string" ]
  }
},
"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeInSeconds": number
},
"Tags": [
  { 
    "Key": "string",
    "Value": "string"
  }
]
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DataQualityAppSpecification (p. 53)**

Specifies the container that runs the monitoring job.

Type: DataQualityAppSpecification (p. 973) object

Required: Yes

**DataQualityBaselineConfig (p. 53)**

Configures the constraints and baselines for the monitoring job.

Type: DataQualityBaselineConfig (p. 975) object

Required: No

**DataQualityJobInput (p. 53)**

A list of inputs for the monitoring job. Currently endpoints are supported as monitoring inputs.

Type: DataQualityJobInput (p. 976) object

Required: Yes

**DataQualityJobOutputConfig (p. 53)**

The output configuration for monitoring jobs.

Type: MonitoringOutputConfig (p. 1209) object

Required: Yes

**JobDefinitionName (p. 53)**

The name for the monitoring job definition.

Type: String

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])+$

Required: Yes

**JobResources (p. 53)**

Identifies the resources to deploy for a monitoring job.

Type: MonitoringResources (p. 1210) object

Required: Yes

**NetworkConfig (p. 53)**

Specifies networking configuration for the monitoring job.

Type: MonitoringNetworkConfig (p. 1207) object

Required: No

**RoleArn (p. 53)**

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/?[a-zA-Z0-9_\-\(\)+,\@\$\:\/]+$

Required: Yes

**StoppingCondition (p. 53)**

A time limit for how long the monitoring job is allowed to run before stopping.

Type: MonitoringStoppingCondition (p. 1219) object

Required: No

**Tags (p. 53)**

(Optional) An array of key-value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```json
{
   "JobDefinitionArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**JobDefinitionArn (p. 55)**

The Amazon Resource Name (ARN) of the job definition.

Type: String

Length Constraints: Maximum length of 256.

Pattern: . *

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateDeviceFleet
Service: Amazon SageMaker Service

Creates a device fleet.

Request Syntax

```
{
  "Description": "string",
  "DeviceFleetName": "string",
  "EnableIotRoleAlias": boolean,
  "OutputConfig": {
    "KmsKeyId": "string",
    "PresetDeploymentConfig": "string",
    "PresetDeploymentType": "string",
    "S3OutputLocation": "string"
  },
  "RoleArn": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

Description (p. 57)
A description of the fleet.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 800.
Pattern: \[\S\s\]+
Required: No

DeviceFleetName (p. 57)
The name of the fleet that the device belongs to.
Type: String
Pattern: ^[a-zA-Z0-9]+(-*[a-zA-Z0-9]*)\{0,62}\$
Required: Yes

EnableIotRoleAlias (p. 57)
Whether to create an AWS IoT Role Alias during device fleet creation. The name of the role alias generated will match this pattern: “SageMakerEdge-{DeviceFleetName}”.

57
For example, if your device fleet is called "demo-fleet", the name of the role alias will be "SageMakerEdge-demo-fleet".

Type: Boolean
Required: No

**OutputConfig (p. 57)**

The output configuration for storing sample data collected by the fleet.

Type: EdgeOutputConfig (p. 1008) object
Required: Yes

**RoleArn (p. 57)**

The Amazon Resource Name (ARN) that has access to AWS Internet of Things (IoT).

Type: String
Pattern: ^arn:aws[a-z-]*:iam::\d{12}:role/\[a-zA-Z_0-9+=,.@\-_\/]+$  
Required: No

**Tags (p. 57)**

Creates tags for the specified fleet.

Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateDomain
Service: Amazon SageMaker Service

Creates a domain used by Amazon SageMaker Studio. A domain consists of an associated Amazon Elastic File System (EFS) volume, a list of authorized users, and a variety of security, application, policy, and Amazon Virtual Private Cloud (VPC) configurations. An AWS account is limited to one domain per region. Users within a domain can share notebook files and other artifacts with each other.

EFS storage

When a domain is created, an EFS volume is created for use by all of the users within the domain. Each user receives a private home directory within the EFS volume for notebooks, Git repositories, and data files.

SageMaker uses the AWS Key Management Service (AWS KMS) to encrypt the EFS volume attached to the domain with an AWS managed key by default. For more control, you can specify a customer managed key. For more information, see Protect Data at Rest Using Encryption.

VPC configuration

All SageMaker Studio traffic between the domain and the EFS volume is through the specified VPC and subnets. For other Studio traffic, you can specify the AppNetworkAccessType parameter. AppNetworkAccessType corresponds to the network access type that you choose when you onboard to Studio. The following options are available:

- PublicInternetOnly - Non-EFS traffic goes through a VPC managed by Amazon SageMaker, which allows internet access. This is the default value.
- VpcOnly - All Studio traffic is through the specified VPC and subnets. Internet access is disabled by default. To allow internet access, you must specify a NAT gateway.

    When internet access is disabled, you won't be able to run a Studio notebook or to train or host models unless your VPC has an interface endpoint to the SageMaker API and runtime or a NAT gateway and your security groups allow outbound connections.

Important

NFS traffic over TCP on port 2049 needs to be allowed in both inbound and outbound rules in order to launch a SageMaker Studio app successfully.

For more information, see Connect SageMaker Studio Notebooks to Resources in a VPC.

Request Syntax

```json
{
    "AppNetworkAccessType": "string",
    "AppSecurityGroupManagement": "string",
    "AuthMode": "string",
    "DefaultUserSettings": {
        "ExecutionRole": "string",
        "JupyterServerAppSettings": {
            "DefaultResourceSpec": {
                "InstanceType": "string",
                "LifecycleConfigArn": "string",
                "SageMakerImageArn": "string",
                "SageMakerImageVersionArn": "string"
            },
            "LifecycleConfigArns": [ "string" ]
        }
    },
    "KernelGatewayAppSettings": {
```
"CustomImages": [
  {
    "AppImageConfigName": "string",
    "ImageName": "string",
    "ImageVersionNumber": number
  }
],
"DefaultResourceSpec": {
  "InstanceType": "string",
  "LifecycleConfigArn": "string",
  "SageMakerImageArn": "string",
  "SageMakerImageVersionArn": "string"
},
"LifecycleConfigArns": [ "string" ],
"RSessionAppSettings": {
  "CustomImages": [
    {
      "AppImageConfigName": "string",
      "ImageName": "string",
      "ImageVersionNumber": number
    }
  ],
  "DefaultResourceSpec": {
    "InstanceType": "string",
    "LifecycleConfigArn": "string",
    "SageMakerImageArn": "string",
    "SageMakerImageVersionArn": "string"
  }
},
"RStudioServerProAppSettings": {
  "AccessStatus": "string",
  "UserGroup": "string"
},
"SecurityGroups": [ "string" ],
"SharingSettings": {
  "NotebookOutputOption": "string",
  "S3KmsKeyId": "string",
  "S3OutputPath": "string"
},
"TensorBoardAppSettings": {
  "DefaultResourceSpec": {
    "InstanceType": "string",
    "LifecycleConfigArn": "string",
    "SageMakerImageArn": "string",
    "SageMakerImageVersionArn": "string"
  }
}
],
"DomainName": "string",
"DomainSettings": {
  "RStudioServerProDomainSettings": {
    "DefaultResourceSpec": {
      "InstanceType": "string",
      "LifecycleConfigArn": "string",
      "SageMakerImageArn": "string",
      "SageMakerImageVersionArn": "string"
    },
    "DomainExecutionRoleArn": "string",
    "RStudioConnectUrl": "string",
    "RStudioPackageManagerUrl": "string"
  }
},
"SecurityGroupIds": [ "string" ],
"HomeEfsFileSystemKmsKeyId": "string",
"KmsKeyId": "string"
"SubnetIds": [ "string" ],
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
],
"VpcId": "string"}

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AppNetworkAccessType (p. 60)

Specifies the VPC used for non-EFS traffic. The default value is PublicInternetOnly.

- PublicInternetOnly - Non-EFS traffic is through a VPC managed by Amazon SageMaker, which allows direct internet access
- VpcOnly - All Studio traffic is through the specified VPC and subnets

Type: String

Valid Values: PublicInternetOnly | VpcOnly

Required: No

AppSecurityGroupManagement (p. 60)

The entity that creates and manages the required security groups for inter-app communication in VPCOnly mode. Required when CreateDomain.AppNetworkAccessType is VPCOnly and DomainSettings.RStudioServerProDomainSettings.DomainExecutionRoleArn is provided.

Type: String

Valid Values: Service | Customer

Required: No

AuthMode (p. 60)

The mode of authentication that members use to access the domain.

Type: String

Valid Values: SSO | IAM

Required: Yes

DefaultUserSettings (p. 60)

The default settings to use to create a user profile when UserSettings isn't specified in the call to the CreateUserProfile API.

SecurityGroups is aggregated when specified in both calls. For all other settings in UserSettings, the values specified in CreateUserProfile take precedence over those specified in CreateDomain.

Type: UserSettings (p. 1452) object
Required: Yes

**DomainName (p. 60)**

A name for the domain.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](\-\*[a-zA-Z0-9])\{0,62\}

Required: Yes

**DomainSettings (p. 60)**

A collection of domain settings.

Type: `DomainSettings (p. 996)` object

Required: No

**HomeEfsFileSystemKmsKeyId (p. 60)**

*This parameter has been deprecated.*

Use `KmsKeyId`.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

**KmsKeyId (p. 60)**

SageMaker uses AWS KMS to encrypt the EFS volume attached to the domain with an AWS managed key by default. For more control, specify a customer managed key.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

**SubnetIds (p. 60)**

The VPC subnets that Studio uses for communication.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 16 items.

Length Constraints: Maximum length of 32.

Pattern: \([-0-9a-zA-Z]+\]

Required: Yes

**Tags (p. 60)**

Tags to associated with the Domain. Each tag consists of a key and an optional value. Tag keys must be unique per resource. Tags are searchable using the `Search API`.
Tags that you specify for the Domain are also added to all Apps that the Domain launches.

**Type**: Array of [Tag](p. 1377) objects

**Array Members**: Minimum number of 0 items. Maximum number of 50 items.

**Required**: No

**VpcId (p. 60)**

The ID of the Amazon Virtual Private Cloud (VPC) that Studio uses for communication.

**Type**: String

**Length Constraints**: Maximum length of 32.

**Pattern**: [-0-9a-zA-Z]+

**Required**: Yes

### Response Syntax

```json
{
   "DomainArn": "string",
   "Url": "string"
}
```

### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**DomainArn (p. 64)**

The Amazon Resource Name (ARN) of the created domain.

**Type**: String

**Length Constraints**: Maximum length of 256.

**Pattern**: `arn:aws[a-zA-Z\-]*:sagemaker:[a-zA-Z0-9\-\*]:[0-9]{12}:domain/.*`

**Url (p. 64)**

The URL to the created domain.

**Type**: String

**Length Constraints**: Maximum length of 1024.

### Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400
ResourceLimitExceeded

You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateEdgePackagingJob
Service: Amazon SageMaker Service

Starts a SageMaker Edge Manager model packaging job. Edge Manager will use the model artifacts from the Amazon Simple Storage Service bucket that you specify. After the model has been packaged, Amazon SageMaker saves the resulting artifacts to an S3 bucket that you specify.

Request Syntax

```json
{
  "CompilationJobName": "string",
  "EdgePackagingJobName": "string",
  "ModelName": "string",
  "ModelVersion": "string",
  "OutputConfig": {
    "KmsKeyId": "string",
    "PresetDeploymentConfig": "string",
    "PresetDeploymentType": "string",
    "S3OutputLocation": "string"
  },
  "ResourceKey": "string",
  "RoleArn": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CompilationJobName (p. 66)**

The name of the SageMaker Neo compilation job that will be used to locate model artifacts for packaging.

Type: String


Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]{0,62}$

Required: Yes

**EdgePackagingJobName (p. 66)**

The name of the edge packaging job.

Type: String


Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]{0,62}$

Required: Yes
ModelName (p. 66)
The name of the model.
Type: String
Pattern: ^[a-zA-Z0-9-]*[a-zA-Z0-9]\{0,62\}$
Required: Yes

ModelVersion (p. 66)
The version of the model.
Type: String
Pattern: [a-zA-Z0-9\ \_\.]+
Required: Yes

OutputConfig (p. 66)
Provides information about the output location for the packaged model.
Type: EdgeOutputConfig (p. 1008) object
Required: Yes

ResourceKey (p. 66)
The AWS KMS key to use when encrypting the EBS volume the edge packaging job runs on.
Type: String
Length Constraints: Maximum length of 2048.
Pattern: . *
Required: No

RoleArn (p. 66)
The Amazon Resource Name (ARN) of an IAM role that enables Amazon SageMaker to download and upload the model, and to contact SageMaker Neo.
Type: String
Pattern: ^arn:aws[a-zA-Z\-]*:iam:\d{12}:role/?[a-zA-Z0-9\-\=\+,@\-\_\/]+$*
Required: Yes

Tags (p. 66)
Creates tags for the packaging job.
Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No
Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateEndpoint
Service: Amazon SageMaker Service

Creates an endpoint using the endpoint configuration specified in the request. SageMaker uses the endpoint to provision resources and deploy models. You create the endpoint configuration with the CreateEndpointConfig (p. 73) API.

Use this API to deploy models using SageMaker hosting services.

For an example that calls this method when deploying a model to SageMaker hosting services, see the Create Endpoint example notebook.

**Note**
You must not delete an EndpointConfig that is in use by an endpoint that is live or while the UpdateEndpoint or CreateEndpoint operations are being performed on the endpoint. To update an endpoint, you must create a new EndpointConfig.

The endpoint name must be unique within an AWS Region in your AWS account.

When it receives the request, SageMaker creates the endpoint, launches the resources (ML compute instances), and deploys the model(s) on them.

**Note**
When you call CreateEndpoint (p. 69), a load call is made to DynamoDB to verify that your endpoint configuration exists. When you read data from a DynamoDB table supporting Eventually Consistent Reads, the response might not reflect the results of a recently completed write operation. The response might include some stale data. If the dependent entities are not yet in DynamoDB, this causes a validation error. If you repeat your read request after a short time, the response should return the latest data. So retry logic is recommended to handle these possible issues. We also recommend that customers call DescribeEndpointConfig (p. 343) before calling CreateEndpoint (p. 69) to minimize the potential impact of a DynamoDB eventually consistent read.

When SageMaker receives the request, it sets the endpoint status to Creating. After it creates the endpoint, it sets the status to InService. SageMaker can then process incoming requests for inferences. To check the status of an endpoint, use the DescribeEndpoint (p. 338) API.

If any of the models hosted at this endpoint get model data from an Amazon S3 location, SageMaker uses AWS Security Token Service to download model artifacts from the S3 path you provided. AWS STS is activated in your IAM user account by default. If you previously deactivated AWS STS for a region, you need to reactivate AWS STS for that region. For more information, see Activating and Deactivating AWS STS in an AWS Region in the AWS Identity and Access Management User Guide.

**Note**
To add the IAM role policies for using this API operation, go to the IAM console, and choose Roles in the left navigation pane. Search the IAM role that you want to grant access to use the CreateEndpoint (p. 69) and CreateEndpointConfig (p. 73) API operations, add the following policies to the role.

- Option 1: For a full SageMaker access, search and attach the AmazonSageMakerFullAccess policy.
- Option 2: For granting a limited access to an IAM role, paste the following Action elements manually into the JSON file of the IAM role:

```
"Action": ["sagemaker:CreateEndpoint",
          "sagemaker:CreateEndpointConfig"]

"Resource": [
  "arn:aws:sagemaker:region:account-id:endpoint/endpointName"
]```
CreateEndpoint

"arn:aws:sagemaker:region:account-id:endpoint-config/endpointConfigName"

For more information, see SageMaker API Permissions: Actions, Permissions, and Resources Reference.

Request Syntax

```json
{
  "DeploymentConfig": {
    "AutoRollbackConfiguration": {
      "Alarms": [
        {
          "AlarmName": "string"
        }
      ],
    },
    "BlueGreenUpdatePolicy": {
      "MaximumExecutionTimeoutInSeconds": number,
      "TerminationWaitInSeconds": number,
      "TrafficRoutingConfiguration": {
        "CanarySize": {
          "Type": "string",
          "Value": number
        },
        "LinearStepSize": {
          "Type": "string",
          "Value": number
        },
        "Type": "string",
        "WaitIntervalInSeconds": number
      }
    },
    "EndpointConfigName": "string",
    "EndpointName": "string",
    "Tags": [
      {
        "Key": "string",
        "Value": "string"
      }
    ]
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeploymentConfig (p. 70)**

The deployment configuration for an endpoint, which contains the desired deployment strategy and rollback configurations.

Type: DeploymentConfig (p. 987) object
CreateEndpoint

Required: No

**EndpointConfigName (p. 70)**

The name of an endpoint configuration. For more information, see CreateEndpointConfig (p. 73).

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

Required: Yes

**EndpointName (p. 70)**

The name of the endpoint. The name must be unique within an AWS Region in your AWS account. The name is case-insensitive in CreateEndpoint, but the case is preserved and must be matched in InvokeEndpoint (p. 817).

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

Required: Yes

**Tags (p. 70)**

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```
{
   "EndpointArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**EndpointArn (p. 71)**

The Amazon Resource Name (ARN) of the endpoint.

Type: String


Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:endpoint/.*
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateEndpointConfig
Service: Amazon SageMaker Service

Creates an endpoint configuration that SageMaker hosting services uses to deploy models. In the configuration, you identify one or more models, created using the CreateModel API, to deploy and the resources that you want SageMaker to provision. Then you call the CreateEndpoint (p. 69) API.

**Note**
Use this API if you want to use SageMaker hosting services to deploy models into production.

In the request, you define a ProductionVariant, for each model that you want to deploy. Each ProductionVariant parameter also describes the resources that you want SageMaker to provision. This includes the number and type of ML compute instances to deploy.

If you are hosting multiple models, you also assign a VariantWeight to specify how much traffic you want to allocate to each model. For example, suppose that you want to host two models, A and B, and you assign traffic weight 2 for model A and 1 for model B. SageMaker distributes two-thirds of the traffic to Model A, and one-third to model B.

**Note**
When you call CreateEndpoint (p. 69), a load call is made to DynamoDB to verify that your endpoint configuration exists. When you read data from a DynamoDB table supporting **Eventually Consistent Reads**, the response might not reflect the results of a recently completed write operation. The response might include some stale data. If the dependent entities are not yet in DynamoDB, this causes a validation error. If you repeat your read request after a short time, the response should return the latest data. So retry logic is recommended to handle these possible issues. We also recommend that customers call DescribeEndpointConfig (p. 343) before calling CreateEndpoint (p. 69) to minimize the potential impact of a DynamoDB eventually consistent read.

Request Syntax

```json
{
   "AsyncInferenceConfig": {
      "ClientConfig": {
         "MaxConcurrentInvocationsPerInstance": number
      },
      "OutputConfig": {
         "KmsKeyId": "string",
         "NotificationConfig": {
            "ErrorTopic": "string",
            "SuccessTopic": "string"
         },
         "S3OutputPath": "string"
      }
   },
   "DataCaptureConfig": {
      "CaptureContentTypeHeader": {
         "CsvContentTypes": [ "string" ],
         "JsonContentTypes": [ "string" ]
      },
      "CaptureOptions": [ {
         "CaptureMode": "string"
      } ],
      "DestinationS3Uri": "string",
      "EnableCapture": boolean,
      "InitialSamplingPercentage": number,
      "KmsKeyId": "string"
   }
}
```
"EndpointConfigName": "string",
"KmsKeyId": "string",
"ProductionVariants": [
  {
    "AcceleratorType": "string",
    "CoreDumpConfig": {
      "DestinationS3Uri": "string",
      "KmsKeyId": "string"
    },
    "InitialInstanceCount": number,
    "InitialVariantWeight": number,
    "InstanceType": "string",
    "ModelName": "string",
    "ServerlessConfig": {
      "MaxConcurrency": number,
      "MemorySizeInMB": number
    },
    "VariantName": "string"
  }
],
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
]}

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AsyncInferenceConfig (p. 73)

Specifies configuration for how an endpoint performs asynchronous inference. This is a required field in order for your Endpoint to be invoked using InvokeEndpointAsync.

Type: AsyncInferenceConfig (p. 899) object

Required: No

DataCaptureConfig (p. 73)

Configuration to control how SageMaker captures inference data.

Type: DataCaptureConfig (p. 966) object

Required: No

EndpointConfigName (p. 73)

The name of the endpoint configuration. You specify this name in a CreateEndpoint (p. 69) request.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](\-[a-zA-Z0-9])*\{0,62}$

Required: Yes
KmsKeyId (p. 73)

The Amazon Resource Name (ARN) of a AWS Key Management Service key that SageMaker uses to encrypt data on the storage volume attached to the ML compute instance that hosts the endpoint.

The KmsKeyId can be any of the following formats:

- Key ID: 1234abcd-12ab-34cd-56ef-1234567890ab
- Key ARN: arn:aws:kms:us-west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab
- Alias name: alias/ExampleAlias

The KMS key policy must grant permission to the IAM role that you specify in your CreateEndpoint, UpdateEndpoint requests. For more information, refer to the AWS Key Management Service section Using Key Policies in AWS KMS.

Note

Certain Nitro-based instances include local storage, dependent on the instance type. Local storage volumes are encrypted using a hardware module on the instance. You can't request a KmsKeyId when using an instance type with local storage. If any of the models that you specify in the ProductionVariants parameter use nitro-based instances with local storage, do not specify a value for the KmsKeyId parameter. If you specify a value for KmsKeyId when using any nitro-based instances with local storage, the call to CreateEndpointConfig fails.

For a list of instance types that support local instance storage, see Instance Store Volumes. For more information about local instance storage encryption, see SSD Instance Store Volumes.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

ProductionVariants (p. 73)

An list of ProductionVariant objects, one for each model that you want to host at this endpoint.

Type: Array of ProductionVariant (p. 1293) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Required: Yes

Tags (p. 73)

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

{...}
"EndpointConfigArn": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**EndpointConfigArn (p. 75)**

The Amazon Resource Name (ARN) of the endpoint configuration.

Type: String


Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:endpoint-config/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateExperiment
Service: Amazon SageMaker Service

Creates an SageMaker experiment. An experiment is a collection of trials that are observed, compared and evaluated as a group. A trial is a set of steps, called trial components, that produce a machine learning model.

The goal of an experiment is to determine the components that produce the best model. Multiple trials are performed, each one isolating and measuring the impact of a change to one or more inputs, while keeping the remaining inputs constant.

When you use SageMaker Studio or the SageMaker Python SDK, all experiments, trials, and trial components are automatically tracked, logged, and indexed. When you use the AWS SDK for Python (Boto), you must use the logging APIs provided by the SDK.

You can add tags to experiments, trials, trial components and then use the Search (p. 691) API to search for the tags.

To add a description to an experiment, specify the optional Description parameter. To add a description later, or to change the description, call the UpdateExperiment (p. 773) API.

To get a list of all your experiments, call the ListExperiments (p. 553) API. To view an experiment's properties, call the DescribeExperiment (p. 346) API. To get a list of all the trials associated with an experiment, call the ListTrials (p. 665) API. To create a trial call the CreateTrial (p. 188) API.

Request Syntax

```json
{
  "Description": "string",
  "DisplayName": "string",
  "ExperimentName": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

Description (p. 77)

The description of the experiment.

Type: String

Length Constraints: Maximum length of 3072.

Pattern: .*

Required: No
DisplayName (p. 77)

The name of the experiment as displayed. The name doesn't need to be unique. If you don't specify
DisplayName, the value in ExperimentName is displayed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9][-\[a-zA-Z0-9\]]{0,119}$

Required: No

ExperimentName (p. 77)

The name of the experiment. The name must be unique in your AWS account and is not case-
sensitive.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9][-\[a-zA-Z0-9\]]{0,119}$

Required: Yes

Tags (p. 77)

A list of tags to associate with the experiment. You can use Search (p. 691) API to search on the
tags.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```
{
  "ExperimentArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ExperimentArn (p. 78)

The Amazon Resource Name (ARN) of the experiment.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-zA-Z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:experiment/.*
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateFeatureGroup
Service: Amazon SageMaker Service

Create a new FeatureGroup. A FeatureGroup is a group of Features defined in the FeatureStore to describe a Record.

The FeatureGroup defines the schema and features contained in the FeatureGroup. A FeatureGroup definition is composed of a list of Features, a RecordIdentifierFeatureName, an EventTimeFeatureName and configurations for its OnlineStore and OfflineStore. Check AWS service quotas to see the FeatureGroups quota for your AWS account.

Important
You must include at least one of OnlineStoreConfig and OfflineStoreConfig to create a FeatureGroup.

Request Syntax

```json
{
    "Description": "string",
    "EventTimeFeatureName": "string",
    "FeatureDefinitions": [
        {
            "FeatureName": "string",
            "FeatureType": "string"
        }
    ],
    "FeatureGroupName": "string",
    "OfflineStoreConfig": {
        "DataCatalogConfig": {
            "Catalog": "string",
            "Database": "string",
            "TableName": "string"
        },
        "DisableGlueTableCreation": boolean,
        "S3StorageConfig": {
            "KmsKeyId": "string",
            "ResolvedOutputS3Uri": "string",
            "S3Uri": "string"
        }
    },
    "OnlineStoreConfig": {
        "EnableOnlineStore": boolean,
        "SecurityConfig": {
            "KmsKeyId": "string"
        }
    },
    "RecordIdentifierFeatureName": "string",
    "RoleArn": "string",
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).
The request accepts the following data in JSON format.

**Description (p. 80)**

A free-form description of a FeatureGroup.

Type: String

Length Constraints: Maximum length of 128.

Required: No

**EventTimeFeatureName (p. 80)**

The name of the feature that stores the EventTime of a Record in a FeatureGroup.

An EventTime is a point in time when a new event occurs that corresponds to the creation or update of a Record in a FeatureGroup. All Records in the FeatureGroup must have a corresponding EventTime.

An EventTime can be a String or Fractional.

- Fractional: EventTime feature values must be a Unix timestamp in seconds.
- String: EventTime feature values must be an ISO-8601 string in the format. The following formats are supported: yyyy-MM-dd'T'HH:mm:ssZ and yyyy-MM-dd'T'HH:mm:ss.SSSZ where yyyy, MM, and dd represent the year, month, and day respectively and HH, mm, ss, and if applicable, SSS represent the hour, month, second and milliseconds respectively. 'T' and Z are constants.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9](\[-_]*[a-zA-Z0-9]){0,63}$`

Required: Yes

**FeatureDefinitions (p. 80)**

A list of Feature names and types. Name and Type is compulsory per Feature.

Valid feature FeatureTypes are Integral, Fractional and String.

FeatureNames cannot be any of the following: is_deleted, write_time, api_invocation_time

You can create up to 2,500 FeatureDefinitions per FeatureGroup.

Type: Array of FeatureDefinition (p. 1036) objects

Array Members: Minimum number of 1 item. Maximum number of 2500 items.

Required: Yes

**FeatureGroupName (p. 80)**

The name of the FeatureGroup. The name must be unique within an AWS Region in an AWS account. The name:

- Must start and end with an alphanumeric character.
- Can only contain alphanumeric character and hyphens. Spaces are not allowed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,63}$
Required: Yes

**OfflineStoreConfig (p. 80)**

Use this to configure an OfflineFeatureStore. This parameter allows you to specify:
- The Amazon Simple Storage Service (Amazon S3) location of an OfflineStore.
- A configuration for an AWS Glue or AWS Hive data catalog.
- An KMS encryption key to encrypt the Amazon S3 location used for OfflineStore. If KMS encryption key is not specified, by default we encrypt all data at rest using AWS KMS key. By defining your bucket-level key for SSE, you can reduce AWS KMS requests costs by up to 99 percent.

To learn more about this parameter, see OfflineStoreConfig (p. 1231).

Type: OfflineStoreConfig (p. 1231) object
Required: No

**OnlineStoreConfig (p. 80)**

You can turn the OnlineStore on or off by specifying True for the EnableOnlineStore flag in OnlineStoreConfig; the default value is False.

You can also include an AWS KMS key ID (KMSKeyId) for at-rest encryption of the OnlineStore.

Type: OnlineStoreConfig (p. 1238) object
Required: No

**RecordIdentifierFeatureName (p. 80)**

The name of the Feature whose value uniquely identifies a Record defined in the FeatureStore. Only the latest record per identifier value will be stored in the OnlineStore. RecordIdentifierFeatureName must be one of feature definitions' names.

You use the RecordIdentifierFeatureName to access data in a FeatureStore.

This name:
- Must start and end with an alphanumeric character.
- Can only contains alphanumeric characters, hyphens, underscores. Spaces are not allowed.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 64.
Pattern: ^[a-zA-Z0-9][\-_]*[a-zA-Z0-9]{0,63}$
Required: Yes

**RoleArn (p. 80)**

The Amazon Resource Name (ARN) of the IAM execution role used to persist data into the OfflineStore if an OfflineStoreConfig is provided.

Type: String
Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/\S[a-zA-Z0-9-_]*=\S,\S=\S,\S=\S,\S=\S$ Required: No
**Tags (p. 80)**

Tags used to identify Features in each FeatureGroup.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```json
{
  "FeatureGroupArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**FeatureGroupArn (p. 83)**

The Amazon Resource Name (ARN) of the FeatureGroup. This is a unique identifier for the feature group.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:feature-group/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreateFlowDefinition
Service: Amazon SageMaker Service

Creates a flow definition.

Request Syntax

```json
{
    "FlowDefinitionName": "string",
    "HumanLoopActivationConfig": {
        "HumanLoopActivationConditionsConfig": {
            "HumanLoopActivationConditions": "string"
        }
    },
    "HumanLoopConfig": {
        "HumanTaskUiArn": "string",
        "PublicWorkforceTaskPrice": {
            "AmountInUsd": {
                "Cents": number,
                "Dollars": number,
                "TenthFractionsOfACent": number
            }
        },
        "TaskAvailabilityLifetimeInSeconds": number,
        "TaskCount": number,
        "TaskDescription": "string",
        "TaskKeywords": [ "string" ],
        "TaskTimeLimitInSeconds": number,
        "TaskTitle": "string",
        "WorkteamArn": "string"
    },
    "HumanLoopRequestSource": {
        "AwsManagedHumanLoopRequestSource": "string"
    },
    "OutputConfig": {
        "KmsKeyId": "string",
        "S3OutputPath": "string"
    },
    "RoleArn": "string",
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**FlowDefinitionName (p. 85)**

The name of your flow definition.

Type: String

CreateFlowDefinition

Pattern: \^[a-z0-9](-*[a-z0-9])\{0,62\}

Required: Yes

**HumanLoopActivationConfig (p. 85)**

An object containing information about the events that trigger a human workflow.

Type: HumanLoopActivationConfig (p. 1057) object

Required: No

**HumanLoopConfig (p. 85)**

An object containing information about the tasks the human reviewers will perform.

Type: HumanLoopConfig (p. 1058) object

Required: Yes

**HumanLoopRequestSource (p. 85)**

Container for configuring the source of human task requests. Use to specify if Amazon Rekognition or Amazon Textract is used as an integration source.

Type: HumanLoopRequestSource (p. 1063) object

Required: No

**OutputConfig (p. 85)**

An object containing information about where the human review results will be uploaded.

Type: FlowDefinitionOutputConfig (p. 1051) object

Required: No

**RoleArn (p. 85)**

The Amazon Resource Name (ARN) of the role needed to call other services on your behalf. For example, arn:aws:iam::1234567890:role/service-role/AmazonSageMaker-ExecutionRole-20180111T151298.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::d\{12\}:role/?[a-zA-Z_0-9+=,.@\-_\/]+$

Required: Yes

**Tags (p. 85)**

An array of key-value pairs that contain metadata to help you categorize and organize a flow definition. Each tag consists of a key and a value, both of which you define.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```json
{
}
```
"FlowDefinitionArn": "string"
}

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**FlowDefinitionArn (p. 86)**

The Amazon Resource Name (ARN) of the flow definition you create.

- **Type:** String
- **Length Constraints:** Maximum length of 1024.
- **Pattern:** \(arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:flow-definition/.\)

## Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateHumanTaskUi
Service: Amazon SageMaker Service

Defines the settings you will use for the human review workflow user interface. Reviewers will see a three-panel interface with an instruction area, the item to review, and an input area.

Request Syntax

```
{
    "HumanTaskUiName": "string",
    "Tags": [
    {
        "Key": "string",
        "Value": "string"
    }
    ],
    "UiTemplate": {
        "Content": "string"
    }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

HumanTaskUiName (p. 88)

The name of the user interface you are creating.
Type: String
Pattern: ^[a-z0-9](-*[a-z0-9])*
Required: Yes

Tags (p. 88)

An array of key-value pairs that contain metadata to help you categorize and organize a human review workflow user interface. Each tag consists of a key and a value, both of which you define.
Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

UiTemplate (p. 88)

The Liquid template for the worker user interface.
Type: UiTemplate (p. 1446) object
Required: Yes
Response Syntax

```json
{
    "HumanTaskUiArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**HumanTaskUiArn (p. 89)**

The Amazon Resource Name (ARN) of the human review workflow user interface you create.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:human-task-ui/.*`

Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](https://docs.aws.amazon.com/AWSESDKs/latest/APIReference/).  

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateHyperParameterTuningJob

Service: Amazon SageMaker Service

Starts a hyperparameter tuning job. A hyperparameter tuning job finds the best version of a model by running many training jobs on your dataset using the algorithm you choose and values for hyperparameters within ranges that you specify. It then chooses the hyperparameter values that result in a model that performs the best, as measured by an objective metric that you choose.

Request Syntax

```json
{
    "HyperParameterTuningJobConfig": {
        "HyperParameterTuningJobObjective": {
            "MetricName": "string",
            "Type": "string"
        },
        "ParameterRanges": {
            "CategoricalParameterRanges": [
                {
                    "Name": "string",
                    "Values": [ "string" ]
                }
            ],
            "ContinuousParameterRanges": [
                {
                    "MaxValue": "string",
                    "MinValue": "string",
                    "Name": "string",
                    "ScalingType": "string"
                }
            ],
            "IntegerParameterRanges": [
                {
                    "MaxValue": "string",
                    "MinValue": "string",
                    "Name": "string",
                    "ScalingType": "string"
                }
            ]
        },
        "ResourceLimits": {
            "MaxNumberOfTrainingJobs": number,
            "MaxParallelTrainingJobs": number
        },
        "Strategy": "string",
        "TrainingJobEarlyStoppingType": "string",
        "TuningJobCompletionCriteria": {
            "TargetObjectiveMetricValue": number
        }
    },
    "HyperParameterTuningJobName": "string",
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
    "TrainingJobDefinition": {
        "AlgorithmSpecification": {
            "AlgorithmName": "string",
            "MetricDefinitions": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    }
}
```
"Regex": "string",
"TrainingImage": "string",
"TrainingInputMode": "string",

"CheckpointConfig": {
  "LocalPath": "string",
  "S3Uri": "string"
},

"DefinitionName": "string",

"EnableInterContainerTrafficEncryption": boolean,
"EnableManagedSpotTraining": boolean,
"EnableNetworkIsolation": boolean,

"HyperParameterRanges": {
  "CategoricalParameterRanges": [
    {
      "Name": "string",
      "Values": [ "string" ]
    }
  ],

  "ContinuousParameterRanges": [
    {
      "MaxValue": "string",
      "MinValue": "string",
      "Name": "string",
      "ScalingType": "string"
    }
  ],

  "IntegerParameterRanges": [
    {
      "MaxValue": "string",
      "MinValue": "string",
      "Name": "string",
      "ScalingType": "string"
    }
  ]
},

"InputDataConfig": [
  {
    "ChannelName": "string",
    "CompressionType": "string",
    "ContentType": "string",
    "DataSource": {
      "FileSystemDataSource": {
        "DirectoryPath": "string",
        "FileSystemAccessMode": "string",
        "FileSystemId": "string",
        "FileSystemType": "string"
      },

      "S3DataSource": {
        "AttributeNames": [ "string" ],
        "S3DataDistributionType": "string",
        "S3DataType": "string",
        "S3Uri": "string"
      }
    },

    "InputMode": "string",
    "RecordWrapperType": "string",
    "ShuffleConfig": {
      "Seed": number
    }
  }
],

"OutputDataConfig": {
  "KmsKeyId": "string",
  "S3OutputMode": "string",
  "S3OutputPath": "string"
}
"S3OutputPath": "string",
"ResourceConfig": {
"InstanceCount": number,
"InstanceType": "string",
"VolumeKmsKeyId": "string",
"VolumeSizeInGB": number
},
"RetryStrategy": {
"MaximumRetryAttempts": number
},
"RoleArn": "string",
"StaticHyperParameters": {
"string": "string"
},
"StoppingCondition": {
"MaxRuntimeInSeconds": number,
"MaxWaitTimeInSeconds": number
},
"TuningObjective": {
"MetricName": "string",
"Type": "string"
},
"VpcConfig": {
"SecurityGroupIds": [ "string" ],
"Subnets": [ "string" ]
},
"TrainingJobDefinitions": [
{
"AlgorithmSpecification": {
"AlgorithmName": "string",
"MetricDefinitions": [
{ "Name": "string",
"Regex": "string"
}
],
"TrainingImage": "string",
"TrainingInputMode": "string"
},
"CheckpointConfig": {
"LocalPath": "string",
"S3Uri": "string"
},
"DefinitionName": "string",
"EnableInterContainerTrafficEncryption": boolean,
"EnableManagedSpotTraining": boolean,
"EnableNetworkIsolation": boolean,
"HyperParameterRanges": {
"CategoricalParameterRanges": [
{ "Name": "string",
"Values": [ "string" ]
}
],
"ContinuousParameterRanges": [
{ "MaxValue": "string",
"MinValue": "string",
"Name": "string",
"ScalingType": "string"
}
],
"IntegerParameterRanges": [

"MaxValue": "string",
"MinValue": "string",
"Name": "string",
"ScalingType": "string"
]
}
"InputDataConfig": [
{
"ChannelName": "string",
"CompressionType": "string",
"ContentType": "string",
"DataSource": {
"FileSystemDataSource": {
"DirectoryPath": "string",
"FileSystemAccessMode": "string",
"FileSystemId": "string",
"FileSystemType": "string"
},
"S3DataSource": {
"AttributeNames": [ "string" ],
"S3DataDistributionType": "string",
"S3DataType": "string",
"S3Uri": "string"
}
},
"InputMode": "string",
"RecordWrapperType": "string",
"ShuffleConfig": {
"Seed": number
}
}
],
"OutputDataConfig": {
"KmsKeyId": "string",
"S3OutputPath": "string"
},
"ResourceConfig": {
"InstanceCount": number,
"InstanceType": "string",
"VolumeKmsKeyId": "string",
"VolumeSizeInGB": number
},
"RetryStrategy": {
"MaximumRetryAttempts": number
},
"RoleArn": "string",
"StaticHyperParameters": {
"string": "string"
},
"StoppingCondition": {
"MaxRuntimeInSeconds": number,
"MaxWaitTimeInSeconds": number
},
"TuningObjective": {
"MetricName": "string",
"Type": "string"
},
"VpcConfig": {
"SecurityGroupIds": [ "string" ],
"Subnets": [ "string" ]
}
],
"WarmStartConfig": {
"ParentHyperParameterTuningJobs": [
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

HyperParameterTuningJobConfig (p. 90)

The HyperParameterTuningJobConfig (p. 1089) object that describes the tuning job, including the search strategy, the objective metric used to evaluate training jobs, ranges of parameters to search, and resource limits for the tuning job. For more information, see How Hyperparameter Tuning Works.

Type: HyperParameterTuningJobConfig (p. 1089) object

Required: Yes

HyperParameterTuningJobName (p. 90)

The name of the tuning job. This name is the prefix for the names of all training jobs that this tuning job launches. The name must be unique within the same AWS account and AWS Region. The name must have 1 to 32 characters. Valid characters are a-z, A-Z, 0-9, and : + = @ _ % - (hyphen). The name is not case sensitive.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}$

Required: Yes

Tags (p. 90)

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Tags that you specify for the tuning job are also added to all training jobs that the tuning job launches.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

TrainingJobDefinition (p. 90)

The HyperParameterTrainingJobDefinition (p. 1082) object that describes the training jobs that this tuning job launches, including static hyperparameters, input data configuration, output data configuration, resource configuration, and stopping condition.

Type: HyperParameterTrainingJobDefinition (p. 1082) object
CreateHyperParameterTuningJob

**Required:** No

**TrainingJobDefinitions (p. 90)**

A list of the HyperParameterTrainingJobDefinition (p. 1082) objects launched for this tuning job.

Type: Array of HyperParameterTrainingJobDefinition (p. 1082) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

**Required:** No

**WarmStartConfig (p. 90)**

Specifies the configuration for starting the hyperparameter tuning job using one or more previous tuning jobs as a starting point. The results of previous tuning jobs are used to inform which combinations of hyperparameters to search over in the new tuning job.

All training jobs launched by the new hyperparameter tuning job are evaluated by using the objective metric. If you specify IDENTICAL_DATA_AND_ALGORITHM as the WarmStartType value for the warm start configuration, the training job that performs the best in the new tuning job is compared to the best training jobs from the parent tuning jobs. From these, the training job that performs the best as measured by the objective metric is returned as the overall best training job.

**Note**

All training jobs launched by parent hyperparameter tuning jobs and the new hyperparameter tuning jobs count against the limit of training jobs for the tuning job.

Type: HyperParameterTuningJobWarmStartConfig (p. 1094) object

**Required:** No

**Response Syntax**

```
{
  "HyperParameterTuningJobArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**HyperParameterTuningJobArn (p. 95)**

The Amazon Resource Name (ARN) of the tuning job. SageMaker assigns an ARN to a hyperparameter tuning job when you create it.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:hyper-parameter-tuning-job/.*

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateImage
Service: Amazon SageMaker Service

Creates a custom SageMaker image. A SageMaker image is a set of image versions. Each image version represents a container image stored in Amazon Elastic Container Registry (ECR). For more information, see Bring your own SageMaker image.

Request Syntax

```
{
   "Description": "string",
   "DisplayName": "string",
   "ImageName": "string",
   "RoleArn": "string",
   "Tags": [
      {
         "Key": "string",
         "Value": "string"
      }
   ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**Description (p. 97)**

The description of the image.

Type: String


Pattern: .*

Required: No

**DisplayName (p. 97)**

The display name of the image. If not provided, ImageName is displayed.

Type: String


Pattern: ^\$\S\(*\S+$

Required: No

**ImageName (p. 97)**

The name of the image. Must be unique to your account.

Type: String

CreateImage

Pattern: ^[a-zA-Z0-9][-.]?[a-zA-Z0-9]{0,62}$
Required: Yes

RoleArn (p. 97)
The Amazon Resource Name (ARN) of an IAM role that enables Amazon SageMaker to perform tasks on your behalf.
Type: String
Pattern: ^arn:aws[-\[\w\]+]*:iam::\d{12}:role/?[a-zA-Z\-\_0-9\+=,.@\-_]+\+$
Required: Yes

Tags (p. 97)
A list of tags to apply to the image.
Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

Response Syntax

```json
{
    "ImageArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

ImageArn (p. 98)
The Amazon Resource Name (ARN) of the image.
Type: String
Length Constraints: Maximum length of 256.
Pattern: ^arn:aws[-[\w]+]*:sagemaker:.+:\[0-9\]{12}:image/[a-zA-Z0-9][-.]?[a-zA-Z0-9]+$*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse
Resource being accessed is in use.
HTTP Status Code: 400
ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateImageVersion
Service: Amazon SageMaker Service

Creates a version of the SageMaker image specified by ImageName. The version represents the Amazon Elastic Container Registry (ECR) container image specified by BaseImage.

Request Syntax

```json
{
  "BaseImage": "string",
  "ClientToken": "string",
  "ImageName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**BaseImage (p. 100)**

The registry path of the container image to use as the starting point for this version. The path is an Amazon Elastic Container Registry (ECR) URI in the following format:

```
<acct-id>.dkr.ecr.<region>.amazonaws.com/<repo-name[:tag] or [@digest]>
```

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .*

Required: Yes

**ClientToken (p. 100)**

A unique ID. If not specified, the AWS CLI and AWS SDKs, such as the SDK for Python (Boto3), add a unique value to the call.

Type: String


Pattern: ^[a-zA-Z0-9-]+$  

Required: Yes

**ImageName (p. 100)**

The ImageName of the Image to create a version of.

Type: String


Pattern: ^[a-zA-Z0-9][-]?(a-zA-Z0-9]{0,62}$

Required: Yes
Response Syntax

```json
{
   "ImageVersionArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ImageVersionArn (p. 101)**

The Amazon Resource Name (ARN) of the image version.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws([-\[\w]+)*:sagemaker:.+:[0-9]{12}:image-version/[a-z0-9]([-.]?[a-z0-9])*/[0-9]+$)

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreateInferenceRecommendationsJob

Service: Amazon SageMaker Service

Starts a recommendation job. You can create either an instance recommendation or load test job.

Request Syntax

```json
{
  "InputConfig": {
    "EndpointConfigurations": [
      {
        "EnvironmentParameterRanges": {
          "CategoricalParameterRanges": [
            {
              "Name": "string",
              "Value": [ "string" ]
            }
          ],
          "InferenceSpecificationName": "string",
          "InstanceType": "string"
        }
      ],
      "JobDurationInSeconds": number,
      "ModelPackageVersionArn": "string",
      "ResourceLimit": {
        "MaxNumberOfTests": number,
        "MaxParallelOfTests": number
      },
      "TrafficPattern": {
        "Phases": [
          {
            "DurationInSeconds": number,
            "InitialNumberOfUsers": number,
            "SpawnRate": number
          }
        ],
        "TrafficType": "string"
      },
      "VolumeKmsKeyId": "string"
    },
    "JobDescription": "string",
    "JobName": "string",
    "JobType": "string",
    "OutputConfig": {
      "CompiledOutputConfig": {
        "S3OutputUri": "string"
      },
      "KmsKeyId": "string"
    },
    "RoleArn": "string",
    "StoppingConditions": {
      "MaxInvocations": number,
      "ModelLatencyThresholds": [
        {
          "Percentile": "string",
          "ValueInMilliseconds": number
        }
      ]
    },
    "Tags": [
      {
        "Key": "string"
      }
    ]
  }
}
```
"Value": "string"
]
}

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**InputConfig (p. 103)**

Provides information about the versioned model package Amazon Resource Name (ARN), the traffic pattern, and endpoint configurations.

Type: RecommendationJobInputConfig (p. 1325) object

Required: Yes

**JobDescription (p. 103)**

Description of the recommendation job.

Type: String

Length Constraints: Maximum length of 128.

Required: No

**JobName (p. 103)**

A name for the recommendation job. The name must be unique within the AWS Region and within your AWS account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9\-\[\]{}\(\)\*\+\|\?\.$\^\`\{\}\|\}\[\]\?]\{0,63}$

Required: Yes

**JobType (p. 103)**

Defines the type of recommendation job. Specify Default to initiate an instance recommendation and Advanced to initiate a load test. If left unspecified, Amazon SageMaker Inference Recommender will run an instance recommendation (DEFAULT) job.

Type: String

Valid Values: Default | Advanced

Required: Yes

**OutputConfig (p. 103)**

Provides information about the output artifacts and the KMS key to use for Amazon S3 server-side encryption.

Type: RecommendationJobOutputConfig (p. 1327) object

Required: No
RoleArn (p. 103)

The Amazon Resource Name (ARN) of an IAM role that enables Amazon SageMaker to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-z\-\]*:iam::\d{12}:role/?[a-zA-Z0-9+,-@\-_/]+$

Required: Yes

StoppingConditions (p. 103)

A set of conditions for stopping a recommendation job. If any of the conditions are met, the job is automatically stopped.

Type: RecommendationJobStoppingConditions (p. 1329) object

Required: No

Tags (p. 103)

The metadata that you apply to AWS resources to help you categorize and organize them. Each tag consists of a key and a value, both of which you define. For more information, see Tagging AWS Resources in the AWS General Reference.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```
{
  "JobArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

JobArn (p. 105)

The Amazon Resource Name (ARN) of the recommendation job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-\]*:sagemaker:[a-z0-9\-\]*:[0-9]\{12\}:inference-recommendations-job/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateLabelingJob

Service: Amazon SageMaker Service

Creates a job that uses workers to label the data objects in your input dataset. You can use the labeled data to train machine learning models.

You can select your workforce from one of three providers:

- A private workforce that you create. It can include employees, contractors, and outside experts. Use a private workforce when you want the data to stay within your organization or when a specific set of skills is required.
- One or more vendors that you select from the AWS Marketplace. Vendors provide expertise in specific areas.
- The Amazon Mechanical Turk workforce. This is the largest workforce, but it should only be used for public data or data that has been stripped of any personally identifiable information.

You can also use automated data labeling to reduce the number of data objects that need to be labeled by a human. Automated data labeling uses active learning to determine if a data object can be labeled by machine or if it needs to be sent to a human worker. For more information, see Using Automated Data Labeling.

The data objects to be labeled are contained in an Amazon S3 bucket. You create a manifest file that describes the location of each object. For more information, see Using Input and Output Data.

The output can be used as the manifest file for another labeling job or as training data for your machine learning models.

You can use this operation to create a static labeling job or a streaming labeling job. A static labeling job stops if all data objects in the input manifest file identified in ManifestS3Uri have been labeled. A streaming labeling job runs perpetually until it is manually stopped, or remains idle for 10 days. You can send new data objects to an active (InProgress) streaming labeling job in real time. To learn how to create a static labeling job, see Create a Labeling Job (API) in the Amazon SageMaker Developer Guide. To learn how to create a streaming labeling job, see Create a Streaming Labeling Job.

Request Syntax

```
{
    "HumanTaskConfig": {
        "AnnotationConsolidationConfig": {
            "AnnotationConsolidationLambdaArn": "string"
        },
        "MaxConcurrentTaskCount": number,
        "NumberOfHumanWorkersPerDataObject": number,
        "PreHumanTaskLambdaArn": "string",
        "PublicWorkforceTaskPrice": {
            "AmountInUsd": {
                "Cents": number,
                "Dollars": number,
                "TenthFractionsOfACent": number
            }
        },
        "TaskAvailabilityLifetimeInSeconds": number,
        "TaskDescription": "string",
        "TaskKeywords": [ "string" ],
        "TaskTimeLimitInSeconds": number,
        "TaskTitle": "string",
        "UiConfig": {
            "HumanTaskUiArn": "string",
            "UiTemplateS3Uri": "string"
        }
    }
}
```
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**HumanTaskConfig (p. 107)**

Configures the labeling task and how it is presented to workers; including, but not limited to price, keywords, and batch size (task count).

Type: HumanTaskConfig (p. 1064) object

Required: Yes

**InputConfig (p. 107)**

Input data for the labeling job, such as the Amazon S3 location of the data objects and the location of the manifest file that describes the data objects.
You must specify at least one of the following: `S3DataSource` or `SnsDataSource`.

- Use `SnsDataSource` to specify an SNS input topic for a streaming labeling job. If you do not specify and SNS input topic ARN, Ground Truth will create a one-time labeling job that stops after all data objects in the input manifest file have been labeled.
- Use `S3DataSource` to specify an input manifest file for both streaming and one-time labeling jobs. Adding an `S3DataSource` is optional if you use `SnsDataSource` to create a streaming labeling job.

If you use the Amazon Mechanical Turk workforce, your input data should not include confidential information, personal information or protected health information. Use `ContentClassifiers` to specify that your data is free of personally identifiable information and adult content.

Type: `LabelingJobInputConfig (p. 1128)` object

Required: Yes

**LabelAttributeName (p. 107)**

The attribute name to use for the label in the output manifest file. This is the key for the key/value pair formed with the label that a worker assigns to the object. The `LabelAttributeName` must meet the following requirements.

- The name can't end with "-metadata".
- If you are using one of the following built-in task types, the attribute name must end with "-ref". If the task type you are using is not listed below, the attribute name must not end with "-ref".
  - Image semantic segmentation (`SemanticSegmentation`), and adjustment (`AdjustmentSemanticSegmentation`) and verification (`VerificationSemanticSegmentation`) labeling jobs for this task type.
  - Video frame object detection (`VideoObjectDetection`), and adjustment and verification (`AdjustmentVideoObjectDetection`) labeling jobs for this task type.
  - Video frame object tracking (`VideoObjectTracking`), and adjustment and verification (`AdjustmentVideoObjectTracking`) labeling jobs for this task type.
  - 3D point cloud semantic segmentation (`3DPointCloudSemanticSegmentation`), and adjustment and verification (`Adjustment3DPointCloudSemanticSegmentation`) labeling jobs for this task type.
  - 3D point cloud object tracking (`3DPointCloudObjectTracking`), and adjustment and verification (`Adjustment3DPointCloudObjectTracking`) labeling jobs for this task type.

**Important**

If you are creating an adjustment or verification labeling job, you must use a different `LabelAttributeName` than the one used in the original labeling job. The original labeling job is the Ground Truth labeling job that produced the labels that you want verified or adjusted. To learn more about adjustment and verification labeling jobs, see Verify and Adjust Labels.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,126}$`

Required: Yes

**LabelCategoryConfigS3Uri (p. 107)**

The S3 URI of the file, referred to as a label category configuration file, that defines the categories used to label the data objects.
For 3D point cloud and video frame task types, you can add label category attributes and frame attributes to your label category configuration file. To learn how, see Create a Labeling Category Configuration File for 3D Point Cloud Labeling Jobs.

For named entity recognition jobs, in addition to "labels", you must provide worker instructions in the label category configuration file using the "instructions" parameter:

```json
{ "shortInstruction": "<h1>Add header</h1><p>Add Instructions</p>", "fullInstruction": "<p>Add additional instructions.</p>" }
```

For details and an example, see Create a Named Entity Recognition Labeling Job (API).

For all other built-in task types and custom tasks, your label category configuration file must be a JSON file in the following format. Identify the labels you want to use by replacing label_1, label_2,...,label_n with your label categories.

```json
{
  "document-version": "2018-11-28",
  "labels": [{ "label": "label_1" },{ "label": "label_2" },...{ "label": "label_n" }]
}
```

Note the following about the label category configuration file:

- For image classification and text classification (single and multi-label) you must specify at least two label categories. For all other task types, the minimum number of label categories required is one.
- Each label category must be unique, you cannot specify duplicate label categories.
- If you create a 3D point cloud or video frame adjustment or verification labeling job, you must include `auditLabelAttributeName` in the label category configuration. Use this parameter to enter the `LabelAttributeName` of the labeling job you want to adjust or verify annotations of.

**Type:** String

**Length Constraints:** Maximum length of 1024.

**Pattern:** `^(https|s3)://([^/]*)((^[^/]+)?(.*))$`

**Required:** No

**LabelingJobAlgorithmsConfig (p. 107)**

Configures the information required to perform automated data labeling.

**Type:** `LabelingJobAlgorithmsConfig (p. 1122)` object

**Required:** No

**LabelingJobName (p. 107)**

The name of the labeling job. This name is used to identify the job in a list of labeling jobs. Labeling job names must be unique within an AWS account and region. `LabelingJobName` is not case sensitive. For example, Example-job and example-job are considered the same labeling job name by Ground Truth.

**Type:** String

**Length Constraints:** Minimum length of 1. Maximum length of 63.

**Pattern:** `^[a-zA-Z0-9-][a-zA-Z0-9-]{0,62}$`

**Required:** Yes
OutputConfig (p. 107)

The location of the output data and the AWS Key Management Service key ID for the key used to encrypt the output data, if any.

Type: LabelingJobOutputConfig (p. 1130) object

Required: Yes

RoleArn (p. 107)

The Amazon Resource Number (ARN) that Amazon SageMaker assumes to perform tasks on your behalf during data labeling. You must grant this role the necessary permissions so that Amazon SageMaker can successfully complete data labeling.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9+=,.@\-_\//]+$

Required: Yes

StoppingConditions (p. 107)

A set of conditions for stopping the labeling job. If any of the conditions are met, the job is automatically stopped. You can use these conditions to control the cost of data labeling.

Type: LabelingJobStoppingConditions (p. 1135) object

Required: No

Tags (p. 107)

An array of key/value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```json
{
  "LabelingJobArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

LabelingJobArn (p. 111)

The Amazon Resource Name (ARN) of the labeling job. You use this ARN to identify the labeling job.

Type: String
Length Constraints: Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:labeling-job/.*`

## Errors

For information about the errors that are common to all actions, see [Common Errors](p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateModel

Service: Amazon SageMaker Service

Creates a model in SageMaker. In the request, you name the model and describe a primary container. For the primary container, you specify the Docker image that contains inference code, artifacts (from prior training), and a custom environment map that the inference code uses when you deploy the model for predictions.

Use this API to create a model if you want to use SageMaker hosting services or run a batch transform job.

To host your model, you create an endpoint configuration with the `CreateEndpointConfig` API, and then create an endpoint with the `CreateEndpoint` API. SageMaker then deploys all of the containers that you defined for the model in the hosting environment.

For an example that calls this method when deploying a model to SageMaker hosting services, see Create a Model (AWS SDK for Python (Boto 3)).

To run a batch transform using your model, you start a job with the `CreateTransformJob` API. SageMaker uses your model and your dataset to get inferences which are then saved to a specified S3 location.

In the request, you also provide an IAM role that SageMaker can assume to access model artifacts and docker image for deployment on ML compute hosting instances or for batch transform jobs. In addition, you also use the IAM role to manage permissions the inference code needs. For example, if the inference code access any other AWS resources, you grant necessary permissions via this role.

Request Syntax

```json
{
    "Containers": [
        {
            "ContainerHostname": "string",
            "Environment": {
                "string": "string"
            },
            "Image": "string",
            "ImageConfig": {
                "RepositoryAccessMode": "string",
                "RepositoryAuthConfig": {
                    "RepositoryCredentialsProviderArn": "string"
                }
            },
            "InferenceSpecificationName": "string",
            "Mode": "string",
            "ModelDataUrl": "string",
            "ModelPackageName": "string",
            "MultiModelConfig": {
                "ModelCacheSetting": "string"
            }
        }
    ],
    "EnableNetworkIsolation": boolean,
    "ExecutionRoleArn": "string",
    "InferenceExecutionConfig": {
        "Mode": "string"
    },
    "ModelName": "string",
    "PrimaryContainer": {
        "ContainerHostname": "string",
        "Environment": {
```

113
"string" : "string",
"Image": "string",
"ImageConfig": {
  "RepositoryAccessMode": "string",
  "RepositoryAuthConfig": {
    "RepositoryCredentialsProviderArn": "string"
  }
},
"InferenceSpecificationName": "string",
"Mode": "string",
"ModelDataUrl": "string",
"ModelPackageName": "string",
"MultiModelConfig": {
  "ModelCacheSetting": "string"
}
},
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
],
"VpcConfig": {
  "SecurityGroupIds": [ "string" ],
  "Subnets": [ "string" ]
}
}

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

Containers (p. 113)

Specifies the containers in the inference pipeline.

Type: Array of ContainerDefinition (p. 956) objects

Array Members: Maximum number of 15 items.

Required: No

EnableNetworkIsolation (p. 113)

Isolates the model container. No inbound or outbound network calls can be made to or from the model container.

Type: Boolean

Required: No

ExecutionRoleArn (p. 113)

The Amazon Resource Name (ARN) of the IAM role that SageMaker can assume to access model artifacts and docker image for deployment on ML compute instances or for batch transform jobs. Deploying on ML compute instances is part of model hosting. For more information, see SageMaker Roles.

Note

To be able to pass this role to SageMaker, the caller of this API must have the iam:PassRole permission.
CreateModel

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9_\-+=,.@\-/]+$  

Required: Yes

**InferenceExecutionConfig (p. 113)**

Specifies details of how containers in a multi-container endpoint are called.

Type: InferenceExecutionConfig (p. 1101) object

Required: No

**ModelName (p. 113)**

The name of the new model.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-]*(\-*[a-zA-Z0-9]*)*$

Required: Yes

**PrimaryContainer (p. 113)**

The location of the primary docker image containing inference code, associated artifacts, and custom environment map that the inference code uses when the model is deployed for predictions.

Type: ContainerDefinition (p. 956) object

Required: No

**Tags (p. 113)**

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**VpcConfig (p. 113)**

A VpcConfig (p. 1456) object that specifies the VPC that you want your model to connect to. Control access to and from your model container by configuring the VPC. VpcConfig is used in hosting services and in batch transform. For more information, see Protect Endpoints by Using an Amazon Virtual Private Cloud and Protect Data in Batch Transform Jobs by Using an Amazon Virtual Private Cloud.

Type: VpcConfig (p. 1456) object

Required: No

**Response Syntax**

```json
{
}
```
"ModelArn": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ModelArn (p. 115)**

The ARN of the model created in SageMaker.

Type: String


Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateModelBiasJobDefinition

Service: Amazon SageMaker Service

Creates the definition for a model bias job.

Request Syntax

```json
{
    "JobDefinitionName": "string",
    "JobResources": {
        "ClusterConfig": {
            "InstanceCount": number,
            "InstanceType": "string",
            "VolumeKmsKeyId": "string",
            "VolumeSizeInGB": number
        }
    },
    "ModelBiasAppSpecification": {
        "ConfigUri": "string",
        "Environment": {
            "string": "string"
        },
        "ImageUri": "string"
    },
    "ModelBiasBaselineConfig": {
        "BaseliningJobName": "string",
        "ConstraintsResource": {
            "S3Uri": "string"
        }
    },
    "ModelBiasJobInput": {
        "EndpointInput": {
            "EndpointName": "string",
            "EndTimeOffset": "string",
            "FeaturesAttribute": "string",
            "InferenceAttribute": "string",
            "LocalPath": "string",
            "ProbabilityAttribute": "string",
            "ProbabilityThresholdAttribute": number,
            "S3DataDistributionType": "string",
            "S3InputMode": "string",
            "StartTimeOffset": "string"
        },
        "GroundTruthS3Input": {
            "S3Uri": "string"
        }
    },
    "ModelBiasJobOutputConfig": {
        "KmsKeyId": "string",
        "MonitoringOutputs": [
            {
                "S3Output": {
                    "LocalPath": "string",
                    "S3UploadMode": "string",
                    "S3Uri": "string"
                }
            }
        ]
    },
    "NetworkConfig": {
        "EnableInterContainerTrafficEncryption": boolean,
        "EnableNetworkIsolation": boolean,
        "VpcConfig": {
```
"SecurityGroupIds": [ "string" ],
"Subnets": [ "string" ]
},
"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeInSeconds": number
},
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
]}

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 117)**

The name of the bias job definition. The name must be unique within an AWS Region in the AWS account.

Type: String


Pattern: ^[a-zA-Z0-9\-\*][a-zA-Z0-9\-\*]+$

Required: Yes

**JobResources (p. 117)**

Identifies the resources to deploy for a monitoring job.

Type: MonitoringResources (p. 1210) object

Required: Yes

**ModelBiasAppSpecification (p. 117)**

Configures the model bias job to run a specified Docker container image.

Type: ModelBiasAppSpecification (p. 1150) object

Required: Yes

**ModelBiasBaselineConfig (p. 117)**

The baseline configuration for a model bias job.

Type: ModelBiasBaselineConfig (p. 1151) object

Required: No

**ModelBiasJobInput (p. 117)**

Inputs for the model bias job.

Type: ModelBiasJobInput (p. 1152) object
Required: Yes

**ModelBiasJobOutputConfig (p. 117)**

The output configuration for monitoring jobs.

Type: MonitoringOutputConfig (p. 1209) object

Required: Yes

**NetworkConfig (p. 117)**

Networking options for a model bias job.

Type: MonitoringNetworkConfig (p. 1207) object

Required: No

**RoleArn (p. 117)**

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z\-0-9+=,.@\-_/]+$

Required: Yes

**StoppingCondition (p. 117)**

A time limit for how long the monitoring job is allowed to run before stopping.

Type: MonitoringStoppingCondition (p. 1219) object

Required: No

**Tags (p. 117)**

(Optional) An array of key-value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```json
{
    "JobDefinitionArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
JobDefinitionArn (p. 119)

The Amazon Resource Name (ARN) of the model bias job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: . *

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateModelExplainabilityJobDefinition
Service: Amazon SageMaker Service

Creates the definition for a model explainability job.

Request Syntax

```
{
    "JobDefinitionName": "string",
    "JobResources": {
        "ClusterConfig": {
            "InstanceCount": number,
            "InstanceType": "string",
            "VolumeKmsKeyId": "string",
            "VolumeSizeInGB": number
        }
    },
    "ModelExplainabilityAppSpecification": {
        "ConfigUri": "string",
        "Environment": {
            "string": "string"
        },
        "ImageUri": "string"
    },
    "ModelExplainabilityBaselineConfig": {
        "BaselineJobName": "string",
        "ConstraintsResource": {
            "S3Uri": "string"
        }
    },
    "ModelExplainabilityJobInput": {
        "EndpointInput": {
            "EndpointName": "string",
            "EndTimeOffset": "string",
            "FeaturesAttribute": "string",
            "InferenceAttribute": "string",
            "LocalPath": "string",
            "ProbabilityAttribute": "string",
            "ProbabilityThresholdAttribute": number,
            "S3DataDistributionType": "string",
            "S3InputMode": "string",
            "StartTimeOffset": "string"
        }
    },
    "ModelExplainabilityJobOutputConfig": {
        "KmsKeyId": "string",
        "MonitoringOutputs": [
            {
                "S3Output": {
                    "LocalPath": "string",
                    "S3UploadMode": "string",
                    "S3Uri": "string"
                }
            }
        ]
    },
    "NetworkConfig": {
        "EnableInterContainerTrafficEncryption": boolean,
        "EnableNetworkIsolation": boolean,
        "VpcConfig": {
            "SecurityGroupId": [ "string" ],
            "Subnets": [ "string" ]
        }
    }
}
```
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 121)**

The name of the model explainability job definition. The name must be unique within an AWS Region in the AWS account.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$

Required: Yes

**JobResources (p. 121)**

Identifies the resources to deploy for a monitoring job.

Type: MonitoringResources (p. 1210) object

Required: Yes

**ModelExplainabilityAppSpecification (p. 121)**

Configures the model explainability job to run a specified Docker container image.

Type: ModelExplainabilityAppSpecification (p. 1159) object

Required: Yes

**ModelExplainabilityBaselineConfig (p. 121)**

The baseline configuration for a model explainability job.

Type: ModelExplainabilityBaselineConfig (p. 1160) object

Required: No

**ModelExplainabilityJobInput (p. 121)**

Inputs for the model explainability job.

Type: ModelExplainabilityJobInput (p. 1161) object

Required: Yes
ModelExplainabilityJobOutputConfig (p. 121)

The output configuration for monitoring jobs.

Type: MonitoringOutputConfig (p. 1209) object

Required: Yes

NetworkConfig (p. 121)

Networking options for a model explainability job.

Type: MonitoringNetworkConfig (p. 1207) object

Required: No

RoleArn (p. 121)

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-z\-\]*:iam::\d{12}:role/?[a-zA-Z0-9\-\+=,.@\-\_\/%]+$

Required: Yes

StoppingCondition (p. 121)

A time limit for how long the monitoring job is allowed to run before stopping.

Type: MonitoringStoppingCondition (p. 1219) object

Required: No

Tags (p. 121)

(Optional) An array of key-value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```json
{
    "JobDefinitionArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobDefinitionArn (p. 123)

The Amazon Resource Name (ARN) of the model explainability job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: . *

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse
Resource being accessed is in use.
HTTP Status Code: 400

ResourceLimitExceeded
You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.
HTTP Status Code: 400

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateModelPackage

Service: Amazon SageMaker Service

Creates a model package that you can use to create SageMaker models or list on AWS Marketplace, or a versioned model that is part of a model group. Buyers can subscribe to model packages listed on AWS Marketplace to create models in SageMaker.

To create a model package by specifying a Docker container that contains your inference code and the Amazon S3 location of your model artifacts, provide values for `InferenceSpecification`. To create a model from an algorithm resource that you created or subscribed to in AWS Marketplace, provide a value for `SourceAlgorithmSpecification`.

**Note**
There are two types of model packages:

- Versioned - a model that is part of a model group in the model registry.
- Unversioned - a model package that is not part of a model group.

**Request Syntax**

```
{
  "AdditionalInferenceSpecifications": [
    {
      "Containers": [
        {
          "ContainerHostname": "string",
          "Environment": {
            "string": "string"
          },
          "Framework": "string",
          "FrameworkVersion": "string",
          "Image": "string",
          "ImageDigest": "string",
          "ModelDataUrl": "string",
          "ModelInput": {
            "DataInputConfig": "string"
          },
          "NearestModelName": "string",
          "ProductId": "string"
        }
      ],
      "Description": "string",
      "Name": "string",
      "SupportedContentTypes": [ "string" ],
      "SupportedRealtimeInferenceInstanceTypes": [ "string" ],
      "SupportedResponseMIMETypes": [ "string" ],
      "SupportedTransformInstanceTypes": [ "string" ]
    }
  ],
  "CertifyForMarketplace": boolean,
  "ClientToken": "string",
  "CustomerMetadataProperties": {
    "string": "string"
  },
  "Domain": "string",
  "DriftCheckBaselines": {
    "Bias": {
      "ConfigFile": {
        "ContentDigest": "string",
        "ContentType": "string"
      }
    }
  }
}
```
"S3Uri": "string"
},
"PostTrainingConstraints": {
  "ContentDigest": "string",
  "ContentType": "string",
  "S3Uri": "string"
},
"PreTrainingConstraints": {
  "ContentDigest": "string",
  "ContentType": "string",
  "S3Uri": "string"
}
},
"Explainability": {
  "ConfigFile": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  },
  "Constraints": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  }
},
"ModelDataQuality": {
  "Constraints": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  },
  "Statistics": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  }
},
"ModelQuality": {
  "Constraints": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  },
  "Statistics": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  }
}
},
"InferenceSpecification": {
  "Containers": [
  {
    "ContainerHostname": "string",
    "Environment": {
      "string": "string"
    },
    "Framework": "string",
    "FrameworkVersion": "string",
    "Image": "string",
    "ImageDigest": "string",
    "ModelDataUrl": "string",
    "ModelInput": {
      "DataInputConfig": "string"
    },
    "NearestModelName": "string",
    "Port": "integer",
    "RoleArn": "string",
    "VolumeSizeInGB": "integer"
  }
]
"ProductId": "string"
},
"SupportedContentTypes": [ "string" ],
"SupportedRealtimeInferenceInstanceTypes": [ "string" ],
"SupportedResponseMIMETypes": [ "string" ],
"SupportedTransformInstanceTypes": [ "string" ]
},
"MetadataProperties": {
"CommitId": "string",
"GeneratedBy": "string",
"ProjectId": "string",
"Repository": "string"
},
"ModelApprovalStatus": "string",
"ModelMetrics": {
"Bias": {
"PostTrainingReport": {
"ContentDigest": "string",
"ContentType": "string",
"S3Uri": "string"
},
"PreTrainingReport": {
"ContentDigest": "string",
"ContentType": "string",
"S3Uri": "string"
},
"Report": {
"ContentDigest": "string",
"ContentType": "string",
"S3Uri": "string"
}
},
"Explainability": {
"Report": {
"ContentDigest": "string",
"ContentType": "string",
"S3Uri": "string"
}
},
"ModelDataQuality": {
"Constraints": {
"ContentDigest": "string",
"ContentType": "string",
"S3Uri": "string"
},
"Statistics": {
"ContentDigest": "string",
"ContentType": "string",
"S3Uri": "string"
}
},
"ModelQuality": {
"Constraints": {
"ContentDigest": "string",
"ContentType": "string",
"S3Uri": "string"
},
"Statistics": {
"ContentDigest": "string",
"ContentType": "string",
"S3Uri": "string"
}
}
},
"ModelPackageDescription": "string"
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.
**AdditionalInferenceSpecifications (p. 125)**

An array of additional Inference Specification objects. Each additional Inference Specification specifies artifacts based on this model package that can be used on inference endpoints. Generally used with SageMaker Neo to store the compiled artifacts.

Type: Array of `AdditionalInferenceSpecificationDefinition (p. 863)` objects

Array Members: Minimum number of 1 item. Maximum number of 15 items.

Required: No

**CertifyForMarketplace (p. 125)**

Whether to certify the model package for listing on AWS Marketplace.

This parameter is optional for unversioned models, and does not apply to versioned models.

Type: Boolean

Required: No

**ClientToken (p. 125)**

A unique token that guarantees that the call to this API is idempotent.

Type: String


Pattern: `^[a-zA-Z0-9-]{1,128}$`

Required: No

**CustomerMetadataProperties (p. 125)**

The metadata properties associated with the model package versions.

Type: String to string map

Map Entries: Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `^[a-zA-Z0-9-_.:/=+@]*$`(1,128)

Value Length Constraints: Minimum length of 1. Maximum length of 256.

Value Pattern: `^[a-zA-Z0-9-_.:/=+@]*$`(1,256)

Required: No

**Domain (p. 125)**

The machine learning domain of your model package and its components. Common machine learning domains include computer vision and natural language processing.

Type: String

Required: No

**DriftCheckBaselines (p. 125)**

Represents the drift check baselines that can be used when the model monitor is set using the model package. For more information, see the topic on Drift Detection against Previous Baselines in SageMaker Pipelines in the Amazon SageMaker Developer Guide.
Type: `DriftCheckBaselines (p. 998)` object

Required: No

**InferenceSpecification (p. 125)**

Specifies details about inference jobs that can be run with models based on this model package, including the following:
- The Amazon ECR paths of containers that contain the inference code and model artifacts.
- The instance types that the model package supports for transform jobs and real-time endpoints used for inference.
- The input and output content formats that the model package supports for inference.

Type: `InferenceSpecification (p. 1105)` object

Required: No

**MetadataProperties (p. 125)**

Metadata properties of the tracking entity, trial, or trial component.

Type: `MetadataProperties (p. 1143)` object

Required: No

**ModelApprovalStatus (p. 125)**

Whether the model is approved for deployment.

This parameter is optional for versioned models, and does not apply to unversioned models.

For versioned models, the value of this parameter must be set to `Approved` to deploy the model.

Type: String

Valid Values: Approved | Rejected | PendingManualApproval

Required: No

**ModelMetrics (p. 125)**

A structure that contains model metrics reports.

Type: `ModelMetrics (p. 1167)` object

Required: No

**ModelPackageDescription (p. 125)**

A description of the model package.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `[^\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*`

Required: No

**ModelPackageGroupName (p. 125)**

The name or Amazon Resource Name (ARN) of the model package group that this model version belongs to.

This parameter is required for versioned models, and does not apply to unversioned models.
Type: String


Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:[a-z\-]*\//)?([a-zA-Z0-9][a-zA-Z0-9-]{0,62})(?<!-)$

Required: No

**ModelPackageName (p. 125)**

The name of the model package. The name must have 1 to 63 characters. Valid characters are a-z, A-Z, 0-9, and - (hyphen).

This parameter is required for unversioned models. It is not applicable to versioned models.

Type: String


Pattern: ^[a-zA-Z0-9](\*[a-zA-Z0-9]){0,62}$

Required: No

**SamplePayloadUrl (p. 125)**

The Amazon Simple Storage Service (Amazon S3) path where the sample payload are stored. This path must point to a single gzip compressed tar archive (.tar.gz suffix).

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/\s]+)/(.*)$

Required: No

**SourceAlgorithmSpecification (p. 125)**

Details about the algorithm that was used to create the model package.

Type: SourceAlgorithmSpecification (p. 1368) object

Required: No

**Tags (p. 125)**

A list of key value pairs associated with the model. For more information, see Tagging AWS resources in the AWS General Reference Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Task (p. 125)**

The machine learning task your model package accomplishes. Common machine learning tasks include object detection and image classification. The following tasks are supported by Inference Recommender: "IMAGE_CLASSIFICATION" | "OBJECT_DETECTION" | "TEXT_GENERATION" | "IMAGE_SEGMENTATION" | "FILL_MASK" | "CLASSIFICATION" | "REGRESSION" | "OTHER".

Specify "OTHER" if none of the tasks listed fit your use case.

Type: String
CreateModelPackage

Required: No

ValidationSpecification (p. 125)

Specifies configurations for one or more transform jobs that SageMaker runs to test the model package.

Type: ModelPackageValidationSpecification (p. 1185) object

Required: No

Response Syntax

```
{
    "ModelPackageArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ModelPackageArn (p. 132)

The Amazon Resource Name (ARN) of the new model package.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:model-package/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ConflictException

There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreateModelPackageGroup
Service: Amazon SageMaker Service

Creates a model group. A model group contains a group of model versions.

Request Syntax

```json
{
    "ModelPackageGroupDescription": "string",
    "ModelPackageGroupName": "string",
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ModelPackageGroupDescription (p. 134)

A description for the model group.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `[\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*`

Required: No

ModelPackageGroupName (p. 134)

The name of the model group.

Type: String


Pattern: `^[a-zA-Z0-9-]*[a-zA-Z0-9]{0,62}$`

Required: Yes

Tags (p. 134)

A list of key value pairs associated with the model group. For more information, see Tagging AWS resources in the AWS General Reference Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No
Response Syntax

```
{
    "ModelPackageGroupArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**ModelPackageGroupArn (p. 135)**

The Amazon Resource Name (ARN) of the model group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package-group/.*`

Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
**CreateModelQualityJobDefinition**

Service: Amazon SageMaker Service

Creates a definition for a job that monitors model quality and drift. For information about model monitor, see Amazon SageMaker Model Monitor.

**Request Syntax**

```json
{
  "JobDefinitionName": "string",
  "JobResources": {
    "ClusterConfig": {
      "InstanceCount": number,
      "InstanceType": "string",
      "VolumeKmsKeyId": "string",
      "VolumeSizeInGB": number
    }
  },
  "ModelQualityAppSpecification": {
    "ContainerArguments": [ "string" ],
    "ContainerEntrypoint": [ "string" ],
    "Environment": {
      "string" : "string"
    },
    "ImageUri": "string",
    "PostAnalyticsProcessorSourceUri": "string",
    "ProblemType": "string",
    "RecordPreprocessorSourceUri": "string"
  },
  "ModelQualityBaselineConfig": {
    "BaseliningJobName": "string",
    "ConstraintsResource": {
      "S3Uri": "string"
    }
  },
  "ModelQualityJobInput": {
    "EndpointInput": {
      "EndpointName": "string",
      "EndTimeOffset": "string",
      "FeaturesAttribute": "string",
      "InferenceAttribute": "string",
      "LocalPath": "string",
      "ProbabilityAttribute": "string",
      "ProbabilityThresholdAttribute": number,
      "S3DataDistributionType": "string",
      "S3InputMode": "string",
      "StartTimeOffset": "string"
    },
    "GroundTruthS3Input": {
      "S3Uri": "string"
    }
  },
  "ModelQualityJobOutputConfig": {
    "KmsKeyId": "string",
    "MonitoringOutputs": [
      {
        "S3Output": {
          "LocalPath": "string",
          "S3UploadMode": "string",
          "S3Uri": "string"
        }
      }
    ]
  }
}
```
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 136)**

The name of the monitoring job definition.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*$`

Required: Yes

**JobResources (p. 136)**

Identifies the resources to deploy for a monitoring job.

Type: MonitoringResources (p. 1210) object

Required: Yes

**ModelQualityAppSpecification (p. 136)**

The container that runs the monitoring job.

Type: ModelQualityAppSpecification (p. 1187) object

Required: Yes

**ModelQualityBaselineConfig (p. 136)**

Specifies the constraints and baselines for the monitoring job.

Type: ModelQualityBaselineConfig (p. 1189) object

Required: No
Amazon SageMaker Amazon Sagemaker API Reference
CreateModelQualityJobDefinition

ModelQualityJobInput (p. 136)
A list of the inputs that are monitored. Currently endpoints are supported.

Type: ModelQualityJobInput (p. 1190) object

Required: Yes

ModelQualityJobOutputConfig (p. 136)
The output configuration for monitoring jobs.

Type: MonitoringOutputConfig (p. 1209) object

Required: Yes

NetworkConfig (p. 136)
Specifies the network configuration for the monitoring job.

Type: MonitoringNetworkConfig (p. 1207) object

Required: No

RoleArn (p. 136)
The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_\/]+$

Required: Yes

StoppingCondition (p. 136)
A time limit for how long the monitoring job is allowed to run before stopping.

Type: MonitoringStoppingCondition (p. 1219) object

Required: No

Tags (p. 136)
(Optional) An array of key-value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```
{
   "JobDefinitionArn": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**JobDefinitionArn (p. 138)**

The Amazon Resource Name (ARN) of the model quality monitoring job.

- Type: String
- Length Constraints: Maximum length of 256.
- Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

- HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

- HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateMonitoringSchedule

Service: Amazon SageMaker Service

Creates a schedule that regularly starts Amazon SageMaker Processing Jobs to monitor the data captured for an Amazon SageMaker Endpoint.

Request Syntax

```json
{
  "MonitoringScheduleConfig": {
    "MonitoringJobDefinition": {
      "BaselineConfig": {
        "BaseliningJobName": "string",
        "ConstraintsResource": {
          "S3Uri": "string"
        },
        "StatisticsResource": {
          "S3Uri": "string"
        }
      },
      "Environment": {
        "string": "string"
      },
      "MonitoringAppSpecification": {
        "ContainerArguments": [ "string" ],
        "ContainerEntrypoint": [ "string" ],
        "ImageUri": "string",
        "PostAnalyticsProcessorSourceUri": "string",
        "RecordPreprocessorSourceUri": "string"
      },
      "MonitoringInputs": [{
        "EndpointInput": {
          "EndpointName": "string",
          "EndTimeOffset": "string",
          "FeaturesAttribute": "string",
          "InferenceAttribute": "string",
          "LocalPath": "string",
          "ProbabilityAttribute": "string",
          "ProbabilityThresholdAttribute": number,
          "S3DataDistributionType": "string",
          "S3InputMode": "string",
          "StartTimeOffset": "string"
        }
      }],
      "MonitoringOutputConfig": {
        "KmsKeyId": "string",
        "MonitoringOutputs": [{
          "S3Output": {
            "LocalPath": "string",
            "S3UploadMode": "string",
            "S3Uri": "string"
          }
        }]
      },
      "MonitoringResources": {
        "ClusterConfig": {
          "InstanceCount": number,
          "InstanceType": "string",
          "VolumeKmsKeyId": "string"
        }
      }
    }
  }
}
```
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**MonitoringScheduleConfig (p. 140)**

The configuration object that specifies the monitoring schedule and defines the monitoring job.

Type: MonitoringScheduleConfig (p. 1215) object

Required: Yes

**MonitoringScheduleName (p. 140)**

The name of the monitoring schedule. The name must be unique within an AWS Region within an AWS account.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

**Tags (p. 140)**

(Optional) An array of key-value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.
Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```
{
  "MonitoringScheduleArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

MonitoringScheduleArn (p. 142)

The Amazon Resource Name (ARN) of the monitoring schedule.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResoureInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreateNotebookInstance
Service: Amazon SageMaker Service

Creates an SageMaker notebook instance. A notebook instance is a machine learning (ML) compute instance running on a Jupyter notebook.

In a CreateNotebookInstance request, specify the type of ML compute instance that you want to run. SageMaker launches the instance, installs common libraries that you can use to explore datasets for model training, and attaches an ML storage volume to the notebook instance.

SageMaker also provides a set of example notebooks. Each notebook demonstrates how to use SageMaker with a specific algorithm or with a machine learning framework.

After receiving the request, SageMaker does the following:

1. Creates a network interface in the SageMaker VPC.
2. (Option) If you specified SubnetId, SageMaker creates a network interface in your own VPC, which is inferred from the subnet ID that you provide in the input. When creating this network interface, SageMaker attaches the security group that you specified in the request to the network interface that it creates in your VPC.
3. Launches an EC2 instance of the type specified in the request in the SageMaker VPC. If you specified SubnetId of your VPC, SageMaker specifies both network interfaces when launching this instance. This enables inbound traffic from your own VPC to the notebook instance, assuming that the security groups allow it.

After creating the notebook instance, SageMaker returns its Amazon Resource Name (ARN). You can't change the name of a notebook instance after you create it.

After SageMaker creates the notebook instance, you can connect to the Jupyter server and work in Jupyter notebooks. For example, you can write code to explore a dataset that you can use for model training, train a model, host models by creating SageMaker endpoints, and validate hosted models.

For more information, see How It Works.

Request Syntax

```
{
    "AcceleratorTypes": [ "string" ],
    "AdditionalCodeRepositories": [ "string" ],
    "DefaultCodeRepository": "string",
    "DirectInternetAccess": "string",
    "InstanceMetadataServiceConfiguration": {
        "MinimumInstanceMetadataServiceVersion": "string"
    },
    "InstanceType": "string",
    "KmsKeyId": "string",
    "LifecycleConfigName": "string",
    "NotebookInstanceName": "string",
    "PlatformIdentifier": "string",
    "RoleArn": "string",
    "RootAccess": "string",
    "SecurityGroupIds": [ "string" ],
    "SubnetId": "string",
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
```

"VolumeSizeInGB": number
}

# Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

## AcceleratorTypes (p. 144)

A list of Elastic Inference (EI) instance types to associate with this notebook instance. Currently, only one instance type can be associated with a notebook instance. For more information, see Using Elastic Inference in Amazon SageMaker.

Type: Array of strings

Valid Values: `ml.eia1.medium` | `ml.eia1.large` | `ml.eia1.xlarge` | `ml.eia2.medium` | `ml.eia2.large` | `ml.eia2.xlarge`

Required: No

## AdditionalCodeRepositories (p. 144)

An array of up to three Git repositories to associate with the notebook instance. These can be either the names of Git repositories stored as resources in your account, or the URL of Git repositories in AWS CodeCommit or in any other Git repository. These repositories are cloned at the same level as the default repository of your notebook instance. For more information, see Associating Git Repositories with SageMaker Notebook Instances.

Type: Array of strings

Array Members: Maximum number of 3 items.


Pattern: `^https://([^/]+)/?([^/]+)$|^([a-zA-Z0-9](-[a-zA-Z0-9]+)*)`

Required: No

## DefaultCodeRepository (p. 144)

A Git repository to associate with the notebook instance as its default code repository. This can be either the name of a Git repository stored as a resource in your account, or the URL of a Git repository in AWS CodeCommit or in any other Git repository. When you open a notebook instance, it opens in the directory that contains this repository. For more information, see Associating Git Repositories with SageMaker Notebook Instances.

Type: String


Pattern: `^https://([^/]+)/?([^/]+)$|^([a-zA-Z0-9](-[a-zA-Z0-9]+)*)`

Required: No

## DirectInternetAccess (p. 144)

Sets whether SageMaker provides internet access to the notebook instance. If you set this to `Disabled` this notebook instance is able to access resources only in your VPC, and is not be able to connect to SageMaker training and endpoint services unless you configure a NAT Gateway in your VPC.
For more information, see Notebook Instances Are Internet-Enabled by Default. You can set the value of this parameter to Disabled only if you set a value for the SubnetId parameter.

Type: String

Valid Values: Enabled | Disabled

Required: No

**InstanceMetadataServiceConfiguration (p. 144)**

Information on the IMDS configuration of the notebook instance

Type: InstanceMetadataServiceConfiguration (p. 1111) object

Required: No

**InstanceType (p. 144)**

The type of ML compute instance to launch for the notebook instance.

Type: String

Valid Values: ml.t2.medium | ml.t2.large | ml.t2.xlarge | ml.t2.2xlarge | ml.t3.medium | ml.t3.large | ml.t3.xlarge | ml.t3.2xlarge | ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge | ml.m4.16xlarge | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge | ml.m5.24xlarge | ml.m5d.large | ml.m5d.xlarge | ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.8xlarge | ml.m5d.12xlarge | ml.m5d.16xlarge | ml.m5d.24xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge | ml.c4.8xlarge | ml.c5.xlarge | ml.c5.2xlarge | ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge | ml.c5d.xlarge | ml.c5d.2xlarge | ml.c5d.4xlarge | ml.c5d.9xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p3.xlarge | ml.p3.16xlarge | ml.p3dn.24xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.g4dn.48xlarge | ml.g5.xlarge | ml.g5.2xlarge | ml.g5.4xlarge | ml.g5.8xlarge | ml.g5.16xlarge | ml.g5.24xlarge | ml.g5.48xlarge

Required: Yes

**KmsKeyId (p. 144)**

The Amazon Resource Name (ARN) of a AWS Key Management Service key that SageMaker uses to encrypt data on the storage volume attached to your notebook instance. The KMS key you provide must be enabled. For information, see Enabling and Disabling Keys in the AWS Key Management Service Developer Guide.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

**LifecycleConfigName (p. 144)**

The name of a lifecycle configuration to associate with the notebook instance. For information about lifestyle configurations, see Step 2.1: (Optional) Customize a Notebook Instance.

Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])* 
Required: No

**NotebookInstanceName (p. 144)**

The name of the new notebook instance.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])* 
Required: Yes

**PlatformIdentifier (p. 144)**

The platform identifier of the notebook instance runtime environment.
Type: String
Length Constraints: Maximum length of 15.
Pattern: ^(notebook-al1-v1|notebook-al2-v1|notebook-al2-v2)$
Required: No

**RoleArn (p. 144)**

When you send any requests to AWS resources from the notebook instance, SageMaker assumes this role to perform tasks on your behalf. You must grant this role necessary permissions so SageMaker can perform these tasks. The policy must allow the SageMaker service principal (sagemaker.amazonaws.com) permissions to assume this role. For more information, see SageMaker Roles.

*Note*
To be able to pass this role to SageMaker, the caller of this API must have the `iam:PassRole` permission.

Type: String
Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/?[a-zA-Z0-9+=,.@\-_\d/]+$
Required: Yes

**RootAccess (p. 144)**

Whether root access is enabled or disabled for users of the notebook instance. The default value is Enabled.

*Note*
Lifecycle configurations need root access to be able to set up a notebook instance. Because of this, lifecycle configurations associated with a notebook instance always run with root access even if you disable root access for users.

Type: String
Valid Values: Enabled | Disabled
Required: No
SecurityGroupIds (p. 144)

The VPC security group IDs, in the form sg-xxxxxxxx. The security groups must be for the same VPC as specified in the subnet.

Type: Array of strings

Array Members: Maximum number of 5 items.

Length Constraints: Maximum length of 32.

Pattern: \([-0-9a-zA-Z]+\]

Required: No

SubnetId (p. 144)

The ID of the subnet in a VPC to which you would like to have a connectivity from your ML compute instance.

Type: String

Length Constraints: Maximum length of 32.

Pattern: \([-0-9a-zA-Z]+\]

Required: No

Tags (p. 144)

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

VolumeSizeInGB (p. 144)

The size, in GB, of the ML storage volume to attach to the notebook instance. The default value is 5 GB.

Type: Integer


Required: No

Response Syntax

```json
{
   "NotebookInstanceArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
**NotebookInstanceArn (p. 148)**

The Amazon Resource Name (ARN) of the notebook instance.

Type: String

Length Constraints: Maximum length of 256.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateNotebookInstanceLifecycleConfig
Service: Amazon SageMaker Service

Creates a lifecycle configuration that you can associate with a notebook instance. A *lifecycle configuration* is a collection of shell scripts that run when you create or start a notebook instance.

Each lifecycle configuration script has a limit of 16384 characters.

The value of the `$PATH` environment variable that is available to both scripts is `/sbin:bin:/usr/sbin:/usr/bin`.


Lifecycle configuration scripts cannot run for longer than 5 minutes. If a script runs for longer than 5 minutes, it fails and the notebook instance is not created or started.

For information about notebook instance lifecycle configurations, see Step 2.1: (Optional) Customize a Notebook Instance.

**Request Syntax**

```json
{
   "NotebookInstanceLifecycleConfigName": "string",
   "OnCreate": [
   {
      "Content": "string"
   }
  ],
  "OnStart": [
   {
      "Content": "string"
   }
  ]
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**NotebookInstanceLifecycleConfigName (p. 150)**

The name of the lifecycle configuration.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `[^a-zA-Z0-9][a-zA-Z0-9]*[^a-zA-Z0-9]` *

Required: Yes

**OnCreate (p. 150)**

A shell script that runs only once, when you create a notebook instance. The shell script must be a base64-encoded string.
Type: Array of NotebookInstanceLifecycleHook (p. 1225) objects

Array Members: Maximum number of 1 item.

Required: No

OnStart (p. 150)

A shell script that runs every time you start a notebook instance, including when you create the notebook instance. The shell script must be a base64-encoded string.

Type: Array of NotebookInstanceLifecycleHook (p. 1225) objects

Array Members: Maximum number of 1 item.

Required: No

Response Syntax

```json
{
   "NotebookInstanceLifecycleConfigArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**NotebookInstanceLifecycleConfigArn (p. 151)**

The Amazon Resource Name (ARN) of the lifecycle configuration.

Type: String

Length Constraints: Maximum length of 256.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreatePipeline
Service: Amazon SageMaker Service

Creates a pipeline using a JSON pipeline definition.

Request Syntax

```
{
  "ClientRequestToken": "string",
  "ParallelismConfiguration": {
    "MaxParallelExecutionSteps": number
  },
  "PipelineDefinition": "string",
  "PipelineDefinitionS3Location": {
    "Bucket": "string",
    "ObjectKey": "string",
    "VersionId": "string"
  },
  "PipelineDescription": "string",
  "PipelineDisplayName": "string",
  "PipelineName": "string",
  "RoleArn": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ClientRequestToken (p. 153)**

A unique, case-sensitive identifier that you provide to ensure the idempotency of the operation. An idempotent operation completes no more than one time.

Type: String


Required: Yes

**ParallelismConfiguration (p. 153)**

This is the configuration that controls the parallelism of the pipeline. If specified, it applies to all runs of this pipeline by default.

Type: ParallelismConfiguration (p. 1247) object

Required: No

**PipelineDefinition (p. 153)**

The JSON pipeline definition of the pipeline.

Type: String
CreatePipeline

Pattern: .*(?:\r\n\t\].*)*
Required: No

**PipelineDefinitionS3Location (p. 153)**

The location of the pipeline definition stored in Amazon S3. If specified, SageMaker will retrieve the pipeline definition from this location.

Type: PipelineDefinitionS3Location (p. 1261) object
Required: No

**PipelineDescription (p. 153)**

A description of the pipeline.

Type: String
Length Constraints: Minimum length of 0. Maximum length of 3072.
Pattern: .*
Required: No

**PipelineDisplayName (p. 153)**

The display name of the pipeline.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])+{0,255}$
Required: No

**PipelineName (p. 153)**

The name of the pipeline.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])+{0,255}$
Required: Yes

**RoleArn (p. 153)**

The Amazon Resource Name (ARN) of the role used by the pipeline to access and create resources.

Type: String
Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z\-_0-9+=,.@\-_\d]+$*
Required: Yes

**Tags (p. 153)**

A list of tags to apply to the created pipeline.
Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```json
{
   "PipelineArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**PipelineArn (p. 155)**

The Amazon Resource Name (ARN) of the created pipeline.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreatePresignedDomainUrl
Service: Amazon SageMaker Service

Creates a URL for a specified UserProfile in a Domain. When accessed in a web browser, the user will be automatically signed in to Amazon SageMaker Studio, and granted access to all of the Apps and files associated with the Domain's Amazon Elastic File System (EFS) volume. This operation can only be called when the authentication mode equals IAM.

The IAM role or user used to call this API defines the permissions to access the app. Once the presigned URL is created, no additional permission is required to access this URL. IAM authorization policies for this API are also enforced for every HTTP request and WebSocket frame that attempts to connect to the app.

You can restrict access to this API and to the URL that it returns to a list of IP addresses, Amazon VPCs or Amazon VPC Endpoints that you specify. For more information, see Connect to SageMaker Studio Through an Interface VPC Endpoint.

Note
The URL that you get from a call to CreatePresignedDomainUrl has a default timeout of 5 minutes. You can configure this value using ExpiresInSeconds. If you try to use the URL after the timeout limit expires, you are directed to the AWS console sign-in page.

Request Syntax

```json
{
  "DomainId": "string",
  "ExpiresInSeconds": number,
  "SessionExpirationDurationInSeconds": number,
  "UserProfileName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainId (p. 157)**

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**ExpiresInSeconds (p. 157)**

The number of seconds until the pre-signed URL expires. This value defaults to 300.

Type: Integer


Required: No

**SessionExpirationDurationInSeconds (p. 157)**

The session expiration duration in seconds. This value defaults to 43200.
Type: Integer


Required: No

UserProfileName (p. 157)

The name of the UserProfile to sign-in as.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Syntax

```
{
   "AuthorizedUrl": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AuthorizedUrl (p. 158)

The presigned URL.

Type: String

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreatePresignedNotebookInstanceUrl

Service: Amazon SageMaker Service

Returns a URL that you can use to connect to the Jupyter server from a notebook instance. In the SageMaker console, when you choose Open next to a notebook instance, SageMaker opens a new tab showing the Jupyter server home page from the notebook instance. The console uses this API to get the URL and show the page.

The IAM role or user used to call this API defines the permissions to access the notebook instance. Once the presigned URL is created, no additional permission is required to access this URL. IAM authorization policies for this API are also enforced for every HTTP request and WebSocket frame that attempts to connect to the notebook instance.

You can restrict access to this API and to the URL that it returns to a list of IP addresses that you specify. Use the NotIpAddress condition operator and the aws:SourceIp condition context key to specify the list of IP addresses that you want to have access to the notebook instance. For more information, see Limit Access to a Notebook Instance by IP Address.

Note
The URL that you get from a call to CreatePresignedNotebookInstanceUrl (p. 160) is valid only for 5 minutes. If you try to use the URL after the 5-minute limit expires, you are directed to the AWS console sign-in page.

Request Syntax

```json
{
    "NotebookInstanceName": "string",
    "SessionExpirationDurationInSeconds": number
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**NotebookInstanceName** (p. 160)

The name of the notebook instance.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])* Required: Yes

**SessionExpirationDurationInSeconds** (p. 160)

The duration of the session, in seconds. The default is 12 hours.

Type: Integer


Required: No
Response Syntax

```json
{
    "AuthorizedUrl": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AuthorizedUrl (p. 161)**

A JSON object that contains the URL string.

Type: String

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateProcessingJob
Service: Amazon SageMaker Service

Creates a processing job.

Request Syntax

```
{
  "AppSpecification": {
    "ContainerArguments": [ "string" ],
    "ContainerEntrypoint": [ "string" ],
    "ImageUri": "string"
  },
  "Environment": {
    "string": "string"
  },
  "ExperimentConfig": {
    "ExperimentName": "string",
    "TrialComponentDisplayName": "string",
    "TrialName": "string"
  },
  "NetworkConfig": {
    "EnableInterContainerTrafficEncryption": boolean,
    "EnableNetworkIsolation": boolean,
    "VpcConfig": {
      "SecurityGroupIds": [ "string" ],
      "Subnets": [ "string" ]
    }
  },
  "ProcessingInputs": [
    {
      "AppManaged": boolean,
      "DatasetDefinition": {
        "AthenaDatasetDefinition": {
          "Catalog": "string",
          "Database": "string",
          "KmsKeyId": "string",
          "OutputCompression": "string",
          "OutputFormat": "string",
          "OutputS3Uri": "string",
          "QueryString": "string",
          "WorkGroup": "string"
        },
        "DataDistributionType": "string",
        "InputMode": "string",
        "LocalPath": "string",
        "RedshiftDatasetDefinition": {
          "ClusterId": "string",
          "ClusterRoleArn": "string",
          "Database": "string",
          "DbUser": "string",
          "KmsKeyId": "string",
          "OutputCompression": "string",
          "OutputFormat": "string",
          "OutputS3Uri": "string",
          "QueryString": "string"
        }
      },
      "InputName": "string",
      "S3Input": {
        "LocalPath": "string",
        "S3CompressionType": "string",
        "S3DataDistributionType": "string",
        "S3InputMode": "string",
        "S3Key": "string"
      }
    }
  ]
}
```
"S3DataType": "string",
"S3InputMode": "string",
"S3Uri": "string"
}
]

"ProcessingJobName": "string",
"ProcessingOutputConfig": {
  "KmsKeyId": "string",
  "Outputs": [
    {
      "AppManaged": boolean,
      "FeatureStoreOutput": {
        "FeatureGroupName": "string"
      },
      "OutputName": "string",
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
}

"ProcessingResources": {
  "ClusterConfig": {
    "InstanceCount": number,
    "InstanceType": "string",
    "VolumeKmsKeyId": "string",
    "VolumeSizeInGB": number
  }
},

"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeInSeconds": number
},
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
]}

---

**Request Parameters**

For information about the parameters that are common to all actions, see [Common Parameters](p. 1470).

The request accepts the following data in JSON format.

**AppSpecification (p. 162)**

Configures the processing job to run a specified Docker container image.

*Type:* AppSpecification (p. 891) object

*Required:* Yes

**Environment (p. 162)**

The environment variables to set in the Docker container. Up to 100 key and values entries in the map are supported.
Type: String to string map

Map Entries: Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: [a-zA-Z_][a-zA-Z0-9_]*

Value Length Constraints: Maximum length of 256.

Value Pattern: [\S\s]*

Required: No

**ExperimentConfig (p. 162)**

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:

- CreateProcessingJob (p. 162)
- CreateTrainingJob (p. 173)
- CreateTransformJob (p. 182)

Type: ExperimentConfig (p. 1029) object

Required: No

**NetworkConfig (p. 162)**

Networking options for a processing job, such as whether to allow inbound and outbound network calls to and from processing containers, and the VPC subnets and security groups to use for VPC-enabled processing jobs.

Type: NetworkConfig (p. 1223) object

Required: No

**ProcessingInputs (p. 162)**

An array of inputs configuring the data to download into the processing container.

Type: Array of ProcessingInput (p. 1278) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Required: No

**ProcessingJobName (p. 162)**

The name of the processing job. The name must be unique within an AWS Region in the AWS account.

Type: String


Pattern: ^[a-zA-Z0-9]*[a-zA-Z0-9](_[a-zA-Z0-9]*){0,62}$

Required: Yes

**ProcessingOutputConfig (p. 162)**

Output configuration for the processing job.

Type: ProcessingOutputConfig (p. 1287) object
Required: No

**ProcessingResources (p. 162)**

Identifies the resources, ML compute instances, and ML storage volumes to deploy for a processing job. In distributed training, you specify more than one instance.

Type: `ProcessingResources (p. 1288)` object

Required: Yes

**RoleArn (p. 162)**

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.

Type: String


Pattern: `^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@-_]+\$`

Required: Yes

**StoppingCondition (p. 162)**

The time limit for how long the processing job is allowed to run.

Type: `ProcessingStoppingCondition (p. 1292)` object

Required: No

**Tags (p. 162)**

(Optional) An array of key-value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of `Tag (p. 1377)` objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```json
{
  "ProcessingJobArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ProcessingJobArn (p. 165)**

The Amazon Resource Name (ARN) of the processing job.

Type: String

Length Constraints: Maximum length of 256.
Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:processing-job/.*`

**Errors**

For information about the errors that are common to all actions, see [Common Errors](#).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateProject
Service: Amazon SageMaker Service

Creates a machine learning (ML) project that can contain one or more templates that set up an ML pipeline from training to deploying an approved model.

Request Syntax

```json
{
    "ProjectDescription": "string",
    "ProjectName": "string",
    "ServiceCatalogProvisioningDetails": {
        "PathId": "string",
        "ProductId": "string",
        "ProvisioningArtifactId": "string",
        "ProvisioningParameters": [
            {
                "Key": "string",
                "Value": "string"
            }
        ]
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ProjectDescription** (p. 167)

A description for the project.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `[^\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*`

Required: No

**ProjectName** (p. 167)

The name of the project.

Type: String


Pattern: `^[a-zA-Z0-9-][\-][a-zA-Z0-9]{0,31}`

Required: Yes
ServiceCatalogProvisioningDetails (p. 167)

The product ID and provisioning artifact ID to provision a service catalog. The provisioning artifact ID will default to the latest provisioning artifact ID of the product, if you don’t provide the provisioning artifact ID. For more information, see What is AWS Service Catalog.

Type: ServiceCatalogProvisioningDetails (p. 1362) object

Required: Yes

Tags (p. 167)

An array of key-value pairs that you want to use to organize and track your AWS resource costs. For more information, see Tagging AWS resources in the AWS General Reference Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```json
{
  "ProjectArn": "string",
  "ProjectId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ProjectArn (p. 168)

The Amazon Resource Name (ARN) of the project.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:project:.*

ProjectId (p. 168)

The ID of the new project.

Type: String


Pattern: ^[a-zA-Z0-9][a-zA-Z0-9\-]{0,49}$

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateStudioLifecycleConfig
Service: Amazon SageMaker Service

Creates a new Studio Lifecycle Configuration.

Request Syntax

```
{
  "StudioLifecycleConfigAppType": "string",
  "StudioLifecycleConfigContent": "string",
  "StudioLifecycleConfigName": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**StudioLifecycleConfigAppType (p. 170)**

The App type that the Lifecycle Configuration is attached to.

Type: String

Valid Values: JupyterServer | KernelGateway

Required: Yes

**StudioLifecycleConfigContent (p. 170)**

The content of your Studio Lifecycle Configuration script. This content must be base64 encoded.

Type: String


Pattern: \[\S\s\]+

Required: Yes

**StudioLifecycleConfigName (p. 170)**

The name of the Studio Lifecycle Configuration to create.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-]{0,63}$

Required: Yes
Tags (p. 170)

Tags to be associated with the Lifecycle Configuration. Each tag consists of a key and an optional value. Tag keys must be unique per resource. Tags are searchable using the Search API.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Response Syntax

```
{
  "StudioLifecycleConfigArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**StudioLifecycleConfigArn (p. 171)**

The ARN of your created Lifecycle Configuration.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:studio-lifecycle-config/.*`

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreateTrainingJob
Service: Amazon SageMaker Service

Starts a model training job. After training completes, SageMaker saves the resulting model artifacts to an Amazon S3 location that you specify.

If you choose to host your model using SageMaker hosting services, you can use the resulting model artifacts as part of the model. You can also use the artifacts in a machine learning service other than SageMaker, provided that you know how to use them for inference.

In the request body, you provide the following:

- **AlgorithmSpecification**: Identifies the training algorithm to use.
- **HyperParameters**: Specify these algorithm-specific parameters to enable the estimation of model parameters during training. Hyperparameters can be tuned to optimize this learning process. For a list of hyperparameters for each training algorithm provided by SageMaker, see Algorithms.
- **InputDataConfig**: Describes the training dataset and the Amazon S3, EFS, or FSx location where it is stored.
- **OutputDataConfig**: Identifies the Amazon S3 bucket where you want SageMaker to save the results of model training.
- **ResourceConfig**: Identifies the resources, ML compute instances, and ML storage volumes to deploy for model training. In distributed training, you specify more than one instance.
- **EnableManagedSpotTraining**: Optimize the cost of training machine learning models by up to 80% by using Amazon EC2 Spot instances. For more information, see Managed Spot Training.
- **RoleArn**: The Amazon Resource Name (ARN) that SageMaker assumes to perform tasks on your behalf during model training. You must grant this role the necessary permissions so that SageMaker can successfully complete model training.
- **StoppingCondition**: To help cap training costs, use MaxRuntimeInSeconds to set a time limit for training. Use MaxWaitTimeInSeconds to specify how long a managed spot training job has to complete.
- **Environment**: The environment variables to set in the Docker container.
- **RetryStrategy**: The number of times to retry the job when the job fails due to an InternalServerError.

For more information about SageMaker, see How It Works.

**Request Syntax**

```
{
  "AlgorithmSpecification": {
    "AlgorithmName": "string",
    "EnableSageMakerMetricsTimeSeries": boolean,
    "MetricDefinitions": [
      {
        "Name": "string",
        "Regex": "string"
      }
    ],
    "TrainingImage": "string",
    "TrainingInputMode": "string"
  },
  "CheckpointConfig": {
    "LocalPath": "string",
    "S3Uri": "string"
  }
}
```
"DebugHookConfig": {
    "CollectionConfigurations": [
        {
            "CollectionName": "string",
            "CollectionParameters": {
                "string": "string"
            }
        },
        "HookParameters": {
            "string": "string"
        },
        "LocalPath": "string",
        "S3OutputPath": "string"
    }
},
"DebugRuleConfigurations": [
    {
        "InstanceType": "string",
        "LocalPath": "string",
        "RuleConfigurationName": "string",
        "RuleEvaluatorImage": "string",
        "RuleParameters": {
            "string": "string"
        },
        "S3OutputPath": "string",
        "VolumeSizeInGB": number
    }
],
"EnableInterContainerTrafficEncryption": boolean,
"EnableManagedSpotTraining": boolean,
"EnableNetworkIsolation": boolean,
"Environment": {
    "string": "string"
},
"ExperimentConfig": {
    "ExperimentName": "string",
    "TrialComponentDisplayName": "string",
    "TrialName": "string"
},
"HyperParameters": {
    "string": "string"
},
"InputDataConfig": [
    {
        "ChannelName": "string",
        "CompressionType": "string",
        "ContentType": "string",
        "DataSource": {
            "FileSystemDataSource": {
                "DirectoryPath": "string",
                "FileSystemAccessMode": "string",
                "FileSystemId": "string",
                "FileSystemType": "string"
            },
            "S3DataSource": {
                "AttributeNames": [ "string" ],
                "S3DataDistributionType": "string",
                "S3DataType": "string",
                "S3Uri": "string"
            }
        },
        "InputMode": "string",
        "RecordWrapperType": "string",
        "ShuffleConfig": {
            "Seed": number
        }
    }
]
CreateTrainingJob

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.
AlgorithmSpecification (p. 173)

The registry path of the Docker image that contains the training algorithm and algorithm-specific metadata, including the input mode. For more information about algorithms provided by SageMaker, see Algorithms. For information about providing your own algorithms, see Using Your Own Algorithms with Amazon SageMaker.

Type: AlgorithmSpecification (p. 867) object
Required: Yes

CheckpointConfig (p. 173)

Contains information about the output location for managed spot training checkpoint data.

Type: CheckpointConfig (p. 945) object
Required: No

DebugHookConfig (p. 173)

Configuration information for the Debugger hook parameters, metric and tensor collections, and storage paths. To learn more about how to configure the DebugHookConfig parameter, see Use the SageMaker and Debugger Configuration API Operations to Create, Update, and Debug Your Training Job.

Type: DebugHookConfig (p. 980) object
Required: No

DebugRuleConfigurations (p. 173)

Configuration information for Debugger rules for debugging output tensors.

Type: Array of DebugRuleConfiguration (p. 982) objects
Array Members: Minimum number of 0 items. Maximum number of 20 items.
Required: No

EnableInterContainerTrafficEncryption (p. 173)

To encrypt all communications between ML compute instances in distributed training, choose True. Encryption provides greater security for distributed training, but training might take longer. How long it takes depends on the amount of communication between compute instances, especially if you use a deep learning algorithm in distributed training. For more information, see Protect Communications Between ML Compute Instances in a Distributed Training Job.

Type: Boolean
Required: No

EnableManagedSpotTraining (p. 173)

To train models using managed spot training, choose True. Managed spot training provides a fully managed and scalable infrastructure for training machine learning models. this option is useful when training jobs can be interrupted and when there is flexibility when the training job is run.

The complete and intermediate results of jobs are stored in an Amazon S3 bucket, and can be used as a starting point to train models incrementally. Amazon SageMaker provides metrics and logs in CloudWatch. They can be used to see when managed spot training jobs are running, interrupted, resumed, or completed.

Type: Boolean
**EnableNetworkIsolation (p. 173)**

Isolates the training container. No inbound or outbound network calls can be made, except for calls between peers within a training cluster for distributed training. If you enable network isolation for training jobs that are configured to use a VPC, SageMaker downloads and uploads customer data and model artifacts through the specified VPC, but the training container does not have network access.

Type: Boolean

**Environment (p. 173)**

The environment variables to set in the Docker container.

Type: String to string map

Map Entries: Maximum number of 48 items.

Key Length Constraints: Maximum length of 512.

Key Pattern: `[a-zA-Z_]\[a-zA-Z0-9_]*`

Value Length Constraints: Maximum length of 512.

Value Pattern: `[\S\s]*`

**ExperimentConfig (p. 173)**

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:

- CreateProcessingJob (p. 162)
- CreateTrainingJob (p. 173)
- CreateTransformJob (p. 182)

Type: ExperimentConfig (p. 1029) object

**HyperParameters (p. 173)**

Algorithm-specific parameters that influence the quality of the model. You set hyperparameters before you start the learning process. For a list of hyperparameters for each training algorithm provided by SageMaker, see Algorithms.

You can specify a maximum of 100 hyperparameters. Each hyperparameter is a key-value pair. Each key and value is limited to 256 characters, as specified by the Length Constraint.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: `. *`

Value Length Constraints: Maximum length of 2500.

Value Pattern: `. *`
**InputDataConfig (p. 173)**

An array of Channel objects. Each channel is a named input source. InputDataConfig describes the input data and its location.

Algorithms can accept input data from one or more channels. For example, an algorithm might have two channels of input data, `training_data` and `validation_data`. The configuration for each channel provides the S3, EFS, or FSx location where the input data is stored. It also provides information about the stored data: the MIME type, compression method, and whether the data is wrapped in RecordIO format.

Depending on the input mode that the algorithm supports, SageMaker either copies input data files from an S3 bucket to a local directory in the Docker container, or makes it available as input streams. For example, if you specify an EFS location, input data files are available as input streams. They do not need to be downloaded.

Type: Array of Channel (p. 941) objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.

**OutputDataConfig (p. 173)**

Specifies the path to the S3 location where you want to store model artifacts. SageMaker creates subfolders for the artifacts.

Type: OutputDataConfig (p. 1244) object

Required: Yes

**ProfilerConfig (p. 173)**

Configuration information for Debugger system monitoring, framework profiling, and storage paths.

Type: ProfilerConfig (p. 1301) object

Required: No

**ProfilerRuleConfigurations (p. 173)**

Configuration information for Debugger rules for profiling system and framework metrics.

Type: Array of ProfilerRuleConfiguration (p. 1305) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

Required: No

**ResourceConfig (p. 173)**

The resources, including the ML compute instances and ML storage volumes, to use for model training.

ML storage volumes store model artifacts and incremental states. Training algorithms might also use ML storage volumes for scratch space. If you want SageMaker to use the ML storage volume to store the training data, choose `File` as the `TrainingInputMode` in the algorithm specification. For distributed training algorithms, specify an instance count greater than 1.

Type: ResourceConfig (p. 1338) object

Required: Yes
RetryStrategy (p. 173)

The number of times to retry the job when the job fails due to an InternalServerError.

Type: RetryStrategy (p. 1344) object

Required: No

RoleArn (p. 173)

The Amazon Resource Name (ARN) of an IAM role that SageMaker can assume to perform tasks on your behalf.

During model training, SageMaker needs your permission to read input data from an S3 bucket, download a Docker image that contains training code, write model artifacts to an S3 bucket, write logs to Amazon CloudWatch Logs, and publish metrics to Amazon CloudWatch. You grant permissions for all of these tasks to an IAM role. For more information, see SageMaker Roles.

Note
To be able to pass this role to SageMaker, the caller of this API must have the iam:PassRole permission.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_/%]+$

Required: Yes

StoppingCondition (p. 173)

Specifies a limit to how long a model training job can run. It also specifies how long a managed Spot training job has to complete. When the job reaches the time limit, SageMaker ends the training job. Use this API to cap model training costs.

To stop a job, SageMaker sends the algorithm the SIGTERM signal, which delays job termination for 120 seconds. Algorithms can use this 120-second window to save the model artifacts, so the results of training are not lost.

Type: StoppingCondition (p. 1370) object

Required: Yes

Tags (p. 173)

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

TensorBoardOutputConfig (p. 173)

Configuration of storage locations for the Debugger TensorBoard output data.

Type: TensorBoardOutputConfig (p. 1381) object

Required: No

TrainingJobName (p. 173)

The name of the training job. The name must be unique within an AWS Region in an AWS account.
CreateTrainingJob

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

VpcConfig (p. 173)

A VpcConfig (p. 1456) object that specifies the VPC that you want your training job to connect to. Control access to and from your training container by configuring the VPC. For more information, see Protect Training Jobs by Using an Amazon Virtual Private Cloud.

Type: VpcConfig (p. 1456) object

Required: No

Response Syntax

```
{
  "TrainingJobArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TrainingJobArn (p. 180)

The Amazon Resource Name (ARN) of the training job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:training-job/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

ResourceNotFound

Resource being access is not found.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateTransformJob

Service: Amazon SageMaker Service

Starts a transform job. A transform job uses a trained model to get inferences on a dataset and saves these results to an Amazon S3 location that you specify.

To perform batch transformations, you create a transform job and use the data that you have readily available.

In the request body, you provide the following:

- **TransformJobName** - Identifies the transform job. The name must be unique within an AWS Region in an AWS account.
- **ModelName** - Identifies the model to use. `ModelName` must be the name of an existing Amazon SageMaker model in the same AWS Region and AWS account. For information on creating a model, see CreateModel.
- **TransformInput** - Describes the dataset to be transformed and the Amazon S3 location where it is stored.
- **TransformOutput** - Identifies the Amazon S3 location where you want Amazon SageMaker to save the results from the transform job.
- **TransformResources** - Identifies the ML compute instances for the transform job.

For more information about how batch transformation works, see Batch Transform.

**Request Syntax**

```json
{
  "BatchStrategy": "string",
  "DataProcessing": {
    "InputFilter": "string",
    "JoinSource": "string",
    "OutputFilter": "string"
  },
  "Environment": {
    "string": "string"
  },
  "ExperimentConfig": {
    "ExperimentName": "string",
    "TrialComponentDisplayName": "string",
    "TrialName": "string"
  },
  "MaxConcurrentTransforms": number,
  "MaxPayloadInMB": number,
  "ModelClientConfig": {
    "InvocationsMaxRetries": number,
    "InvocationsTimeoutInSeconds": number
  },
  "ModelName": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "TransformInput": {
    "CompressionType": "string",
    "ContentType": "string",
    "DataSource": {
```
Amazon SageMaker Amazon Sagemaker API Reference
CreateTransformJob

```json
"S3DataSource": {
  "S3DataType": "string",
  "S3Uri": "string"
},
"SplitType": "string",
"TransformJobName": "string",
"TransformOutput": {
  "Accept": "string",
  "AssembleWith": "string",
  "KmsKeyId": "string",
  "S3OutputPath": "string"
},
"TransformResources": {
  "InstanceCount": number,
  "InstanceType": "string",
  "VolumeKmsKeyId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**BatchStrategy (p. 182)**

Specifies the number of records to include in a mini-batch for an HTTP inference request. A record is a single unit of input data that inference can be made on. For example, a single line in a CSV file is a record.

To enable the batch strategy, you must set the SplitType property to Line, RecordIO, or TFRecord.

To use only one record when making an HTTP invocation request to a container, set BatchStrategy to SingleRecord and SplitType to Line.

To fit as many records in a mini-batch as can fit within the MaxPayloadInMB limit, set BatchStrategy to MultiRecord and SplitType to Line.

Type: String

Valid Values: MultiRecord | SingleRecord

Required: No

**DataProcessing (p. 182)**

The data structure used to specify the data to be used for inference in a batch transform job and to associate the data that is relevant to the prediction results in the output. The input filter provided allows you to exclude input data that is not needed for inference in a batch transform job. The output filter provided allows you to include input data relevant to interpreting the predictions in the output from the job. For more information, see Associate Prediction Results with their Corresponding Input Records.

Type: DataProcessing (p. 971) object

Required: No
**Environment (p. 182)**

The environment variables to set in the Docker container. We support up to 16 key and values entries in the map.

Type: String to string map

Map Entries: Maximum number of 16 items.

Key Length Constraints: Maximum length of 1024.

Key Pattern: `[a-zA-Z_][a-zA-Z0-9_]{0,1023}`

Value Length Constraints: Maximum length of 10240.

Value Pattern: `[^\s]*`

Required: No

**ExperimentConfig (p. 182)**

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:

- CreateProcessingJob (p. 162)
- CreateTrainingJob (p. 173)
- CreateTransformJob (p. 182)

Type: ExperimentConfig (p. 1029) object

Required: No

**MaxConcurrentTransforms (p. 182)**

The maximum number of parallel requests that can be sent to each instance in a transform job. If `MaxConcurrentTransforms` is set to 0 or left unset, Amazon SageMaker checks the optional execution-parameters to determine the settings for your chosen algorithm. If the execution-parameters endpoint is not enabled, the default value is 1. For more information on execution-parameters, see How Containers Serve Requests. For built-in algorithms, you don’t need to set a value for `MaxConcurrentTransforms`.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**MaxPayloadInMB (p. 182)**

The maximum allowed size of the payload, in MB. A payload is the data portion of a record (without metadata). The value in `MaxPayloadInMB` must be greater than, or equal to, the size of a single record. To estimate the size of a record in MB, divide the size of your dataset by the number of records. To ensure that the records fit within the maximum payload size, we recommend using a slightly larger value. The default value is 6 MB.

The value of `MaxPayloadInMB` cannot be greater than 100 MB. If you specify the `MaxConcurrentTransforms` parameter, the value of `(MaxConcurrentTransforms * MaxPayloadInMB)` also cannot exceed 100 MB.

For cases where the payload might be arbitrarily large and is transmitted using HTTP chunked encoding, set the value to 0. This feature works only in supported algorithms. Currently, Amazon SageMaker built-in algorithms do not support HTTP chunked encoding.
CreateTransformJob

Type: Integer

Valid Range: Minimum value of 0.

Required: No

ModelClientConfig (p. 182)

Configures the timeout and maximum number of retries for processing a transform job invocation.

Type: ModelClientConfig (p. 1153) object

Required: No

ModelName (p. 182)

The name of the model that you want to use for the transform job. ModelName must be the name of an existing Amazon SageMaker model within an AWS Region in an AWS account.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*

Required: Yes

Tags (p. 182)

(Optional) An array of key-value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

TransformInput (p. 182)

Describes the input source and the way the transform job consumes it.

Type: TransformInput (p. 1403) object

Required: Yes

TransformJobName (p. 182)

The name of the transform job. The name must be unique within an AWS Region in an AWS account.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0, 62\}

Required: Yes

TransformOutput (p. 182)

Describes the results of the transform job.

Type: TransformOutput (p. 1415) object

Required: Yes
TransformResources (p. 182)

Describes the resources, including ML instance types and ML instance count, to use for the transform job.

Type: TransformResources (p. 1417) object

Required: Yes

Response Syntax

```json
{
  "TransformJobArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TransformJobArn (p. 186)

The Amazon Resource Name (ARN) of the transform job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:transform-job/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
CreateTrial
Service: Amazon SageMaker Service

Creates an SageMaker trial. A trial is a set of steps called trial components that produce a machine learning model. A trial is part of a single SageMaker experiment.

When you use SageMaker Studio or the SageMaker Python SDK, all experiments, trials, and trial components are automatically tracked, logged, and indexed. When you use the AWS SDK for Python (Boto), you must use the logging APIs provided by the SDK.

You can add tags to a trial and then use the Search (p. 691) API to search for the tags.

To get a list of all your trials, call the ListTrials (p. 665) API. To view a trial's properties, call the DescribeTrial (p. 464) API. To create a trial component, call the CreateTrialComponent (p. 191) API.

Request Syntax

```
{
  "DisplayName": "string",
  "ExperimentName": "string",
  "MetadataProperties": {
    "CommitId": "string",
    "GeneratedBy": "string",
    "ProjectId": "string",
    "Repository": "string"
  },
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "TrialName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

DisplayName (p. 188)

The name of the trial as displayed. The name doesn't need to be unique. If DisplayName isn't specified, TrialName is displayed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9\-]{0,119}$

Required: No

ExperimentName (p. 188)

The name of the experiment to associate the trial with.

Type: String
CreateTrial

Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$
Required: Yes

MetadataProperties (p. 188)
Metadata properties of the tracking entity, trial, or trial component.
Type: MetadataProperties (p. 1143) object
Required: No

Tags (p. 188)
A list of tags to associate with the trial. You can use Search (p. 691) API to search on the tags.
Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

TrialName (p. 188)
The name of the trial. The name must be unique in your AWS account and is not case-sensitive.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$
Required: Yes

Response Syntax

```{ "TrialArn": "string" }
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

TrialArn (p. 189)
The Amazon Resource Name (ARN) of the trial.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateTrialComponent
Service: Amazon SageMaker Service

Creates a trial component, which is a stage of a machine learning trial. A trial is composed of one or more trial components. A trial component can be used in multiple trials.

Trial components include pre-processing jobs, training jobs, and batch transform jobs.

When you use SageMaker Studio or the SageMaker Python SDK, all experiments, trials, and trial components are automatically tracked, logged, and indexed. When you use the AWS SDK for Python (Boto), you must use the logging APIs provided by the SDK.

You can add tags to a trial component and then use the Search (p. 691) API to search for the tags.

Request Syntax

```
{
    "DisplayName": "string",
    "EndTime": number,
    "InputArtifacts": {
        "string": {
            "MediaType": "string",
            "Value": "string"
        }
    },
    "MetadataProperties": {
        "CommitId": "string",
        "GeneratedBy": "string",
        "ProjectId": "string",
        "Repository": "string"
    },
    "OutputArtifacts": {
        "string": {
            "MediaType": "string",
            "Value": "string"
        }
    },
    "Parameters": {
        "string": {
            "NumberValue": number,
            "StringValue": "string"
        }
    },
    "StartTime": number,
    "Status": {
        "Message": "string",
        "PrimaryStatus": "string"
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
    "TrialComponentName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).
The request accepts the following data in JSON format.

**DisplayName (p. 191)**

The name of the component as displayed. The name doesn't need to be unique. If `displayName` isn't specified, `TrialComponentName` is displayed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]{0,119}){0,119}`

Required: No

**EndTime (p. 191)**

When the component ended.

Type: Timestamp

Required: No

**InputArtifacts (p. 191)**

The input artifacts for the component. Examples of input artifacts are datasets, algorithms, hyperparameters, source code, and instance types.

Type: String to `TrialComponentArtifact (p. 1428)` object map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 64.

Key Pattern: `.*`

Required: No

**MetadataProperties (p. 191)**

Metadata properties of the tracking entity, trial, or trial component.

Type: `MetadataProperties (p. 1143)` object

Required: No

**OutputArtifacts (p. 191)**

The output artifacts for the component. Examples of output artifacts are metrics, snapshots, logs, and images.

Type: String to `TrialComponentArtifact (p. 1428)` object map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 64.

Key Pattern: `.*`

Required: No

**Parameters (p. 191)**

The hyperparameters for the component.

Type: String to `TrialComponentParameterValue (p. 1431)` object map
Map Entries: Maximum number of 150 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: .*

Required: No

**StartTime (p. 191)**

When the component started.

Type: Timestamp

Required: No

**Status (p. 191)**

The status of the component. States include:

- InProgress
- Completed
- Failed

Type: `TrialComponentStatus (p. 1436)` object

Required: No

**Tags (p. 191)**

A list of tags to associate with the component. You can use `Search (p. 691)` API to search on the tags.

Type: Array of `Tag (p. 1377)` objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**TrialComponentName (p. 191)**

The name of the component. The name must be unique in your AWS account and is not case-sensitive.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9]([-]*[a-zA-Z0-9]){0,119}`

Required: Yes

**Response Syntax**

```json
{
  "TrialComponentArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**TrialComponentArn (p. 193)**

The Amazon Resource Name (ARN) of the trial component.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial-component/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateUserProfile
Service: Amazon SageMaker Service

Creates a user profile. A user profile represents a single user within a domain, and is the main way to reference a "person" for the purposes of sharing, reporting, and other user-oriented features. This entity is created when a user onboards to Amazon SageMaker Studio. If an administrator invites a person by email or imports them from SSO, a user profile is automatically created. A user profile is the primary holder of settings for an individual user and has a reference to the user's private Amazon Elastic File System (EFS) home directory.

Request Syntax

```
{
  "DomainId": "string",
  "SingleSignOnUserIdentifier": "string",
  "SingleSignOnUserValue": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "UserProfileName": "string",
  "UserSettings": {
    "ExecutionRole": "string",
    "JupyterServerAppSettings": {
      "DefaultResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      },
      "LifecycleConfigArns": [ "string" ]
    },
    "KernelGatewayAppSettings": {
      "CustomImages": [
        {
          "AppImageConfigName": "string",
          "ImageName": "string",
          "ImageVersionNumber": number
        }
      ],
      "DefaultResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      },
      "LifecycleConfigArns": [ "string" ]
    },
    "RSessionAppSettings": {
      "CustomImages": [
        {
          "AppImageConfigName": "string",
          "ImageName": "string",
          "ImageVersionNumber": number
        }
      ],
      "DefaultResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      }
    }
  }
}
```
"SageMakerImageVersionArn": "string"
},
"RStudioServerProAppSettings": {
  "AccessStatus": "string",
  "UserGroup": "string"
},
"SecurityGroups": [ "string" ],
"SharingSettings": {
  "NotebookOutputOption": "string",
  "S3KmsKeyId": "string",
  "S3OutputPath": "string"
},
"TensorBoardAppSettings": {
  "DefaultResourceSpec": {
    "InstanceType": "string",
    "LifecycleConfigArn": "string",
    "SageMakerImageArn": "string",
    "SageMakerImageVersionArn": "string"
  }
}

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainId (p. 195)**

The ID of the associated Domain.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**SingleSignOnUserIdentifier (p. 195)**

A specifier for the type of value specified in SingleSignOnUserValue. Currently, the only supported value is "UserName". If the Domain's AuthMode is SSO, this field is required. If the Domain's AuthMode is not SSO, this field cannot be specified.

Type: String

Pattern: UserName

Required: No

**SingleSignOnUserValue (p. 195)**

The username of the associated AWS Single Sign-On User for this UserProfile. If the Domain's AuthMode is SSO, this field is required, and must match a valid username of a user in your directory. If the Domain's AuthMode is not SSO, this field cannot be specified.

Type: String

Length Constraints: Maximum length of 256.
Required: No

**Tags (p. 195)**

Each tag consists of a key and an optional value. Tag keys must be unique per resource.

Tags that you specify for the User Profile are also added to all Apps that the User Profile launches.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**UserProfileName (p. 195)**

A name for the UserProfile. This value is not case sensitive.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]{0,62}

Required: Yes

**UserSettings (p. 195)**

A collection of settings.

Type: UserSettings (p. 1452) object

Required: No

**Response Syntax**

```
{
  "UserProfileArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**UserProfileArn (p. 197)**

The user profile Amazon Resource Name (ARN).

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:user-profile/.*

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceLimitExceeded

You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateWorkforce
Service: Amazon SageMaker Service

Use this operation to create a workforce. This operation will return an error if a workforce already exists in the AWS Region that you specify. You can only create one workforce in each AWS Region per AWS account.

If you want to create a new workforce in an AWS Region where a workforce already exists, use the DeleteWorkforce (p. 273) API operation to delete the existing workforce and then use CreateWorkforce to create a new workforce.

To create a private workforce using Amazon Cognito, you must specify a Cognito user pool in CognitoConfig. You can also create an Amazon Cognito workforce using the Amazon SageMaker console. For more information, see Create a Private Workforce (Amazon Cognito).

To create a private workforce using your own OIDC Identity Provider (IdP), specify your IdP configuration in OidcConfig. Your OIDC IdP must support groups because groups are used by Ground Truth and Amazon A2I to create work teams. For more information, see Create a Private Workforce (OIDC IdP).

Request Syntax

```json
{
  "CognitoConfig": {
    "ClientId": "string",
    "UserPool": "string"
  },
  "OidcConfig": {
    "AuthorizationEndpoint": "string",
    "ClientId": "string",
    "ClientSecret": "string",
    "Issuer": "string",
    "JwksUri": "string",
    "LogoutEndpoint": "string",
    "TokenEndpoint": "string",
    "UserInfoEndpoint": "string"
  },
  "SourceIpConfig": {
    "Cidrs": [ "string" ]
  },
  "Tags": [
    { "Key": "string",
      "Value": "string"
    }
  ],
  "WorkforceName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CognitoConfig (p. 199)**

Use this parameter to configure an Amazon Cognito private workforce. A single Cognito workforce is created using and corresponds to a single Amazon Cognito user pool.
CreateWorkforce

Do not use OidcConfig if you specify values for CognitoConfig.

Type: CognitoConfig (p. 950) object

Required: No

OidcConfig (p. 199)

Use this parameter to configure a private workforce using your own OIDC Identity Provider.

Do not use CognitoConfig if you specify values for OidcConfig.

Type: OidcConfig (p. 1233) object

Required: No

SourceIpConfig (p. 199)

A list of IP address ranges (CIDRs). Used to create an allow list of IP addresses for a private workforce. Workers will only be able to login to their worker portal from an IP address within this range. By default, a workforce isn't restricted to specific IP addresses.

Type: SourceIpConfig (p. 1369) object

Required: No

Tags (p. 199)

An array of key-value pairs that contain metadata to help you categorize and organize our workforce. Each tag consists of a key and a value, both of which you define.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

WorkforceName (p. 199)

The name of the private workforce.

Type: String


Pattern: ^[a-zA-Z0-9](^[a-zA-Z0-9\-]){0,62}$

Required: Yes

Response Syntax

```json
{
  "WorkforceArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
**WorkforceArn (p. 200)**

The Amazon Resource Name (ARN) of the workforce.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workforce/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
CreateWorkteam

Service: Amazon SageMaker Service

Creates a new work team for labeling your data. A work team is defined by one or more Amazon Cognito user pools. You must first create the user pools before you can create a work team.

You cannot create more than 25 work teams in an account and region.

Request Syntax

```
{
    "Description": "string",
    "MemberDefinitions": [
        {
            "CognitoMemberDefinition": {
                "ClientId": "string",
                "UserGroup": "string",
                "UserPool": "string"
            },
            "OidcMemberDefinition": {
                "Groups": [ "string" ]
            }
        }
    ],
    "NotificationConfiguration": {
        "NotificationTopicArn": "string"
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
    "WorkforceName": "string",
    "WorkteamName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**Description (p. 202)**

A description of the work team.

Type: String


Pattern: .+

Required: Yes

**MemberDefinitions (p. 202)**

A list of MemberDefinition objects that contains objects that identify the workers that make up the work team.
Workforces can be created using Amazon Cognito or your own OIDC Identity Provider (IdP). For private workforces created using Amazon Cognito use `CognitoMemberDefinition`. For workforces created using your own OIDC identity provider (IdP) use `OidcMemberDefinition`. Do not provide input for both of these parameters in a single request.

For workforces created using Amazon Cognito, private work teams correspond to Amazon Cognito user groups within the user pool used to create a workforce. All of the `CognitoMemberDefinition` objects that make up the member definition must have the same `ClientId` and `UserPool` values. To add a Amazon Cognito user group to an existing worker pool, see Adding groups to a User Pool. For more information about user pools, see Amazon Cognito User Pools.

For workforces created using your own OIDC IdP, specify the user groups that you want to include in your private work team in `OidcMemberDefinition` by listing those groups in `Groups`.

Type: Array of `MemberDefinition` objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Required: Yes

**NotificationConfiguration (p. 202)**

Configures notification of workers regarding available or expiring work items.

Type: `NotificationConfiguration` object

Required: No

**Tags (p. 202)**

An array of key-value pairs.

For more information, see Resource Tag and Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of `Tag` objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**WorkforceName (p. 202)**

The name of the workforce.

Type: String


Pattern: `^[a-zA-Z0-9\-\[]{0,62}$`

Required: No

**WorkteamName (p. 202)**

The name of the work team. Use this name to identify the work team.

Type: String


Pattern: `^[a-zA-Z0-9\-\_]*[a-zA-Z0-9\-\_]{0,62}$`

Required: Yes
Response Syntax

```
{
   "WorkteamArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**WorkteamArn (p. 204)**

The Amazon Resource Name (ARN) of the work team. You can use this ARN to identify the work team.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[\da-z\/-]*:sagemaker:[\da-z0-9\/-]*:[0-9]{12}:workteam/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java
- AWS SDK for JavaScript
- AWS SDK for PHP
- AWS SDK for Python
- AWS SDK for Ruby
DeleteAction
Service: Amazon SageMaker Service

Deletes an action.

Request Syntax

```json
{
   "ActionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ActionName (p. 205)**

- The name of the action to delete.
- Type: String
- Length Constraints: Minimum length of 1. Maximum length of 120.
- Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}
- Required: Yes

Response Syntax

```json
{
   "ActionArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ActionArn (p. 205)**

- The Amazon Resource Name (ARN) of the action.
- Type: String
- Length Constraints: Maximum length of 256.
- Pattern: arn:aws[a-z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:action/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteAlgorithm
Service: Amazon SageMaker Service

Removes the specified algorithm from your account.

Request Syntax

```
{
    "AlgorithmName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AlgorithmName (p. 207)

The name of the algorithm to delete.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteApp
Service: Amazon SageMaker Service

Used to stop and delete an app.

Request Syntax

```json
{
    "AppName": "string",
    "AppType": "string",
    "DomainId": "string",
    "UserProfileName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AppName (p. 208)

The name of the app.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

AppType (p. 208)

The type of app.

Type: String

Valid Values: JupyterServer | KernelGateway | TensorBoard | RStudioServerPro | RSessionGateway

Required: Yes

DomainId (p. 208)

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

UserProfileName (p. 208)

The user profile name.

Type: String

Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\$

Required: Yes

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](https://aws.amazon.com/documentation/sagemaker/).  

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteAppImageConfig

Service: Amazon SageMaker Service

Deletes an AppImageConfig.

Request Syntax

```
{
  "AppImageConfigName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AppImageConfigName (p. 210)**

The name of the AppImageConfig to delete.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DeleteArtifact
Service: Amazon SageMaker Service

Deletes an artifact. Either ArtifactArn or Source must be specified.

Request Syntax

```
{
    "ArtifactArn": "string",
    "Source": {
        "SourceTypes": [
            {
                "SourceIdType": "string",
                "Value": "string"
            }
        ],
        "SourceUri": "string"
    }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ArtifactArn (p. 212)

The Amazon Resource Name (ARN) of the artifact to delete.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:artifact/.*`

Required: No

Source (p. 212)

The URI of the source.

Type: ArtifactSource (p. 892) object

Required: No

Response Syntax

```
{
    "ArtifactArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**ArtifactArn (p. 212)**

The Amazon Resource Name (ARN) of the artifact.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:artifact/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteAssociation
Service: Amazon SageMaker Service

Deletes an association.

Request Syntax

```
{
   "DestinationArn": "string",
   "SourceArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DestinationArn (p. 214)**

The Amazon Resource Name (ARN) of the destination.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[\w-]*:sagemaker:[\w-]+:[0-9]{12}:\{experiment|experiment-trial-component|artifact|action|context\}/.*`

Required: Yes

**SourceArn (p. 214)**

The ARN of the source.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[\w-]*:sagemaker:[\w-]+:[0-9]{12}:\{experiment|experiment-trial-component|artifact|action|context\}/.*`

Required: Yes

Response Syntax

```
{
   "DestinationArn": "string",
   "SourceArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
**DestinationArn (p. 214)**

The Amazon Resource Name (ARN) of the destination.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

**SourceArn (p. 214)**

The ARN of the source.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteCodeRepository

Service: Amazon SageMaker Service

Deletes the specified Git repository from your account.

Request Syntax

```json
{
   "CodeRepositoryName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CodeRepositoryName (p. 216)**

The name of the Git repository to delete.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteContext
Service: Amazon SageMaker Service
Deletes an context.

Request Syntax

```
{
    "ContextName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ContextName (p. 217)**

The name of the context to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: Yes

Response Syntax

```
{
    "ContextArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ContextArn (p. 217)**

The Amazon Resource Name (ARN) of the context.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:context/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteDataQualityJobDefinition

Service: Amazon SageMaker Service

Deletes a data quality monitoring job definition.

Request Syntax

```json
{
  "JobDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 219)**

- The name of the data quality monitoring job definition to delete.
- Type: String
- Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*$`
- Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

- Resource being access is not found.
- HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteDeviceFleet
Service: Amazon SageMaker Service

Deletes a fleet.

Request Syntax

```json
{
  "DeviceFleetName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeviceFleetName (p. 221)**

The name of the fleet to delete.

Type: String

- Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

- Resource being accessed is in use.
- HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
DeleteDeviceFleet

- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteDomain
Service: Amazon SageMaker Service

Used to delete a domain. If you onboarded with IAM mode, you will need to delete your domain to onboard again using SSO. Use with caution. All of the members of the domain will lose access to their EFS volume, including data, notebooks, and other artifacts.

Request Syntax

```json
{
   "DomainId": "string",
   "RetentionPolicy": {
      "HomeEfsFileSystem": "string"
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainId (p. 223)**

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**RetentionPolicy (p. 223)**

The retention policy for this domain, which specifies whether resources will be retained after the Domain is deleted. By default, all resources are retained (not automatically deleted).

Type: RetentionPolicy (p. 1343) object

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.
HTTP Status Code: 400

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteEndpoint

Service: Amazon SageMaker Service

Deletes an endpoint. SageMaker frees up all of the resources that were deployed when the endpoint was created.

SageMaker retires any custom KMS key grants associated with the endpoint, meaning you don't need to use the RevokeGrant API call.

When you delete your endpoint, SageMaker asynchronously deletes associated endpoint resources such as KMS key grants. You might still see these resources in your account for a few minutes after deleting your endpoint. Do not delete or revoke the permissions for your ExecutionRoleArn, otherwise SageMaker cannot delete these resources.

Request Syntax

```
{
  "EndpointName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**EndpointName (p. 225)**

The name of the endpoint that you want to delete.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-\[\]a-zA-Z0-9\)\]+\[a-zA-Z0-9\-\[\]a-zA-Z0-9\(\]+{0,62}

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DeleteEndpointConfig
Service: Amazon SageMaker Service

Deletes an endpoint configuration. The DeleteEndpointConfig API deletes only the specified configuration. It does not delete endpoints created using the configuration.

You must not delete an EndpointConfig in use by an endpoint that is live or while the UpdateEndpoint or CreateEndpoint operations are being performed on the endpoint. If you delete the EndpointConfig of an endpoint that is active or being created or updated you may lose visibility into the instance type the endpoint is using. The endpoint must be deleted in order to stop incurring charges.

Request Syntax

```json
{
   "EndpointConfigName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

EndpointConfigName (p. 227)

The name of the endpoint configuration that you want to delete.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteExperiment
Service: Amazon SageMaker Service

Deletes an SageMaker experiment. All trials associated with the experiment must be deleted first. Use the ListTrials (p. 665) API to get a list of the trials associated with the experiment.

Request Syntax

```json
{
   "ExperimentName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ExperimentName (p. 229)**

The name of the experiment to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: Yes

Response Syntax

```json
{
   "ExperimentArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ExperimentArn (p. 229)**

The Amazon Resource Name (ARN) of the experiment that is being deleted.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteFeatureGroup

Service: Amazon SageMaker Service

Delete the FeatureGroup and any data that was written to the OnlineStore of the FeatureGroup. Data cannot be accessed from the OnlineStore immediately after DeleteFeatureGroup is called.

Data written into the OfflineStore will not be deleted. The AWS Glue database and tables that are automatically created for your OfflineStore are not deleted.

Request Syntax

```json
{
    "FeatureGroupName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

FeatureGroupName (p. 231)

The name of the FeatureGroup you want to delete. The name must be unique within an AWS Region in an AWS account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9](^-*[a-zA-Z0-9]){0,63}$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteFlowDefinition

Service: Amazon SageMaker Service

Deletion the specified flow definition.

Request Syntax

```json
{
    "FlowDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**FlowDefinitionName (p. 233)**

The name of the flow definition you are deleting.

Type: String


Pattern: ^[a-z0-9](-*[a-z0-9]){0,62}

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DeleteHumanTaskUi
Service: Amazon SageMaker Service

Use this operation to delete a human task user interface (worker task template).

To see a list of human task user interfaces (work task templates) in your account, use ListHumanTaskUis (p. 562). When you delete a worker task template, it no longer appears when you call ListHumanTaskUis.

Request Syntax

```
{
   "HumanTaskUiName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**HumanTaskUiName** (p. 235)

The name of the human task user interface (work task template) you want to delete.

Type: String


Pattern: \^[a-z0-9](-*[a-z0-9])*\n
Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteImage

Service: Amazon SageMaker Service

Deletes a SageMaker image and all versions of the image. The container images aren't deleted.

Request Syntax

```json
{
   "ImageName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ImageName (p. 237)

The name of the image to delete.

Type: String


Pattern: `^[a-zA-Z0-9][\-.]?[a-zA-Z0-9]{0,62}$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DeleteImageVersion
Service: Amazon SageMaker Service

Deletes a version of a SageMaker image. The container image the version represents isn't deleted.

Request Syntax

```
{
   "ImageName": "string",
   "Version": number
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ImageName (p. 239)**

- The name of the image.
- Type: String
- Pattern: ^[a-zA-Z0-9][-._]?[a-zA-Z0-9]{0,62}$
- Required: Yes

**Version (p. 239)**

- The version to delete.
- Type: Integer
- Valid Range: Minimum value of 0.
- Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

- Resource being accessed is in use.
- HTTP Status Code: 400

**ResourceNotFound**

- Resource being access is not found.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteModel
Service: Amazon SageMaker Service

Deletes a model. The `DeleteModel` API deletes only the model entry that was created in SageMaker when you called the `CreateModel` API. It does not delete model artifacts, inference code, or the IAM role that you specified when creating the model.

Request Syntax

```
{
  "ModelName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

`ModelName` (p. 241)

The name of the model to delete.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*`  

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteModelBiasJobDefinition
Service: Amazon SageMaker Service

Deletes an Amazon SageMaker model bias job definition.

Request Syntax

```
{
  "JobDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 243)**

The name of the model bias job definition to delete.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFoundException**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DeleteModelExplainabilityJobDefinition

Service: Amazon SageMaker Service

Deletes an Amazon SageMaker model explainability job definition.

Request Syntax

```
{
   "JobDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 245)**

The name of the model explainability job definition to delete.

Type: String


Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9])*#

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteModelPackage

Service: Amazon SageMaker Service

Deletes a model package.

A model package is used to create SageMaker models or list on AWS Marketplace. Buyers can subscribe to model packages listed on AWS Marketplace to create models in SageMaker.

Request Syntax

```
{
    "ModelPackageName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ModelPackageName (p. 247)**

The name or Amazon Resource Name (ARN) of the model package to delete.

When you specify a name, the name must have 1 to 63 characters. Valid characters are a-z, A-Z, 0-9, and -(hyphen).

Type: String


Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:[a-z\-]*\//)?([a-zA-Z0-9][a-zA-Z0-9-]*)([a-zA-Z0-9](\/[0-9]{1,5})?)?

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ConflictException**

There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
DeleteModelPackage

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteModelPackageGroup
Service: Amazon SageMaker Service

Deletes the specified model group.

Request Syntax

```json
{
   "ModelPackageGroupName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ModelPackageGroupName (p. 249)

The name of the model group to delete.

Type: String


Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:\[a-z-]*\/)?(\[a-zA-Z0-9-\]\{0,62\})(?<!-)$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ConflictException

There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DeleteModelPackageGroupPolicy

Service: Amazon SageMaker Service

Deletes a model group resource policy.

Request Syntax

```json
{
    "ModelPackageGroupName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ModelPackageGroupName (p. 251)**

The name of the model group for which to delete the policy.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteModelQualityJobDefinition
Service: Amazon SageMaker Service

Deletes the specified model quality monitoring job definition.

Request Syntax

```json
{
   "JobDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName** (p. 252)

The name of the model quality monitoring job definition to delete.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFoundException**

Resource being accessed is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteMonitoringSchedule
Service: Amazon SageMaker

Deletes a monitoring schedule. Also stops the schedule had not already been stopped. This does not delete the job execution history of the monitoring schedule.

Request Syntax

```json
{
   "MonitoringScheduleName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

MonitoringScheduleName (p. 254)

The name of the monitoring schedule to delete.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteNotebookInstance
Service: Amazon SageMaker Service

Deletes an SageMaker notebook instance. Before you can delete a notebook instance, you must call the StopNotebookInstance API.

**Important**
When you delete a notebook instance, you lose all of your data. SageMaker removes the ML compute instance, and deletes the ML storage volume and the network interface associated with the notebook instance.

**Request Syntax**

```json
{
   "NotebookInstanceName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**NotebookInstanceName (p. 256)**

The name of the SageMaker notebook instance to delete.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*`

Required: Yes

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
AWS SDK for PHP V3
AWS SDK for Python
AWS SDK for Ruby V3
DeleteNotebookInstanceLifecycleConfig
Service: Amazon SageMaker Service
Deletion a notebook instance lifecycle configuration.

Request Syntax

```json
{
   "NotebookInstanceLifecycleConfigName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

NotebookInstanceLifecycleConfigName (p. 258)

The name of the lifecycle configuration to delete.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]*$
Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeletePipeline
Service: Amazon SageMaker Service

Deletes a pipeline if there are no running instances of the pipeline. To delete a pipeline, you must stop all running instances of the pipeline using the StopPipelineExecution API. When you delete a pipeline, all instances of the pipeline are deleted.

Request Syntax

```json
{
    "ClientRequestToken": "string",
    "PipelineName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ClientRequestToken (p. 259)

A unique, case-sensitive identifier that you provide to ensure the idempotency of the operation. An idempotent operation completes no more than one time.

Type: String


Required: Yes

PipelineName (p. 259)

The name of the pipeline to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: ^[a-zA-Z0-9-\[a-zA-Z0-9\]+\[a-zA-Z0-9\]+\]0,255

Required: Yes

Response Syntax

```json
{
    "PipelineArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
PipelineArn (p. 259)

The Amazon Resource Name (ARN) of the pipeline to delete.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteProject

Service: Amazon SageMaker Service

Delete the specified project.

Request Syntax

```json
{
    "ProjectName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ProjectName (p. 261)**

- The name of the project to delete.
- Type: String
- Pattern: `^[a-zA-Z0-9][-]*[a-zA-Z0-9]{0,31}`
- Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ConflictException**

- There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.
- HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteStudioLifecycleConfig
Service: Amazon SageMaker Service

Deletes the Studio Lifecycle Configuration. In order to delete the Lifecycle Configuration, there must be no running apps using the Lifecycle Configuration. You must also remove the Lifecycle Configuration from UserSettings in all Domains and UserProfiles.

Request Syntax

```
{
  "StudioLifecycleConfigName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**StudioLifecycleConfigName (p. 263)**

The name of the Studio Lifecycle Configuration to delete.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9-]*[a-zA-Z0-9]{0,62}

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteTags

Service: Amazon SageMaker Service

Deletes the specified tags from an SageMaker resource.

To list a resource's tags, use the ListTags API.

**Note**
When you call this API to delete tags from a hyperparameter tuning job, the deleted tags are not removed from training jobs that the hyperparameter tuning job launched before you called this API.

**Note**
When you call this API to delete tags from a SageMaker Studio Domain or User Profile, the deleted tags are not removed from Apps that the SageMaker Studio Domain or User Profile launched before you called this API.

Request Syntax

```json
{
    "ResourceArn": "string",
    "TagKeys": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ResourceArn (p. 265)**

  The Amazon Resource Name (ARN) of the resource whose tags you want to delete.

  Type: String

  Length Constraints: Maximum length of 256.

  Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:.+

  Required: Yes

**TagKeys (p. 265)**

  An array or one or more tag keys to delete.

  Type: Array of strings

  Array Members: Minimum number of 1 item. Maximum number of 50 items.


  Pattern: ^([\p{L}\p{Z}\p{N}\p{S}]_::+=@\-])$.

  Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.
Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteTrial
Service: Amazon SageMaker Service

Deletes the specified trial. All trial components that make up the trial must be deleted first. Use the DescribeTrialComponent (p. 467) API to get the list of trial components.

Request Syntax

```
{
    "TrialName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**TrialName (p. 267)**

The name of the trial to delete.

- Type: String
- Length Constraints: Minimum length of 1. Maximum length of 120.
- Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}`
- Required: Yes

Response Syntax

```
{
    "TrialArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**TrialArn (p. 267)**

The Amazon Resource Name (ARN) of the trial that is being deleted.

- Type: String
- Length Constraints: Maximum length of 256.
- Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial/.*`

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteTrialComponent

Service: Amazon SageMaker Service

Deletes the specified trial component. A trial component must be disassociated from all trials before the trial component can be deleted. To disassociate a trial component from a trial, call the DisassociateTrialComponent (p. 481) API.

Request Syntax

```json
{
  "TrialComponentName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**TrialComponentName (p. 269)**

  The name of the component to delete.

  Type: String

  Length Constraints: Minimum length of 1. Maximum length of 120.

  Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}`

  Required: Yes

Response Syntax

```json
{
  "TrialComponentArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**TrialComponentArn (p. 269)**

  The Amazon Resource Name (ARN) of the component is being deleted.

  Type: String

  Length Constraints: Maximum length of 256.

  Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:\[0-9\]{12}:experiment-trial-component/.*`
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

  Resource being access is not found.

  HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteUserProfile
Service: Amazon SageMaker Service

Deletes a user profile. When a user profile is deleted, the user loses access to their EFS volume, including data, notebooks, and other artifacts.

Request Syntax

```json
{
    "DomainId": "string",
    "UserProfileName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainId (p. 271)**

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**UserProfileName (p. 271)**

The user profile name.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeleteWorkforce

Service: Amazon SageMaker Service

Use this operation to delete a workforce.

If you want to create a new workforce in an AWS Region where a workforce already exists, use this operation to delete the existing workforce and then use CreateWorkforce (p. 199) to create a new workforce.

**Important**
If a private workforce contains one or more work teams, you must use the DeleteWorkteam (p. 275) operation to delete all work teams before you delete the workforce. If you try to delete a workforce that contains one or more work teams, you will recieve a ResourceInUse error.

**Request Syntax**

```json
{
  "WorkforceName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**WorkforceName (p. 273)**

  The name of the workforce.
  Type: String
  Pattern: `^[a-zA-Z0-9\-]{0,62}$`
  Required: Yes

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DeleteWorkteam
Service: Amazon SageMaker Service
Deletes an existing work team. This operation can't be undone.

Request Syntax

```json
{
   "WorkteamName": "string"
}
```

Request Parameters
For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**WorkteamName (p. 275)**

- The name of the work team to delete.
- Type: String
- Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`
- Required: Yes

Response Syntax

```json
{
   "Success": boolean
}
```

Response Elements
If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**Success (p. 275)**

- Returns true if the work team was successfully deleted; otherwise, returns false.
- Type: Boolean

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

- You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DeregisterDevices

Service: Amazon SageMaker Service

Deregisters the specified devices. After you deregister a device, you will need to re-register the devices.

Request Syntax

```
{
   "DeviceFleetName": "string",
   "DeviceNames": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeviceFleetName (p. 277)**

The name of the fleet the devices belong to.

Type: String


Pattern: `^[a-zA-Z0-9]([-\[a-zA-Z0-9]\]+)\{0,62}$`

Required: Yes

**DeviceNames (p. 277)**

The unique IDs of the devices.

Type: Array of strings


Pattern: `^[a-zA-Z0-9]([-\[a-zA-Z0-9]\]+)\{0,62}$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
**DescribeAction**
Service: Amazon SageMaker Service

Describes an action.

**Request Syntax**

```json
{
   "ActionName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ActionName (p. 279)**

The name of the action to describe.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$`

Required: Yes

**Response Syntax**

```json
{
   "ActionArn": "string",
   "ActionName": "string",
   "ActionType": "string",
   "CreatedBy": {
      "DomainId": "string",
      "UserProfileArn": "string",
      "UserProfileName": "string"
   },
   "CreationTime": number,
   "Description": "string",
   "LastModifiedBy": {
      "DomainId": "string",
      "UserProfileArn": "string",
      "UserProfileName": "string"
   },
   "LastModifiedTime": number,
   "LineageGroupArn": "string",
   "MetadataProperties": {
      "CommitId": "string",
      "GeneratedBy": "string",
      "ProjectId": "string",
      "Repository": "string"
   },
   "Properties": {
      "string": "string"
   }
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ActionArn (p. 279)**

The Amazon Resource Name (ARN) of the action.

Type: String

Length Constraints: Maximum length of 256.

**Pattern:** arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:action/.*

**ActionName (p. 279)**

The name of the action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

**Pattern:** (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment| experiment-trial|experiment-trial-component|artifact|action|context)\/\?)?([a-zA-Z0-9](-*[a-zA-Z0-9])\{0,119})

**ActionType (p. 279)**

The type of the action.

Type: String

Length Constraints: Maximum length of 256.

**CreatedBy (p. 279)**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

**CreationTime (p. 279)**

When the action was created.

Type: Timestamp

**Description (p. 279)**

The description of the action.

Type: String

Length Constraints: Maximum length of 3072.
Pattern: . *

**LastModifiedBy (p. 279)**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

**LastModifiedTime (p. 279)**

When the action was last modified.

Type: Timestamp

**LineageGroupArn (p. 279)**

The Amazon Resource Name (ARN) of the lineage group.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:lineage-group/.*

**MetadataProperties (p. 279)**

Metadata properties of the tracking entity, trial, or trial component.

Type: MetadataProperties (p. 1143) object

**Properties (p. 279)**

A list of the action's properties.

Type: String to string map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: . *

Value Length Constraints: Maximum length of 256.

Value Pattern: . *

**Source (p. 279)**

The source of the action.

Type: ActionSource (p. 860) object

**Status (p. 279)**

The status of the action.

Type: String

Valid Values: Unknown | InProgress | Completed | Failed | Stopping | Stopped

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeAlgorithm
Service: Amazon SageMaker Service

Returns a description of the specified algorithm that is in your account.

Request Syntax

```json
{
    "AlgorithmName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AlgorithmName (p. 283)**

The name of the algorithm to describe.

Type: String


Pattern: `arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:[a-z-]*\/?([a-zA-Z0-9][a-zA-Z0-9-]*){0,62}(?:[a-zA-Z0-9](?<!-)[a-zA-Z0-9-]*){0,62}(?:[a-zA-Z0-9\-]{0,62}?<[-]*?)*$

Required: Yes

Response Syntax

```json
{
    "AlgorithmArn": "string",
    "AlgorithmDescription": "string",
    "AlgorithmName": "string",
    "AlgorithmStatus": "string",
    "AlgorithmStatusDetails": {
        "ImageScanStatuses": [
            {
                "FailureReason": "string",
                "Name": "string",
                "Status": "string"
            }
        ],
        "ValidationStatuses": [
            {
                "FailureReason": "string",
                "Name": "string",
                "Status": "string"
            }
        ],
        "CertifyForMarketplace": boolean,
        "CreationTime": number,
        "InferenceSpecification": {
            "Containers": [
```
```
{
    "ContainerHostname": "string",
    "Environment": {
        "string": "string"
    },
    "Framework": "string",
    "FrameworkVersion": "string",
    "Image": "string",
    "ImageDigest": "string",
    "ModelDataUrl": "string",
    "ModelInput": {
        "DataInputConfig": "string"
    },
    "NearestModelName": "string",
    "ProductId": "string"
}
```

```
"SupportedContentTypes": [ "string" ],
"SupportedRealtimeInferenceInstanceTypes": [ "string" ],
"SupportedResponseMIMETypes": [ "string" ],
"SupportedTransformInstanceTypes": [ "string" ]
```

```
"ProductId": "string",
"TrainingSpecification": {
    "MetricDefinitions": [
        {
            "Name": "string",
            "Regex": "string"
        }
    ],
    "SupportedHyperParameters": [
        {
            "DefaultValue": "string",
            "Description": "string",
            "IsRequired": boolean,
            "IsTunable": boolean,
            "Name": "string",
            "Range": {
                "CategoricalParameterRangeSpecification": {
                    "Values": [ "string" ]
                },
                "ContinuousParameterRangeSpecification": {
                    "MaxValue": "string",
                    "MinValue": "string"
                },
                "IntegerParameterRangeSpecification": {
                    "MaxValue": "string",
                    "MinValue": "string"
                }
            },
            "Type": "string"
        }
    ],
    "SupportedTrainingInstanceTypes": [ "string" ],
    "SupportedTuningJobObjectiveMetrics": [
        {
            "MetricName": "string",
            "Type": "string"
        }
    ],
    "SupportsDistributedTraining": boolean,
    "TrainingChannels": [
        {
            "Description": "string",
            "IsRequired": boolean,
            "Name": "string"
        }
    ]
```
"SupportedCompressionTypes": [ "string" ],
"SupportedContentTypes": [ "string" ],
"SupportedInputModes": [ "string" ]
},
"TrainingImage": "string",
"TrainingImageDigest": "string"
},
"ValidationSpecification": {
  "ValidationProfiles": [ {
    "ProfileName": "string",
    "TrainingJobDefinition": {
      "HyperParameters": { 
        "string": "string"
      },
      "InputDataConfig": [ {
        "ChannelName": "string",
        "CompressionType": "string",
        "ContentType": "string",
        "DataSource": { 
          "FileSystemDataSource": { 
            "DirectoryPath": "string",
            "FileSystemAccessMode": "string",
            "FileSystemId": "string",
            "FileSystemType": "string"
          },
          "S3DataSource": { 
            "AttributeNames": [ "string" ],
            "S3DataDistributionType": "string",
            "S3DataType": "string",
            "S3Uri": "string"
          }
        },
        "InputMode": "string",
        "RecordWrapperType": "string",
        "ShuffleConfig": { 
          "Seed": number
        }
      }
    },
    "OutputDataConfig": { 
      "KmsKeyId": "string",
      "S3OutputPath": "string"
    },
    "ResourceConfig": { 
      "InstanceCount": number,
      "InstanceType": "string",
      "VolumeKmsKeyId": "string",
      "VolumeSizeInGB": number
    },
    "StoppingCondition": { 
      "MaxRuntimeInSeconds": number,
      "MaxWaitTimeInSeconds": number
    },
    "TrainingInputMode": "string"
  }
},
"TransformJobDefinition": { 
  "BatchStrategy": "string",
  "Environment": { 
    "string": "string"
  },
  "MaxConcurrentTransforms": number,
  "MaxPayloadInMB": number,
  "TransformInput": { 
    "string": "string" 
  } 
}
"CompressionType": "string",
"ContentType": "string",
"DataSource": {
  "S3DataSource": {
    "S3DataType": "string",
    "S3Uri": "string"
  }
},
"SplitType": "string"
},
"TransformOutput": {
  "Accept": "string",
  "AssembleWith": "string",
  "KmsKeyId": "string",
  "S3OutputPath": "string"
},
"TransformResources": {
  "InstanceCount": number,
  "InstanceType": "string",
  "VolumeKmsKeyId": "string"
}
],
"ValidationRole": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AlgorithmArn (p. 283)**

The Amazon Resource Name (ARN) of the algorithm.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:algorithm/.*

**AlgorithmDescription (p. 283)**

A brief summary about the algorithm.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*

**AlgorithmName (p. 283)**

The name of the algorithm being described.

Type: String


Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9]{0,62}$
**AlgorithmStatus (p. 283)**

The current status of the algorithm.

Type: String

Valid Values: Pending | InProgress | Completed | Failed | Deleting

**AlgorithmStatusDetails (p. 283)**

Details about the current status of the algorithm.

Type: AlgorithmStatusDetails (p. 870) object

**CertifyForMarketplace (p. 283)**

Whether the algorithm is certified to be listed in AWS Marketplace.

Type: Boolean

**CreationTime (p. 283)**

A timestamp specifying when the algorithm was created.

Type: Timestamp

**InferenceSpecification (p. 283)**

Details about inference jobs that the algorithm runs.

Type: InferenceSpecification (p. 1105) object

**ProductId (p. 283)**

The product identifier of the algorithm.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]+$

**TrainingSpecification (p. 283)**

Details about training jobs run by this algorithm.

Type: TrainingSpecification (p. 1400) object

**ValidationSpecification (p. 283)**

Details about configurations for one or more training jobs that SageMaker runs to test the algorithm.

Type: AlgorithmValidationSpecification (p. 875) object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeApp
Service: Amazon SageMaker Service

Describes the app.

Request Syntax

```json
{
    "AppName": "string",
    "AppType": "string",
    "DomainId": "string",
    "UserProfileName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common
Parameters (p. 1470).

The request accepts the following data in JSON format.

**AppName (p. 289)**

The name of the app.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-[a-zA-Z0-9]*){0,62}`

Required: Yes

**AppType (p. 289)**

The type of app.

Type: String

Valid Values: JupyterServer | KernelGateway

Required: Yes

**DomainId (p. 289)**

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**UserProfileName (p. 289)**

The user profile name.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-[a-zA-Z0-9]*){0,62}`
Required: Yes

Response Syntax

```
{
  "AppArn": "string",
  "AppName": "string",
  "AppType": "string",
  "CreationTime": number,
  "DomainId": "string",
  "FailureReason": "string",
  "LastHealthCheckTimestamp": number,
  "LastUserActivityTimestamp": number,
  "ResourceSpec": {
    "InstanceType": "string",
    "LifecycleConfigArn": "string",
    "SageMakerImageArn": "string",
    "SageMakerImageVersionArn": "string"
  },
  "Status": "string",
  "UserProfileName": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AppArn (p. 290)**

The Amazon Resource Name (ARN) of the app.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:app/.*`

**AppName (p. 290)**

The name of the app.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9-]*[a-zA-Z0-9]*\{0,62}`

**AppType (p. 290)**

The type of app.

Type: String

Valid Values: JupyterServer | KernelGateway

**CreationTime (p. 290)**

The creation time.

Type: Timestamp
DomainId (p. 290)

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

FailureReason (p. 290)

The failure reason.

Type: String

Length Constraints: Maximum length of 1024.

LastHealthCheckTimestamp (p. 290)

The timestamp of the last health check.

Type: Timestamp

LastUserActivityTimestamp (p. 290)

The timestamp of the last user's activity. LastUserActivityTimestamp is also updated when SageMaker performs health checks without user activity. As a result, this value is set to the same value as LastHealthCheckTimestamp.

Type: Timestamp

ResourceSpec (p. 290)

The instance type and the Amazon Resource Name (ARN) of the SageMaker image created on the instance.

Type: ResourceSpec (p. 1341) object

Status (p. 290)

The status.

Type: String

Valid Values: Deleted | Deleting | Failed | InService | Pending

UserProfileName (p. 290)

The user profile name.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-[a-zA-Z0-9]+)*$
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeAppImageConfig

Service: Amazon SageMaker Service

Describes an AppImageConfig.

Request Syntax

```
{
  "AppImageConfigName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AppImageConfigName (p. 293)**

The name of the AppImageConfig to describe.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}`

Required: Yes

Response Syntax

```
{
  "AppImageConfigArn": "string",
  "AppImageConfigName": "string",
  "CreationTime": number,
  "KernelGatewayImageConfig": {
    "FileSystemConfig": {
      "DefaultGid": number,
      "DefaultUid": number,
      "MountPath": "string"
    },
    "KernelSpecs": [
      {
        "DisplayName": "string",
        "Name": "string"
      }
    ],
    "LastModifiedTime": number
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
AppImageConfigArn (p. 293)

The Amazon Resource Name (ARN) of the AppImageConfig.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:app-image-config/.*

AppImageConfigName (p. 293)

The name of the AppImageConfig.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-](\*[^a-zA-Z0-9\-])\{0,62}\}

CreationTime (p. 293)

When the AppImageConfig was created.

Type: Timestamp

KernelGatewayImageConfig (p. 293)

The configuration of a KernelGateway app.

Type: KernelGatewayImageConfig (p. 1117) object

LastModifiedTime (p. 293)

When the AppImageConfig was last modified.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DescribeArtifact
Service: Amazon SageMaker Service

Describes an artifact.

Request Syntax

```
{
    "ArtifactArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ArtifactArn (p. 296)**

The Amazon Resource Name (ARN) of the artifact to describe.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:artifact/.*`

Required: Yes

Response Syntax

```
{
    "ArtifactArn": "string",
    "ArtifactName": "string",
    "ArtifactType": "string",
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "LastModifiedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "LastModifiedTime": number,
    "LineageGroupArn": "string",
    "MetadataProperties": {
        "CommitId": "string",
        "GeneratedBy": "string",
        "ProjectId": "string",
        "Repository": "string"
    },
    "Properties": {
        "string": "string"
    }
}
```
"Source": {
    "SourceTypes": [
      {
        "SourceIdType": "string",
        "Value": "string"
      },
      "SourceUri": "string"
    ]
  }

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ArtifactArn (p. 296)

The Amazon Resource Name (ARN) of the artifact.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:artifact/.*

ArtifactName (p. 296)

The name of the artifact.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial|experiment-trial-component|artifact|action|context)\//?([a-zA-Z0-9\-\*]{0,119})

ArtifactType (p. 296)

The type of the artifact.

Type: String

Length Constraints: Maximum length of 256.

CreatedBy (p. 296)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

CreationTime (p. 296)

When the artifact was created.

Type: Timestamp

LastModifiedBy (p. 296)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
Type: UserContext (p. 1449) object

**LastModifiedTime (p. 296)**

When the artifact was last modified.

Type: Timestamp

**LineageGroupArn (p. 296)**

The Amazon Resource Name (ARN) of the lineage group.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:lineage-group/.*`

**MetadataProperties (p. 296)**

Metadata properties of the tracking entity, trial, or trial component.

Type: MetadataProperties (p. 1143) object

**Properties (p. 296)**

A list of the artifact's properties.

Type: String to string map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: `. *`

Value Length Constraints: Maximum length of 256.

Value Pattern: `. *`

**Source (p. 296)**

The source of the artifact.

Type: ArtifactSource (p. 892) object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFoundException**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
**DescribeAutoMLJob**

Service: Amazon SageMaker Service

Returns information about an Amazon SageMaker AutoML job.

**Request Syntax**

```json
{
   "AutoMLJobName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see [Common Parameters (p. 1470)](#).

The request accepts the following data in JSON format.

**AutoMLJobName (p. 300)**

Requests information about an AutoML job using its unique name.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}$

Required: Yes

**Response Syntax**

```json
{
   "AutoMLJobArn": "string",
   "AutoMLJobArtifacts": {
      "CandidateDefinitionNotebookLocation": "string",
      "DataExplorationNotebookLocation": "string"
   },
   "AutoMLJobConfig": {
      "CandidateGenerationConfig": {
         "FeatureSpecificationsS3Uri": "string"
      },
      "CompletionCriteria": {
         "MaxAutoMLJobRuntimeInSeconds": number,
         "MaxCandidates": number,
         "MaxRuntimePerTrainingJobInSeconds": number
      },
      "DataSplitConfig": {
         "ValidationFraction": number
      },
      "SecurityConfig": {
         "EnableInterContainerTrafficEncryption": boolean,
         "VolumeKmsKeyId": "string",
         "VpcConfig": {
            "SecurityGroupIds": [ "string" ],
            "Subnets": [ "string" ]
         }
      }
   }
}
```
"AutoMLJobName": "string",
"AutoMLJobObjective": {
    "MetricName": "string"
},
"AutoMLJobSecondaryStatus": "string",
"AutoMLJobStatus": "string",
"BestCandidate": {
    "CandidateName": "string",
    "CandidateProperties": {
        "CandidateArtifactLocations": {
            "Explainability": "string",
            "ModelInsights": "string"
        },
        "CandidateMetrics": {
            "MetricName": "string",
            "Set": "string",
            "StandardMetricName": "string",
            "Value": number
        }
    },
    "CandidateStatus": "string",
    "CandidateSteps": [
        {
            "CandidateStepArn": "string",
            "CandidateStepName": "string",
            "CandidateStepType": "string"
        }
    ],
    "CreationTime": number,
    "EndTime": number,
    "FailureReason": "string",
    "FinalAutoMLJobObjectiveMetric": {
        "MetricName": "string",
        "Type": "string",
        "Value": number
    },
    "InferenceContainers": [
        {
            "Environment": {
                "string": "string"
            },
            "Image": "string",
            "ModelDataUrl": "string"
        }
    ],
    "LastModifiedTime": number,
    "ObjectiveStatus": "string"
},
"CreationTime": number,
"EndTime": number,
"FailureReason": "string",
"GenerateCandidateDefinitionsOnly": boolean,
"InputDataConfig": [
    {
        "ChannelType": "string",
        "CompressionType": "string",
        "ContentType": "string",
        "DataSource": {
            "S3DataSource": {
                "S3DataType": "string",
                "S3Uri": "string"
            }
        },
        "TargetAttributeName": "string"
    }
]
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AutoMLJobArn (p. 300)**

Returns the ARN of the AutoML job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*

**AutoMLJobArtifacts (p. 300)**

Returns information on the job's artifacts found in AutoMLJobArtifacts.

Type: `AutoMLJobArtifacts (p. 914)` object

**AutoMLJobConfig (p. 300)**

Returns the configuration for the AutoML job.

Type: `AutoMLJobConfig (p. 916)` object

**AutoMLJobName (p. 300)**

Returns the name of the AutoML job.
Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}

**AutoMLJobObjective (p. 300)**

Returns the job's objective.

Type: `AutoMLJobObjective (p. 917)` object

**AutoMLJobSecondaryStatus (p. 300)**

Returns the secondary status of the AutoML job.

Type: String


**AutoMLJobStatus (p. 300)**

Returns the status of the AutoML job.

Type: String

Valid Values: Completed | InProgress | Failed | Stopped | Stopping

**BestCandidate (p. 300)**

Returns the job's best AutoMLCandidate.

Type: `AutoMLCandidate (p. 904)` object

**CreationTime (p. 300)**

Returns the creation time of the AutoML job.

Type: Timestamp

**EndTime (p. 300)**

Returns the end time of the AutoML job.

Type: Timestamp

**FailureReason (p. 300)**

Returns the failure reason for an AutoML job, when applicable.

Type: String

Length Constraints: Maximum length of 1024.

**GenerateCandidateDefinitionsOnly (p. 300)**

Indicates whether the output for an AutoML job generates candidate definitions only.

Type: Boolean

**InputDataConfig (p. 300)**

Returns the input data configuration for the AutoML job.

Type: Array of `AutoMLChannel (p. 908)` objects
Array Members: Minimum number of 1 item. Maximum number of 2 items.

**LastModifiedTime (p. 300)**

Returns the job's last modified time.

Type: Timestamp

**ModelDeployConfig (p. 300)**

Indicates whether the model was deployed automatically to an endpoint and the name of that endpoint if deployed automatically.

Type: `ModelDeployConfig (p. 1156)` object

**ModelDeployResult (p. 300)**

Provides information about endpoint for the model deployment.

Type: `ModelDeployResult (p. 1157)` object

**OutputDataConfig (p. 300)**

Returns the job's output data config.

Type: `AutoMLOutputDataConfig (p. 921)` object

**PartialFailureReasons (p. 300)**

Returns a list of reasons for partial failures within an AutoML job.

Type: Array of `AutoMLPartialFailureReason (p. 922)` objects

Array Members: Minimum number of 1 item. Maximum number of 5 items.

**ProblemType (p. 300)**

Returns the job's problem type.

Type: String

Valid Values: `BinaryClassification` | `MulticlassClassification` | `Regression`

**ResolvedAttributes (p. 300)**

This contains `ProblemType`, `AutoMLJobObjective`, and `CompletionCriteria`. If you do not provide these values, they are auto-inferred. If you do provide them, the values used are the ones you provide.

Type: `ResolvedAttributes (p. 1337)` object

**RoleArn (p. 300)**

The Amazon Resource Name (ARN) of the AWS Identity and Access Management (IAM) role that has read permission to the input data location and write permission to the output data location in Amazon S3.

Type: String


Pattern: `^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9+=,.@\-_\/%]+$`

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](mailto:).
Resource Not Found

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeCodeRepository
Service: Amazon SageMaker Service

Gets details about the specified Git repository.

Request Syntax

```
{
   "CodeRepositoryName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CodeRepositoryName (p. 306)**

The name of the Git repository to describe.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Syntax

```
{
   "CodeRepositoryArn": "string",
   "CodeRepositoryName": "string",
   "CreationTime": number,
   "GitConfig": {
      "Branch": "string",
      "RepositoryUrl": "string",
      "SecretArn": "string"
   },
   "LastModifiedTime": number
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CodeRepositoryArn (p. 306)**

The Amazon Resource Name (ARN) of the Git repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.
DescribeCodeRepository

**Pattern:** arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:code-repository/.*

**CodeRepositoryName (p. 306)**

The name of the Git repository.

Type: String


Pattern: ^[a-zA-Z0-9]{0,62}$

**CreationTime (p. 306)**

The date and time that the repository was created.

Type: Timestamp

**GitConfig (p. 306)**

Configuration details about the repository, including the URL where the repository is located, the default branch, and the Amazon Resource Name (ARN) of the AWS Secrets Manager secret that contains the credentials used to access the repository.

Type: GitConfig (p. 1054) object

**LastModifiedTime (p. 306)**

The date and time that the repository was last changed.

Type: Timestamp

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeCompilationJob

Service: Amazon SageMaker Service

Returns information about a model compilation job.

To create a model compilation job, use CreateCompilationJob (p. 46). To get information about multiple model compilation jobs, use ListCompilationJobs (p. 525).

Request Syntax

```
{
    "CompilationJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CompilationJobName (p. 308)**

The name of the model compilation job that you want information about.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Syntax

```
{
    "CompilationEndTime": number,
    "CompilationJobArn": "string",
    "CompilationJobName": "string",
    "CompilationJobStatus": "string",
    "CompilationStartTime": number,
    "CreationTime": number,
    "FailureReason": "string",
    "InferenceImage": "string",
    "InputConfig": {
        "DataInputConfig": "string",
        "Framework": "string",
        "FrameworkVersion": "string",
        "S3Uri": "string"
    },
    "LastModifiedTime": number,
    "ModelArtifacts": {
        "S3ModelArtifacts": "string"
    },
    "ModelDigests": {
        "ArtifactDigest": "string"
    },
    "ModelPackageVersionArn": "string",

```
"OutputConfig": {  
  "CompilerOptions": "string",  
  "KmsKeyId": "string",  
  "S3OutputLocation": "string",  
  "TargetDevice": "string",  
  "TargetPlatform": {  
    "Accelerator": "string",  
    "Arch": "string",  
    "Os": "string"  
  },  
  "RoleArn": "string",  
  "StoppingCondition": {  
    "MaxRuntimeInSeconds": number,  
    "MaxWaitTimeInSeconds": number  
  },  
  "VpcConfig": {  
    "SecurityGroupIds": [ "string" ],  
    "Subnets": [ "string" ]  
  }  
},  
"CompilationEndTime": number,  
"CompilationJobArn": string,  
"CompilationJobName": string,  
"CompilationJobStatus": string

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CompilationEndTime (p. 308)

The time when the model compilation job on a compilation job instance ended. For a successful or stopped job, this is when the job's model artifacts have finished uploading. For a failed job, this is when Amazon SageMaker detected that the job failed.

Type: Timestamp

CompilationJobArn (p. 308)

The Amazon Resource Name (ARN) of the model compilation job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:compilation-job/.*

CompilationJobName (p. 308)

The name of the model compilation job.

Type: String


Pattern: ^[a-zA-iz0-9\-]*[a-zA-Z0-9]{0,62}$

CompilationJobStatus (p. 308)

The status of the model compilation job.

Type: String

Valid Values: INPROGRESS | COMPLETED | FAILED | STARTING | STOPPING | STOPPED
**CompilationStartTime (p. 308)**

The time when the model compilation job started the CompilationJob instances.

You are billed for the time between this timestamp and the timestamp in the DescribeCompilationJob:CompilationEndTime (p. 309) field. In Amazon CloudWatch Logs, the start time might be later than this time. That's because it takes time to download the compilation job, which depends on the size of the compilation job container.

Type: Timestamp

**CreationTime (p. 308)**

The time that the model compilation job was created.

Type: Timestamp

**FailureReason (p. 308)**

If a model compilation job failed, the reason it failed.

Type: String

Length Constraints: Maximum length of 1024.

**InferenceImage (p. 308)**

The inference image to use when compiling a model. Specify an image only if the target device is a cloud instance.

Type: String

Length Constraints: Maximum length of 256.

**InputConfig (p. 308)**

Information about the location in Amazon S3 of the input model artifacts, the name and shape of the expected data inputs, and the framework in which the model was trained.

Type: InputConfig (p. 1107) object

**LastModifiedTime (p. 308)**

The time that the status of the model compilation job was last modified.

Type: Timestamp

**ModelArtifacts (p. 308)**

Information about the location in Amazon S3 that has been configured for storing the model artifacts used in the compilation job.

Type: ModelArtifacts (p. 1149) object

**ModelDigests (p. 308)**

Provides a BLAKE2 hash value that identifies the compiled model artifacts in Amazon S3.

Type: ModelDigests (p. 1158) object

**ModelPackageVersionArn (p. 308)**

The Amazon Resource Name (ARN) of the versioned model package that was provided to SageMaker Neo when you initiated a compilation job.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package/.*

**OutputConfig (p. 308)**

Information about the output location for the compiled model and the target device that the model runs on.

Type: OutputConfig (p. 1240) object

**RoleArn (p. 308)**

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker assumes to perform the model compilation job.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_]+$

**StoppingCondition (p. 308)**

Specifies a limit to how long a model compilation job can run. When the job reaches the time limit, Amazon SageMaker ends the compilation job. Use this API to cap model training costs.

Type: StoppingCondition (p. 1370) object

**VpcConfig (p. 308)**

A VpcConfig (p. 1456) object that specifies the VPC that you want your compilation job to connect to. Control access to your models by configuring the VPC. For more information, see Protect Compilation Jobs by Using an Amazon Virtual Private Cloud.

Type: NeoVpcConfig (p. 1221) object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
• AWS SDK for Ruby V3
DescribeContext
Service: Amazon SageMaker Service

Describes a context.

Request Syntax

```
{
    "ContextName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ContextName (p. 313)**

The name of the context to describe.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: (arn:aws[a-z\-]{1,}\*:sagemaker:[a-z0-9\-]{1,}:0-9\{12\}:(experiment|experiment-trial|experiment-trial-component|artifact|action|context)\(/\)?(a-zA-Z0-9\{-*[a-zA-Z0-9]\}\{0,119\})

Required: Yes

Response Syntax

```
{
    "ContextArn": "string",
    "ContextName": "string",
    "ContextType": "string",
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "Description": "string",
    "LastModifiedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "LastModifiedTime": number,
    "LineageGroupArn": "string",
    "Properties": {
        "string": "string"
    },
    "Source": {
        "SourceId": "string",
        "SourceType": "string"
    }
}
```
"SourceUri": "string"
}
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ContextArn (p. 313)**

The Amazon Resource Name (ARN) of the context.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:context/.*

**ContextName (p. 313)**

The name of the context.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9\-]{0,119}

**ContextType (p. 313)**

The type of the context.

Type: String

Length Constraints: Maximum length of 256.

**CreatedBy (p. 313)**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

**CreationTime (p. 313)**

When the context was created.

Type: Timestamp

**Description (p. 313)**

The description of the context.

Type: String

Length Constraints: Maximum length of 3072.

Pattern: .*

**LastModifiedBy (p. 313)**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
DescribeContext

Type: UserContext (p. 1449) object

_LastModifiedTime (p. 313)_

When the context was last modified.

Type: Timestamp

_LineageGroupArn (p. 313)_

The Amazon Resource Name (ARN) of the lineage group.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:lineage-group/.*

_Properties (p. 313)_

A list of the context's properties.

Type: String to string map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: . *

Value Length Constraints: Maximum length of 256.

Value Pattern: . *

_Source (p. 313)_

The source of the context.

Type: ContextSource (p. 959) object

_Errors_

For information about the errors that are common to all actions, see Common Errors (p. 1472).

_ResourceNotFoundError_

Resource being access is not found.

HTTP Status Code: 400

_See Also_

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DescribeDataQualityJobDefinition
Service: Amazon SageMaker Service

Gets the details of a data quality monitoring job definition.

Request Syntax

```
{
  "JobDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

*JobDefinitionName* (p. 317)

The name of the data quality monitoring job definition to describe.

Type: String
Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*$`
Required: Yes

Response Syntax

```
{
  "CreationTime": number,
  "DataQualityAppSpecification": {
    "ContainerArguments": [ "string" ],
    "ContainerEntrypoint": [ "string" ],
    "Environment": {
      "string" : "string"
    },
    "ImageUri": "string",
    "PostAnalyticsProcessorSourceUri": "string",
    "RecordPreprocessorSourceUri": "string"
  },
  "DataQualityBaselineConfig": {
    "BaseliningJobName": "string",
    "ConstraintsResource": {
      "S3Uri": "string"
    },
    "StatisticsResource": {
      "S3Uri": "string"
    }
  },
  "DataQualityJobInput": {
    "EndpointInput": {
      "EndpointName": "string",
      "EndTimeOffset": "string",
      "FeaturesAttribute": "string",
      "InferenceAttribute": "string",
```

317
"LocalPath": "string",
"ProbabilityAttribute": "string",
"ProbabilityThresholdAttribute": number,
"S3DataDistributionType": "string",
"S3InputMode": "string",
"StartTimeOffset": "string"
},

"DataQualityJobOutputConfig": {
  "KmsKeyId": "string",
  "MonitoringOutputs": [
    {
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
},

"JobDefinitionArn": "string",
"JobDefinitionName": "string",
"JobResources": {
  "ClusterConfig": {
    "InstanceCount": number,
    "InstanceType": "string",
    "VolumeKmsKeyId": "string",
    "VolumeSizeInGB": number
  }
},

"NetworkConfig": {
  "EnableInterContainerTrafficEncryption": boolean,
  "EnableNetworkIsolation": boolean,
  "VpcConfig": {
    "SecurityGroupIds": [ "string" ],
    "Subnets": [ "string" ]
  }
},

"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeInSeconds": number
}
}

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 317)**

The time that the data quality monitoring job definition was created.

*Type: Timestamp*

**DataQualityAppSpecification (p. 317)**

Information about the container that runs the data quality monitoring job.

*Type: DataQualityAppSpecification (p. 973) object*

**DataQualityBaselineConfig (p. 317)**

The constraints and baselines for the data quality monitoring job definition.
Type: `DataQualityBaselineConfig (p. 975)` object

**DataQualityJobInput (p. 317)**

The list of inputs for the data quality monitoring job. Currently endpoints are supported.

Type: `DataQualityJobInput (p. 976)` object

**DataQualityJobOutputConfig (p. 317)**

The output configuration for monitoring jobs.

Type: `MonitoringOutputConfig (p. 1209)` object

**JobDefinitionArn (p. 317)**

The Amazon Resource Name (ARN) of the data quality monitoring job definition.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

**JobDefinitionName (p. 317)**

The name of the data quality monitoring job definition.

Type: String


Pattern: ^[a-zA-Z0-9](-[a-zA-Z0-9])*$

**JobResources (p. 317)**

Identifies the resources to deploy for a monitoring job.

Type: `MonitoringResources (p. 1210)` object

**NetworkConfig (p. 317)**

The networking configuration for the data quality monitoring job.

Type: `MonitoringNetworkConfig (p. 1207)` object

**RoleArn (p. 317)**

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/?[a-zA-Z0-9+=,.@-_\//]+$

**StoppingCondition (p. 317)**

A time limit for how long the monitoring job is allowed to run before stopping.

Type: `MonitoringStoppingCondition (p. 1219)` object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeDevice

Service: Amazon SageMaker Service

Describes the device.

Request Syntax

```json
{
   "DeviceFleetName": "string",
   "DeviceName": "string",
   "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeviceFleetName (p. 321)**

The name of the fleet the devices belong to.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

**DeviceName (p. 321)**

The unique ID of the device.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

**NextToken (p. 321)**

Next token of device description.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: `.*`

Required: No

Response Syntax

```json
{
}
```
DescribeDevice

```json
"AgentVersion": "string",
"Description": "string",
"DeviceArn": "string",
"DeviceFleetName": "string",
"DeviceName": "string",
"IotThingName": "string",
"LatestHeartbeat": number,
"MaxModels": number,
"Models": [
  {
    "LatestInference": number,
    "LatestSampleTime": number,
    "ModelName": "string",
    "ModelVersion": "string"
  }
],
"NextToken": "string",
"RegistrationTime": number
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AgentVersion (p. 321)**

Edge Manager agent version.

Type: String


Pattern: \[a-zA-Z0-9\ \_.\]+ 

**Description (p. 321)**

A description of the device.

Type: String


Pattern: \[\S\s\]+ 

**DeviceArn (p. 321)**

The Amazon Resource Name (ARN) of the device.

Type: String


Pattern: ^arn:aws[a-z\-]*:[a-z\-]*:[a-z\-]*:\d{12}:[a-z\-]*[/][a-zA-Z0-9-]+=,\@[a-zA-Z0-9]+\d$/

**DeviceFleetName (p. 321)**

The name of the fleet the device belongs to.

Type: String

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

**DeviceName (p. 321)**

The unique identifier of the device.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

**IotThingName (p. 321)**

The AWS Internet of Things (IoT) object thing name associated with the device.

Type: String

Length Constraints: Maximum length of 128.

Pattern: [a-zA-Z0-9:_-]+

**LatestHeartbeat (p. 321)**

The last heartbeat received from the device.

Type: Timestamp

**MaxModels (p. 321)**

The maximum number of models.

Type: Integer

**Models (p. 321)**

Models on the device.

Type: Array of EdgeModel (p. 1004) objects

**NextToken (p. 321)**

The response from the last list when returning a list large enough to need tokening.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**RegistrationTime (p. 321)**

The timestamp of the last registration or de-reregistration.

Type: Timestamp

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeDeviceFleet
Service: Amazon SageMaker Service

A description of the fleet the device belongs to.

Request Syntax

```json
{
   "DeviceFleetName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeviceFleetName (p. 325)**

The name of the fleet.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Syntax

```json
{
   "CreationTime": number,
   "Description": "string",
   "DeviceFleetArn": "string",
   "DeviceFleetName": "string",
   "IotRoleAlias": "string",
   "LastModifiedTime": number,
   "OutputConfig": {
      "KmsKeyId": "string",
      "PresetDeploymentConfig": "string",
      "PresetDeploymentType": "string",
      "S3OutputLocation": "string"
   },
   "RoleArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 325)**

Timestamp of when the device fleet was created.
**Type:** Timestamp

**Description (p. 325)**

A description of the fleet.

**Type:** String

**Length Constraints:** Minimum length of 1. Maximum length of 800.

**Pattern:** \[\S\s\]+

**DeviceFleetArn (p. 325)**

The Amazon Resource Name (ARN) of the fleet.

**Type:** String

**Pattern:** ^arn:aws[a-z\-]*:iam::d\{12\}:device-fleet/?[a-zA-Z_0-9+=,.@\-_/]+$\#

**DeviceFleetName (p. 325)**

The name of the fleet.

**Type:** String

**Length Constraints:** Minimum length of 1. Maximum length of 63.

**Pattern:** ^[a-zA-Z\-20-9\-\_]\{0,62\}$

**IotRoleAlias (p. 325)**

The Amazon Resource Name (ARN) alias created in AWS Internet of Things (IoT).

**Type:** String

**Pattern:** ^arn:aws[a-z\-]*:iam::d\{12\}:rolealias/?[a-zA-Z_0-9+=,.@\-_/]+$\#

**LastModifiedTime (p. 325)**

Timestamp of when the device fleet was last updated.

**Type:** Timestamp

**OutputConfig (p. 325)**

The output configuration for storing sampled data.

**Type:** EdgeOutputConfig (p. 1008) object

**RoleArn (p. 325)**

The Amazon Resource Name (ARN) that has access to AWS Internet of Things (IoT).

**Type:** String

**Length Constraints:** Minimum length of 20. Maximum length of 2048.

**Pattern:** ^arn:aws[a-z\-]*:iam::d\{12\}:role/?[a-zA-Z_0-9+=,.@\-_/]+$\#

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeDomain

The description of the domain.

Request Syntax

```
{
  "DomainId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainId (p. 328)**

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

Response Syntax

```
{
  "AppNetworkAccessType": "string",
  "AppSecurityGroupManagement": "string",
  "AuthMode": "string",
  "CreationTime": number,
  "DefaultUserSettings": {
    "ExecutionRole": "string",
    "JupyterServerAppSettings": {
      "DefaultResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      },
      "LifecycleConfigArns": [ "string" ]
    },
    "KernelGatewayAppSettings": {
      "CustomImages": [ {
        "AppImageConfigName": "string",
        "ImageName": "string",
        "ImageVersionNumber": number
      } ],
      "DefaultResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      }
    }
  }
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AppNetworkAccessType (p. 328)**

Specifies the VPC used for non-EFS traffic. The default value is `PublicInternetOnly`.

- **PublicInternetOnly** - Non-EFS traffic is through a VPC managed by Amazon SageMaker, which allows direct internet access
- **VpcOnly** - All Studio traffic is through the specified VPC and subnets

Type: String

Valid Values: `PublicInternetOnly` | `VpcOnly`

**AppSecurityGroupManagement (p. 328)**

The entity that creates and manages the required security groups for inter-app communication in `VpcOnly` mode. Required when `CreateDomain.AppNetworkAccessType` is `VpcOnly` and `DomainSettings.RStudioServerProDomainSettings.DomainExecutionRoleArn` is provided.

Type: String

Valid Values: `Service` | `Customer`

**AuthMode (p. 328)**

The domain's authentication mode.

Type: String

Valid Values: `SSO` | `IAM`

**CreationTime (p. 328)**

The creation time.

Type: Timestamp

**DefaultUserSettings (p. 328)**

Settings which are applied to UserProfiles in this domain if settings are not explicitly specified in a given UserProfile.

Type: `UserSettings (p. 1452)` object

**DomainArn (p. 328)**

The domain's Amazon Resource Name (ARN).

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:domain/.*`

**DomainId (p. 328)**

The domain ID.
DescribeDomain

Type: String

DomainName (p. 328)

The domain name.

Type: String

DomainSettings (p. 328)

A collection of Domain settings.

Type: DomainSettings (p. 996) object

FailureReason (p. 328)

The failure reason.

Type: String

HomeEfsFileSystemId (p. 328)

The ID of the Amazon Elastic File System (EFS) managed by this Domain.

Type: String

HomeEfsFileSystemKmsKeyId (p. 328)

This parameter has been deprecated.

Use KmsKeyId.

Type: String

KmsKeyId (p. 328)

The AWS KMS customer managed key used to encrypt the EFS volume attached to the domain.

Type: String

LastModifiedTime (p. 328)

The last modified time.

Type: Timestamp

SecurityGroupIdForDomainBoundary (p. 328)

The ID of the security group that authorizes traffic between the RSessionGateway apps and the RStudioServerPro app.
Type: String
Length Constraints: Maximum length of 32.
Pattern: [-0-9a-zA-Z]+  

**SingleSignOnManagedApplicationInstanceId (p. 328)**

The SSO managed application instance ID.

Type: String
Length Constraints: Maximum length of 256.

**Status (p. 328)**

The status.

Type: String

Valid Values: Deleting | Failed | InService | Pending | Updating | Update_Failed | Delete_Failed

**SubnetIds (p. 328)**

The VPC subnets that Studio uses for communication.

Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 16 items.

Length Constraints: Maximum length of 32.

Pattern: [-0-9a-zA-Z]+  

**Url (p. 328)**

The domain's URL.

Type: String

Length Constraints: Maximum length of 1024.

**VpcId (p. 328)**

The ID of the Amazon Virtual Private Cloud (VPC) that Studio uses for communication.

Type: String

Length Constraints: Maximum length of 32.

Pattern: [-0-9a-zA-Z]+  

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeEdgePackagingJob

Service: Amazon SageMaker Service

A description of edge packaging jobs.

Request Syntax

```json
{
    "EdgePackagingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**EdgePackagingJobName (p. 334)**

The name of the edge packaging job.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Syntax

```json
{
    "CompilationJobName": "string",
    "CreationTime": number,
    "EdgePackagingJobArn": "string",
    "EdgePackagingJobName": "string",
    "EdgePackagingJobStatus": "string",
    "EdgePackagingJobStatusMessage": "string",
    "LastModifiedTime": number,
    "ModelArtifact": "string",
    "ModelName": "string",
    "ModelSignature": "string",
    "ModelVersion": "string",
    "OutputConfig": {
        "KmsKeyId": "string",
        "PresetDeploymentConfig": "string",
        "PresetDeploymentType": "string",
        "S3OutputLocation": "string"
    },
    "PresetDeploymentOutput": {
        "Artifact": "string",
        "Status": "string",
        "StatusMessage": "string",
        "Type": "string"
    },
    "ResourceKey": "string",
    "RoleArn": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CompilationJobName (p. 334)

The name of the SageMaker Neo compilation job that is used to locate model artifacts that are being packaged.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\$

CreationTime (p. 334)

The timestamp of when the packaging job was created.

Type: Timestamp

EdgePackagingJobArn (p. 334)

The Amazon Resource Name (ARN) of the edge packaging job.

Type: String


Pattern: ^arn:aws\[a-z\-]*:sagemaker:\[a-z\-]*:\d\{12\}:edge-packaging-job/?[a-zA-Z_0-9+=,.@\-_/]+\$

EdgePackagingJobName (p. 334)

The name of the edge packaging job.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\$

EdgePackagingJobStatus (p. 334)

The current status of the packaging job.

Type: String

Valid Values: STARTING | INPROGRESS | COMPLETED | FAILED | STOPPING | STOPPED

EdgePackagingJobStatusMessage (p. 334)

Returns a message describing the job status and error messages.

Type: String

LastModifiedTime (p. 334)

The timestamp of when the job was last updated.

Type: Timestamp
ModelArtifact (p. 334)

The Amazon Simple Storage (S3) URI where model artifacts are stored.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(http[s]://[^/]+)/(\.*$)

ModelName (p. 334)

The name of the model.

Type: String


Pattern: ^[a-zA-Z0-9]([-a-zA-Z0-9]*[-a-zA-Z0-9])\{0,62}$

ModelSignature (p. 334)

The signature document of files in the model artifact.

Type: String

ModelVersion (p. 334)

The version of the model.

Type: String


Pattern: [a-zA-Z0-9\_\.-]+

OutputConfig (p. 334)

The output configuration for the edge packaging job.

Type: EdgeOutputConfig (p. 1008) object

PresetDeploymentOutput (p. 334)

The output of a SageMaker Edge Manager deployable resource.

Type: EdgePresetDeploymentOutput (p. 1012) object

ResourceKey (p. 334)

The AWS KMS key to use when encrypting the EBS volume the job run on.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

RoleArn (p. 334)

The Amazon Resource Name (ARN) of an IAM role that enables Amazon SageMaker to download and upload the model, and to contact Neo.

Type: String

Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_\/]++$

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeEndpoint

Returns the description of an endpoint.

Request Syntax

```
{
   "EndpointName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**EndpointName (p. 338)**

The name of the endpoint.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

Required: Yes

Response Syntax

```
{
   "AsyncInferenceConfig": {
      "ClientConfig": {
         "MaxConcurrentInvocationsPerInstance": number
      },
   "OutputConfig": {
      "KmsKeyId": "string",
      "NotificationConfig": {
         "ErrorTopic": "string",
         "SuccessTopic": "string"
      },
      "S3OutputPath": "string"
   },
   "CreationTime": number,
   "DataCaptureConfig": {
      "CaptureStatus": "string",
      "CurrentSamplingPercentage": number,
      "DestinationS3Uri": "string",
      "EnableCapture": boolean,
      "KmsKeyId": "string"
   },
   "EndpointArn": "string",
   "EndpointConfigName": "string",
   "EndpointName": "string",
   "EndpointStatus": "string"
}
```
"FailureReason": "string",
"LastDeploymentConfig": {
  "AutoRollbackConfiguration": {
    "Alarms": [
      {
        "AlarmName": "string"
      }
    ],
  },
  "BlueGreenUpdatePolicy": {
    "MaximumExecutionTimeoutInSeconds": number,
    "TerminationWaitInSeconds": number,
    "TrafficRoutingConfiguration": {
      "CanarySize": {
        "Type": "string",
        "Value": number
      },
      "LinearStepSize": {
        "Type": "string",
        "Value": number
      },
      "Type": "string",
      "WaitIntervalInSeconds": number
    }
  },
  "LastModifiedTime": number,
  "PendingDeploymentSummary": {
    "EndpointConfigName": "string",
    "ProductionVariants": [
      {
        "AcceleratorType": "string",
        "CurrentInstanceCount": number,
        "CurrentServerlessConfig": {
          "MaxConcurrency": number,
          "MemorySizeInMB": number
        },
        "CurrentWeight": number,
        "DeployedImages": [
          {
            "ResolutionTime": number,
            "ResolvedImage": "string",
            "SpecifiedImage": "string"
          }
        ],
        "DesiredInstanceCount": number,
        "DesiredServerlessConfig": {
          "MaxConcurrency": number,
          "MemorySizeInMB": number
        },
        "DesiredWeight": number,
        "InstanceType": "string",
        "VariantName": "string",
        "VariantStatus": [
          {
            "StartTime": number,
            "Status": "string",
            "StatusMessage": "string"
          }
        ],
      }
    ],
    "StartTime": number
  }
},
"ProductionVariants": [
"CurrentInstanceCount": number,
"CurrentServerlessConfig": {
  "MaxConcurreny": number,
  "MemorySizeInMB": number
},
"CurrentWeight": number,
"DeployedImages": [
  {
    "ResolutionTime": number,
    "ResolvedImage": "string",
    "SpecifiedImage": "string"
  }
],
"DesiredInstanceCount": number,
"DesiredServerlessConfig": {
  "MaxConcurreny": number,
  "MemorySizeInMB": number
},
"DesiredWeight": number,
"VariantName": "string",
"VariantStatus": [
  {
    "StartTime": number,
    "Status": "string",
    "StatusMessage": "string"
  }
]
]
]

### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AsyncInferenceConfig (p. 338)**

Returns the description of an endpoint configuration created using the CreateEndpointConfig API.

Type: AsyncInferenceConfig (p. 899) object

**CreationTime (p. 338)**

A timestamp that shows when the endpoint was created.

Type: Timestamp

**DataCaptureConfig (p. 338)**

The currently active data capture configuration used by your Endpoint.

Type: DataCaptureConfigSummary (p. 968) object

**EndpointArn (p. 338)**

The Amazon Resource Name (ARN) of the endpoint.

Type: String


Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:endpoint/.*`
**EndpointConfigName (p. 338)**

The name of the endpoint configuration associated with this endpoint.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

**EndpointName (p. 338)**

Name of the endpoint.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

**EndpointStatus (p. 338)**

The status of the endpoint.
- **OutOfService**: Endpoint is not available to take incoming requests.
- **Creating**: CreateEndpoint (p. 69) is executing.
- **Updating**: UpdateEndpoint (p. 767) or UpdateEndpointWeightsAndCapacities (p. 771) is executing.
- **SystemUpdating**: Endpoint is undergoing maintenance and cannot be updated or deleted or re-scaled until it has completed. This maintenance operation does not change any customer-specified values such as VPC config, KMS encryption, model, instance type, or instance count.
- **RollingBack**: Endpoint fails to scale up or down or change its variant weight and is in the process of rolling back to its previous configuration. Once the rollback completes, endpoint returns to an **InService** status. This transitional status only applies to an endpoint that has autoscaling enabled and is undergoing variant weight or capacity changes as part of an **UpdateEndpointWeightsAndCapacities (p. 771)** call or when the **UpdateEndpointWeightsAndCapacities (p. 771)** operation is called explicitly.
- **InService**: Endpoint is available to process incoming requests.
- **Deleting**: DeleteEndpoint (p. 225) is executing.
- **Failed**: Endpoint could not be created, updated, or re-scaled. Use DescribeEndpoint:FailureReason (p. 341) for information about the failure. DeleteEndpoint (p. 225) is the only operation that can be performed on a failed endpoint.

Type: String

Valid Values: OutOfService | Creating | Updating | SystemUpdating | RollingBack | InService | Deleting | Failed

**FailureReason (p. 338)**

If the status of the endpoint is Failed, the reason why it failed.

Type: String

Length Constraints: Maximum length of 1024.

**LastDeploymentConfig (p. 338)**

The most recent deployment configuration for the endpoint.

Type: DeploymentConfig (p. 987) object


**LastModifiedTime (p. 338)**

A timestamp that shows when the endpoint was last modified.

Type: Timestamp

**PendingDeploymentSummary (p. 338)**

Returns the summary of an in-progress deployment. This field is only returned when the endpoint is creating or updating with a new endpoint configuration.

Type: PendingDeploymentSummary (p. 1253) object

**ProductionVariants (p. 338)**

An array of ProductionVariantSummary (p. 1299) objects, one for each model hosted behind this endpoint.

Type: Array of ProductionVariantSummary (p. 1299) objects

Array Members: Minimum number of 1 item.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeEndpointConfig
Service: Amazon SageMaker Service

Returns the description of an endpoint configuration created using the CreateEndpointConfig API.

Request Syntax

```json
{
    "EndpointConfigName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**EndpointConfigName (p. 343)**

The name of the endpoint configuration.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Syntax

```json
{
    "AsyncInferenceConfig": {
        "ClientConfig": {
            "MaxConcurrentInvocationsPerInstance": number
        },
        "OutputConfig": {
            "KmsKeyId": "string",
            "NotificationConfig": {
                "ErrorTopic": "string",
                "SuccessTopic": "string"
            },
            "S3OutputPath": "string"
        }
    },
    "CreationTime": number,
    "DataCaptureConfig": {
        "CaptureContentTypeHeader": {
            "CsvContentTypes": [ "string" ],
            "JsonContentTypes": [ "string" ]
        },
        "CaptureOptions": [
            {
                "CaptureMode": "string"
            }
        ],
        "DestinationS3Uri": "string",
```
"EnableCapture": boolean,
  "InitialSamplingPercentage": number,
  "KmsKeyId": "string"
},
"EndpointConfigArn": "string",
"EndpointConfigName": "string",
"KmsKeyId": "string",
"ProductionVariants": [
  {
    "AcceleratorType": "string",
    "CoreDumpConfig": {
      "DestinationS3Uri": "string",
      "KmsKeyId": "string"
    },
    "InitialInstanceCount": number,
    "InitialVariantWeight": number,
    "InstanceType": "string",
    "ModelName": "string",
    "ServerlessConfig": {
      "MaxConcurrency": number,
      "MemorySizeInMB": number
    },
    "VariantName": "string"
  }
]
]

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AsyncInferenceConfig (p. 343)

Returns the description of an endpoint configuration created using the CreateEndpointConfig API.

Type: AsyncInferenceConfig (p. 899) object

CreationTime (p. 343)

A timestamp that shows when the endpoint configuration was created.

Type: Timestamp

DataCaptureConfig (p. 343)

Configuration to control how SageMaker captures inference data.

Type: DataCaptureConfig (p. 966) object

EndpointConfigArn (p. 343)

The Amazon Resource Name (ARN) of the endpoint configuration.

Type: String


Pattern: arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:endpoint-config/.*

EndpointConfigName (p. 343)

Name of the SageMaker endpoint configuration.
Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$

**KmsKeyId (p. 343)**

AWS KMS key ID Amazon SageMaker uses to encrypt data when storing it on the ML storage volume attached to the instance.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

**ProductionVariants (p. 343)**

An array of ProductionVariant objects, one for each model that you want to host at this endpoint.

Type: Array of ProductionVariant (p. 1293) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeExperiment
Service: Amazon SageMaker Service
Provides a list of an experiment’s properties.

Request Syntax

```json
{
    "ExperimentName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ExperimentName** (p. 346)

The name of the experiment to describe.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: Yes

Response Syntax

```json
{
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "Description": "string",
    "DisplayName": "string",
    "ExperimentArn": "string",
    "ExperimentName": "string",
    "LastModifiedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "LastModifiedTime": number,
    "Source": {
        "SourceArn": "string",
        "SourceType": "string"
    }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**CreatedBy (p. 346)**
Who created the experiment.
Type: UserContext (p. 1449) object

**CreationTime (p. 346)**
When the experiment was created.
Type: Timestamp

**Description (p. 346)**
The description of the experiment.
Type: String
Length Constraints: Maximum length of 3072.
Pattern: .*

**DisplayName (p. 346)**
The name of the experiment as displayed. If `DisplayName` isn't specified, `ExperimentName` is displayed.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

**ExperimentArn (p. 346)**
The Amazon Resource Name (ARN) of the experiment.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-zA-Z\-]*:sagemaker:[a-zA-Z\-]*:[0-9]\{12\}:experiment/.*

**ExperimentName (p. 346)**
The name of the experiment.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

**LastModifiedBy (p. 346)**
Who last modified the experiment.
Type: UserContext (p. 1449) object

**LastModifiedTime (p. 346)**
When the experiment was last modified.
Type: Timestamp
Source (p. 346)

The ARN of the source and, optionally, the type.

Type: ExperimentSource (p. 1031) object

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeFeatureGroup

Use this operation to describe a FeatureGroup. The response includes information on the creation time, FeatureGroup name, the unique identifier for each FeatureGroup, and more.

Request Syntax

```json
{
  "FeatureGroupName": "string",
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**FeatureGroupName (p. 349)**

The name of the FeatureGroup you want described.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,63}`

Required: Yes

**NextToken (p. 349)**

A token to resume pagination of the list of Features (FeatureDefinitions). 2,500 Features are returned by default.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: `.*`

Required: No

Response Syntax

```json
{
  "CreationTime": number,
  "Description": "string",
  "EventTimeFeatureName": "string",
  "FailureReason": "string",
  "FeatureDefinitions": [
    {
      "FeatureName": "string",
      "FeatureType": "string"
    }
  ],
}
```
DescribeFeatureGroup

"FeatureGroupArn": "string",
"FeatureGroupName": "string",
"FeatureGroupStatus": "string",
"NextToken": "string",
"OfflineStoreConfig": {
  "DataCatalogConfig": {
    "Catalog": "string",
    "Database": "string",
    "TableName": "string"
  },
  "DisableGlueTableCreation": boolean,
  "S3StorageConfig": {
    "KmsKeyId": "string",
    "ResolvedOutputS3Uri": "string",
    "S3Uri": "string"
  }
},
"OfflineStoreStatus": {
  "BlockedReason": "string",
  "Status": "string"
},
"OnlineStoreConfig": {
  "EnableOnlineStore": boolean,
  "SecurityConfig": {
    "KmsKeyId": "string"
  }
},
"RecordIdentifierFeatureName": "string",
"RoleArn": "string"

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreationTime (p. 349)

A timestamp indicating when SageMaker created the FeatureGroup.

Type: Timestamp

Description (p. 349)

A free form description of the feature group.

Type: String

Length Constraints: Maximum length of 128.

EventTimeFeatureName (p. 349)

The name of the feature that stores the EventTime of a Record in a FeatureGroup.

An EventTime is a point in time when a new event occurs that corresponds to the creation or update of a Record in a FeatureGroup. All Records in the FeatureGroup have a corresponding EventTime.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9][\-_][a-zA-Z0-9]{0,63}$
**FailureReason (p. 349)**

The reason that the FeatureGroup failed to be replicated in the OfflineStore. This is failure can occur because:

- The FeatureGroup could not be created in the OfflineStore.
- The FeatureGroup could not be deleted from the OfflineStore.

Type: String

Length Constraints: Maximum length of 1024.

**FeatureDefinitions (p. 349)**

A list of the Features in the FeatureGroup. Each feature is defined by a FeatureName and FeatureType.

Type: Array of FeatureDefinition (p. 1036) objects

Array Members: Minimum number of 1 item. Maximum number of 2500 items.

**FeatureGroupArn (p. 349)**

The Amazon Resource Name (ARN) of the FeatureGroup.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:0-9{12}:feature-group/.*

**FeatureGroupName (p. 349)**

The name of the FeatureGroup.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9]([-]*[a-zA-Z0-9])\{0,63}\$

**FeatureGroupStatus (p. 349)**

The status of the feature group.

Type: String

Valid Values: Creating | Created | CreateFailed | Deleting | DeleteFailed

**NextToken (p. 349)**

A token to resume pagination of the list of Features (FeatureDefinitions).

Type: String

Length Constraints: Maximum length of 8192.

Pattern: \.

**OfflineStoreConfig (p. 349)**

The configuration of the OfflineStore, inducing the S3 location of the OfflineStore, AWS Glue or AWS Hive data catalogue configurations, and the security configuration.

Type: OfflineStoreConfig (p. 1231) object
**OfflineStoreStatus (p. 349)**

The status of the OfflineStore. Notifies you if replicating data into the OfflineStore has failed. Returns either: Active or Blocked

Type: OfflineStoreStatus (p. 1232) object

**OnlineStoreConfig (p. 349)**

The configuration for the OnlineStore.

Type: OnlineStoreConfig (p. 1238) object

**RecordIdentifierFeatureName (p. 349)**

The name of the Feature used for RecordIdentifier, whose value uniquely identifies a record stored in the feature store.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9][\-_]*[a-zA-Z0-9]{0,63}

**RoleArn (p. 349)**

The Amazon Resource Name (ARN) of the IAM execution role used to persist data into the OfflineStore if an OfflineStoreConfig is provided.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z-0-9+=,.@\-_/]+$

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

352
DescribeFlowDefinition
Service: Amazon SageMaker Service

Returns information about the specified flow definition.

Request Syntax

```json
{
   "FlowDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**FlowDefinitionName (p. 354)**

The name of the flow definition.

Type: String


Pattern: ^[a-z0-9](-*[a-z0-9]){0,62}

Required: Yes

Response Syntax

```json
{
   "CreationTime": number,
   "FailureReason": "string",
   "FlowDefinitionArn": "string",
   "FlowDefinitionName": "string",
   "FlowDefinitionStatus": "string",
   "HumanLoopActivationConfig": {
      "HumanLoopActivationConditionsConfig": {
         "HumanLoopActivationConditions": "string"
      }
   },
   "HumanLoopConfig": {
      "HumanTaskUiArn": "string",
      "PublicWorkforceTaskPrice": {
         "AmountInUsd": {
            "Cents": number,
            "Dollars": number,
            "TenthFractionsOfACent": number
         }
      },
      "TaskAvailabilityLifetimeInSeconds": number,
      "TaskCount": number,
      "TaskDescription": "string",
      "TaskKeywords": [ "string" ],
      "TaskTimeLimitInSeconds": number,
      "TaskTitle": "string"
   }
}
```
### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 354)**

The timestamp when the flow definition was created.

Type: Timestamp

**FailureReason (p. 354)**

The reason your flow definition failed.

Type: String

Length Constraints: Maximum length of 1024.

**FlowDefinitionArn (p. 354)**

The Amazon Resource Name (ARN) of the flow definition.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:flow-definition/.*`

**FlowDefinitionName (p. 354)**

The Amazon Resource Name (ARN) of the flow definition.

Type: String


Pattern: `^[a-z0-9\-_a-z0-9]{0,62}`

**FlowDefinitionStatus (p. 354)**

The status of the flow definition. Valid values are listed below.

Type: String

Valid Values: Initializing | Active | Failed | Deleting

**HumanLoopActivationConfig (p. 354)**

An object containing information about what triggers a human review workflow.

Type: `HumanLoopActivationConfig (p. 1057)` object

---

```json
{
   "WorkteamArn": "string",
   "HumanLoopRequestSource": {
      "AwsManagedHumanLoopRequestSource": "string"
   },
   "OutputConfig": {
      "KmsKeyId": "string",
      "S3OutputPath": "string"
   },
   "RoleArn": "string"
}
```
**HumanLoopConfig** (p. 354)

An object containing information about who works on the task, the workforce task price, and other task details.

Type: `HumanLoopConfig` (p. 1058) object

**HumanLoopRequestSource** (p. 354)

Container for configuring the source of human task requests. Used to specify if Amazon Rekognition or Amazon Textract is used as an integration source.

Type: `HumanLoopRequestSource` (p. 1063) object

**OutputConfig** (p. 354)

An object containing information about the output file.

Type: `FlowDefinitionOutputConfig` (p. 1051) object

**RoleArn** (p. 354)

The Amazon Resource Name (ARN) of the AWS Identity and Access Management (IAM) execution role for the flow definition.

Type: String


Pattern: `^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_/]+$`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
**DescribeHumanTaskUi**

Service: Amazon SageMaker Service

Returns information about the requested human task user interface (worker task template).

**Request Syntax**

```json
{
   "HumanTaskUiName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**HumanTaskUiName (p. 357)**

The name of the human task user interface (worker task template) you want information about.

Type: String


Pattern: `^[a-z0-9][a-z0-9]*$`

Required: Yes

**Response Syntax**

```json
{
   "CreationTime": number,
   "HumanTaskUiArn": "string",
   "HumanTaskUiName": "string",
   "HumanTaskUiStatus": "string",
   "UiTemplate": {
      "ContentSha256": "string",
      "Url": "string"
   }
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 357)**

The timestamp when the human task user interface was created.

Type: Timestamp

**HumanTaskUiArn (p. 357)**

The Amazon Resource Name (ARN) of the human task user interface (worker task template).
Type: String

Length Constraints: Maximum length of 1024.

Pattern: \(arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:human-task-ui/.*\)

**HumanTaskUiName (p. 357)**

The name of the human task user interface (worker task template).

Type: String


Pattern: ^[a-z0-9](-*[a-z0-9])*\n
**HumanTaskUiStatus (p. 357)**

The status of the human task user interface (worker task template). Valid values are listed below.

Type: String

Valid Values: Active | Deleting

**UiTemplate (p. 357)**

Container for user interface template information.

Type: UiTemplateInfo (p. 1447) object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeHyperParameterTuningJob

Service: Amazon SageMaker Service

Gets a description of a hyperparameter tuning job.

Request Syntax

```json
{
    "HyperParameterTuningJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**HyperParameterTuningJobName (p. 359)**

The name of the tuning job.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}$

Required: Yes

Response Syntax

```json
{
    "BestTrainingJob": {
        "CreationTime": number,
        "FailureReason": "string",
        "FinalHyperParameterTuningJobObjectiveMetric": {
            "MetricName": "string",
            "Type": "string",
            "Value": number
        },
        "ObjectiveStatus": "string",
        "TrainingEndTime": number,
        "TrainingJobArn": "string",
        "TrainingJobDefinitionName": "string",
        "TrainingJobName": "string",
        "TrainingJobStatus": "string",
        "TrainingStartTime": number,
        "TunedHyperParameters": {
            "string": "string"
        },
        "TuningJobName": "string"
    },
    "CreationTime": number,
    "FailureReason": "string",
    "HyperParameterTuningEndTime": number,
    "HyperParameterTuningJobArn": "string",
    "HyperParameterTuningJobConfig": {
```
"HyperParameterTuningJobObjective": {  
  "MetricName": "string",  
  "Type": "string"  
},  
"ParameterRanges": {  
  "CategoricalParameterRanges": [  
    {  
      "Values": [ "string" ]  
    }  
  ],  
  "ContinuousParameterRanges": [  
    {  
      "MaxValue": "string",  
      "MinValue": "string",  
      "Name": "string",  
      " ScalingType": "string"  
    }  
  ],  
  "IntegerParameterRanges": [  
    {  
      "MaxValue": "string",  
      "MinValue": "string",  
      "Name": "string",  
      "ScalingType": "string"  
    }  
  ]  
},  
"ResourceLimits": {  
  "MaxNumberOfTrainingJobs": number,  
  "MaxParallelTrainingJobs": number  
},  
"Strategy": "string",  
"TrainingJobEarlyStoppingType": "string",  
"TuningJobCompletionCriteria": {  
  "TargetObjectiveMetricValue": number  
},  
"HyperParameterTuningJobName": "string",  
"HyperParameterTuningJobStatus": "string",  
"LastModifiedTime": number,  
"ObjectiveStatusCounters": {  
  "Failed": number,  
  "Pending": number,  
  "Succeeded": number  
},  
"OverallBestTrainingJob": {  
  "CreationTime": number,  
  "FailureReason": "string",  
  "FinalHyperParameterTuningJobObjectiveMetric": {  
    "MetricName": "string",  
    "Type": "string",  
    "Value": number  
  },  
  "ObjectiveStatus": "string",  
  "TrainingEndTime": number,  
  "TrainingJobArn": "string",  
  "TrainingJobDefinitionName": "string",  
  "TrainingJobName": "string",  
  "TrainingJobStatus": "string",  
  "TrainingStartTime": number,  
  "TunedHyperParameters": {  
    "string": "string"  
  },  
  "TuningJobName": "string"  
},
"TrainingJobDefinition": {
    "AlgorithmSpecification": {
        "AlgorithmName": "string",
        "MetricDefinitions": [
            {
                "Name": "string",
                "Regex": "string"
            }
        ],
        "TrainingImage": "string",
        "TrainingInputMode": "string"
    },
    "CheckpointConfig": {
        "LocalPath": "string",
        "S3Uri": "string"
    },
    "DefinitionName": "string",
    "EnableInterContainerTrafficEncryption": boolean,
    "EnableManagedSpotTraining": boolean,
    "EnableNetworkIsolation": boolean,
    "HyperParameterRanges": {
        "CategoricalParameterRanges": [
            {
                "Name": "string",
                "Values": [ "string" ]
            }
        ],
        "ContinuousParameterRanges": [
            {
                "MaxValue": "string",
                "MinValue": "string",
                "Name": "string",
                "ScalingType": "string"
            }
        ],
        "IntegerParameterRanges": [
            {
                "MaxValue": "string",
                "MinValue": "string",
                "Name": "string",
                "ScalingType": "string"
            }
        ]
    },
    "InputDataConfig": [
        {
            "ChannelName": "string",
            "CompressionType": "string",
            "ContentType": "string",
            "DataSource": {
                "FileSystemDataSource": {
                    "DirectoryPath": "string",
                    "FileSystemAccessMode": "string",
                    "FileSystemId": "string",
                    "FileSystemType": "string"
                },
                "S3DataSource": {
                    "AttributeNames": [ "string" ],
                    "S3DataDistributionType": "string",
                    "S3DataType": "string",
                    "S3Uri": "string"
                }
            },
            "InputMode": "string",
            "RecordWrapperType": "string",
            "ShuffleConfig": {
            }
        }
    ]
}
"Seed": number

"OutputDataConfig": {
  "KmsKeyId": "string",
  "S3OutputPath": "string"
},

"ResourceConfig": {
  "InstanceCount": number,
  "InstanceType": "string",
  "VolumeKmsKeyId": "string",
  "VolumeSizeInGB": number
},

"RetryStrategy": {
  "MaximumRetryAttempts": number
},

"RoleArn": "string",

"StaticHyperParameters": {
  "string": "string"
},

"StoppingCondition": {
  "MaxRuntimeInSeconds": number,
  "MaxWaitTimeInSeconds": number
},

"TuningObjective": {
  "MetricName": "string",
  "Type": "string"
},

"VpcConfig": {
  "SecurityGroupIds": [ "string" ],
  "Subnets": [ "string" ]
}

"TrainingJobDefinitions": [
  {
    "AlgorithmSpecification": {
      "AlgorithmName": "string",
      "MetricDefinitions": [
        {
          "Name": "string",
          "Regex": "string"
        }
      ],
      "TrainingImage": "string",
      "TrainingInputMode": "string"
    },
    "CheckpointConfig": {
      "LocalPath": "string",
      "S3Uri": "string"
    },
    "DefinitionName": "string",
    "EnableInterContainerTrafficEncryption": boolean,
    "EnableManagedSpotTraining": boolean,
    "EnableNetworkIsolation": boolean,
    "HyperParameterRanges": {
      "CategoricalParameterRanges": [
        {
          "Name": "string",
          "Values": [ "string" ]
        }
      ],
      "ContinuousParameterRanges": [
        {
          "MaxValue": "string",
          "MinValue": "string",
          "Name": "string",
          "Values": [ "string" ]
        }
      ]
    }
  }
]
"Name": "string",
"ScalingType": "string"
],
"IntegerParameterRanges": [ 
  
  
  "MaxValue": "string",
  "MinValue": "string",
  "Name": "string",
  "ScalingType": "string"
  
],
"InputDataConfig": [ 
  
  "ChannelName": "string",
  "CompressionType": "string",
  "ContentType": "string",
  "DataSource": { 
    "FileSystemDataSource": { 
      "DirectoryPath": "string",
      "FileSystemAccessMode": "string",
      "FileSystemId": "string",
      "FileSystemType": "string"
    },
    "S3DataSource": { 
      "AttributeNames": [ "string" ],
      "S3DataDistributionType": "string",
      "S3DataType": "string",
      "S3Uri": "string"
    }
  },
  "InputMode": "string",
  "RecordWrapperType": "string",
  "ShuffleConfig": { 
    "Seed": number
  }
],
"OutputDataConfig": { 
  "KmsKeyId": "string",
  "S3OutputPath": "string"
},
"ResourceConfig": { 
  "InstanceCount": number,
  "InstanceType": "string",
  "VolumeKmsKeyId": "string",
  "VolumeSizeInGB": number
},
"RetryStrategy": { 
  "MaximumRetryAttempts": number
},
"RoleArn": "string",
"StaticHyperParameters": { 
  "string": "string"
},
"StoppingCondition": { 
  "MaxRuntimeInSeconds": number,
  "MaxWaitTimeInSeconds": number
},
"TuningObjective": { 
  "MetricName": "string",
  "Type": "string"
},
"VpcConfig": { 
  "SecurityGroupIds": [ "string" ],
"Subnets": [ "string" ]
}

"TrainingJobStatusCounters": {
  "Completed": number,
  "InProgress": number,
  "NonRetryableError": number,
  "RetryableError": number,
  "Stopped": number
},
"WarmStartConfig": {
  "ParentHyperParameterTuningJobs": [
    {
      "HyperParameterTuningJobName": "string"
    }
  ],
  "WarmStartType": "string"
}
}

### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**BestTrainingJob (p. 359)**

A TrainingJobSummary (p. 1398) object that describes the training job that completed with the best current HyperParameterTuningJobObjective (p. 1091).

Type: HyperParameterTrainingJobSummary (p. 1086) object

**CreationTime (p. 359)**

The date and time that the tuning job started.

Type: Timestamp

**FailureReason (p. 359)**

If the tuning job failed, the reason it failed.

Type: String

Length Constraints: Maximum length of 1024.

**HyperParameterTuningEndTime (p. 359)**

The date and time that the tuning job ended.

Type: Timestamp

**HyperParameterTuningJobArn (p. 359)**

The Amazon Resource Name (ARN) of the tuning job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:hyper-parameter-tuning-job/.*
HyperParameterTuningJobConfig (p. 359)

The HyperParameterTuningJobConfig (p. 1089) object that specifies the configuration of the tuning job.

Type: HyperParameterTuningJobConfig (p. 1089) object

HyperParameterTuningJobName (p. 359)

The name of the tuning job.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}

HyperParameterTuningJobStatus (p. 359)

The status of the tuning job: InProgress, Completed, Failed, Stopping, or Stopped.

Type: String

Valid Values: Completed | InProgress | Failed | Stopped | Stopping

LastModifiedTime (p. 359)

The date and time that the status of the tuning job was modified.

Type: Timestamp

ObjectiveStatusCounters (p. 359)

The ObjectiveStatusCounters (p. 1230) object that specifies the number of training jobs, categorized by the status of their final objective metric, that this tuning job launched.

Type: ObjectiveStatusCounters (p. 1230) object

OverallBestTrainingJob (p. 359)

If the hyperparameter tuning job is an warm start tuning job with a WarmStartType of IDENTICAL_DATA_AND_ALGORITHM, this is the TrainingJobSummary (p. 1398) for the training job with the best objective metric value of all training jobs launched by this tuning job and all parent jobs specified for the warm start tuning job.

Type: HyperParameterTrainingJobSummary (p. 1086) object

TrainingJobDefinition (p. 359)

The HyperParameterTrainingJobDefinition (p. 1082) object that specifies the definition of the training jobs that this tuning job launches.

Type: HyperParameterTrainingJobDefinition (p. 1082) object

TrainingJobDefinitions (p. 359)

A list of the HyperParameterTrainingJobDefinition (p. 1082) objects launched for this tuning job.

Type: Array of HyperParameterTrainingJobDefinition (p. 1082) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

TrainingJobStatusCounters (p. 359)

The TrainingJobStatusCounters (p. 1395) object that specifies the number of training jobs, categorized by status, that this tuning job launched.
Type: TrainingJobStatusCounters (p. 1395) object

**WarmStartConfig (p. 359)**

The configuration for starting the hyperparameter parameter tuning job using one or more previous tuning jobs as a starting point. The results of previous tuning jobs are used to inform which combinations of hyperparameters to search over in the new tuning job.

Type: HyperParameterTuningJobWarmStartConfig (p. 1094) object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeImage
Service: Amazon SageMaker Service

Describes a SageMaker image.

Request Syntax

```
{
  "ImageName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ImageName (p. 367)**

The name of the image to describe.

Type: String


Pattern: ^[a-zA-Z0-9]([-.]?[a-zA-Z0-9]{0,62}$

Required: Yes

Response Syntax

```
{
  "CreationTime": number,
  "Description": "string",
  "DisplayName": "string",
  "FailureReason": "string",
  "ImageArn": "string",
  "ImageName": "string",
  "ImageStatus": "string",
  "LastModifiedTime": number,
  "RoleArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 367)**

When the image was created.

Type: Timestamp
Description (p. 367)
The description of the image.
Type: String
Pattern: .*

DisplayName (p. 367)
The name of the image as displayed.
Type: String
Pattern: ^\S(.*\S)?$ 

FailureReason (p. 367)
When a create, update, or delete operation fails, the reason for the failure.
Type: String
Length Constraints: Maximum length of 1024.

ImageArn (p. 367)
The Amazon Resource Name (ARN) of the image.
Type: String
Length Constraints: Maximum length of 256.
Pattern: ^arn:aws([-\w]+)*:sagemaker:.+:[0-9]{12}:image/[a-z0-9](\[-.\]?[a-z0-9])*$

ImageName (p. 367)
The name of the image.
Type: String
Pattern: ^[a-zA-Z0-9][-._]?[a-zA-Z0-9]{0,62}$

ImageStatus (p. 367)
The status of the image.
Type: String
Valid Values: CREATING | CREATED | CREATE FAILED | UPDATING | UPDATE FAILED | DELETING | DELETE FAILED

LastModifiedTime (p. 367)
When the image was last modified.
Type: Timestamp

RoleArn (p. 367)
The Amazon Resource Name (ARN) of the IAM role that enables Amazon SageMaker to perform tasks on your behalf.
Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@-\_/]+$
DescribeImageVersion
Service: Amazon SageMaker Service
Describes a version of a SageMaker image.

Request Syntax

```json
{
    "ImageName": "string",
    "Version": number
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ImageName (p. 370)**

The name of the image.

Type: String


Pattern: `^[a-zA-Z0-9]([-._]?[a-zA-Z0-9]{0,62}$

Required: Yes

**Version (p. 370)**

The version of the image. If not specified, the latest version is described.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Response Syntax

```json
{
    "BaseImage": "string",
    "ContainerImage": "string",
    "CreationTime": number,
    "FailureReason": "string",
    "ImageArn": "string",
    "ImageVersionArn": "string",
    "ImageVersionStatus": "string",
    "LastModifiedTime": number,
    "Version": number
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**BaseImage (p. 370)**

The registry path of the container image on which this image version is based.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .*

**ContainerImage (p. 370)**

The registry path of the container image that contains this image version.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

**CreationTime (p. 370)**

When the version was created.

Type: Timestamp

**FailureReason (p. 370)**

When a create or delete operation fails, the reason for the failure.

Type: String

Length Constraints: Maximum length of 1024.

**ImageArn (p. 370)**

The Amazon Resource Name (ARN) of the image the version is based on.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws(-[\w]+)*:sagemaker:.+:[0-9]{12}:image/[a-z0-9][-_.]?[a-z0-9]+$

**ImageVersionArn (p. 370)**

The ARN of the version.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws(-[\w]+)*:sagemaker:.+:[0-9]{12}:image-version/[a-z0-9][-_.]?[a-z0-9]+$/[0-9]+$

**ImageVersionStatus (p. 370)**

The status of the version.

Type: String

Valid Values: CREATING | CREATED | CREATE_FAILED | DELETING | DELETE_FAILED

**LastModifiedTime (p. 370)**

When the version was last modified.
Type: Timestamp

**Version (p. 370)**

The version number.

Type: Integer

Valid Range: Minimum value of 0.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeInferenceRecommendationsJob

Service: Amazon SageMaker Service

Provides the results of the Inference Recommender job. One or more recommendation jobs are returned.

Request Syntax

```
{
  "JobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobName (p. 373)**

The name of the job. The name must be unique within an AWS Region in the AWS account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,63}$

Required: Yes

Response Syntax

```
{
  "CompletionTime": number,
  "CreationTime": number,
  "FailureReason": "string",
  "InferenceRecommendations": [
    {
      "EndpointConfiguration": {
        "EndpointName": "string",
        "InitialInstanceCount": number,
        "InstanceType": "string",
        "VariantName": "string"
      },
      "Metrics": {
        "CostPerHour": number,
        "CostPerInference": number,
        "MaxInvocations": number,
        "ModelLatency": number
      },
      "ModelConfiguration": {
        "EnvironmentParameters": [
          {
            "Key": "string",
            "Value": "string",
            "ValueType": "string"
          }
        ]
      }
    }
  ]
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CompletionTime (p. 373)

A timestamp that shows when the job completed.
Type: Timestamp

**CreationTime (p. 373)**

A timestamp that shows when the job was created.

Type: Timestamp

**FailureReason (p. 373)**

If the job fails, provides information why the job failed.

Type: String

Length Constraints: Maximum length of 1024.

**InferenceRecommendations (p. 373)**

The recommendations made by Inference Recommender.

Type: Array of **InferenceRecommendation (p. 1102)** objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

**InputConfig (p. 373)**

Returns information about the versioned model package Amazon Resource Name (ARN), the traffic pattern, and endpoint configurations you provided when you initiated the job.

Type: **RecommendationJobInputConfig (p. 1325)** object

**JobArn (p. 373)**

The Amazon Resource Name (ARN) of the job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:inference-recommendations-job/.*`

**JobDescription (p. 373)**

The job description that you provided when you initiated the job.

Type: String

Length Constraints: Maximum length of 128.

**JobName (p. 373)**

The name of the job. The name must be unique within an AWS Region in the AWS account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9]{0,63}$`

**JobType (p. 373)**

The job type that you provided when you initiated the job.

Type: String

Valid Values: Default | Advanced
LastModifiedTime (p. 373)

A timestamp that shows when the job was last modified.

Type: Timestamp

RoleArn (p. 373)

The Amazon Resource Name (ARN) of the AWS Identity and Access Management (IAM) role you provided when you initiated the job.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_\/]++$

Status (p. 373)

The status of the job.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | FAILED | STOPPING | STOPPED

StoppingConditions (p. 373)

The stopping conditions that you provided when you initiated the job.

Type: RecommendationJobStoppingConditions (p. 1329) object

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeLabelingJob

Service: Amazon SageMaker Service

Gets information about a labeling job.

Request Syntax

```json
{
    "LabelingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**LabelingJobName (p. 377)**

The name of the labeling job to return information for.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Syntax

```json
{
    "CreationTime": number,
    "FailureReason": "string",
    "HumanTaskConfig": {
        "AnnotationConsolidationConfig": {
            "AnnotationConsolidationLambdaArn": "string"
        },
        "MaxConcurrentTaskCount": number,
        "NumberOfHumanWorkersPerDataObject": number,
        "PreHumanTaskLambdaArn": "string",
        "PublicWorkforceTaskPrice": {
            "AmountInUsd": {
                "Cents": number,
                "Dollars": number,
                "TenthFractionsOfACent": number
            }
        },
        "TaskAvailabilityLifetimeInSeconds": number,
        "TaskDescription": "string",
        "TaskKeywords": [ "string" ],
        "TaskTimeLimitInSeconds": number,
        "TaskTitle": "string",
        "UiConfig": {
            "HumanTaskUiArn": "string",
            "UiTemplateS3Uri": "string"
        }
    }
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
CreationTime (p. 377)

The date and time that the labeling job was created.

Type: Timestamp

FailureReason (p. 377)

If the job failed, the reason that it failed.

Type: String

Length Constraints: Maximum length of 1024.

HumanTaskConfig (p. 377)

Configuration information required for human workers to complete a labeling task.

Type: HumanTaskConfig (p. 1064) object

InputConfig (p. 377)

Input configuration information for the labeling job, such as the Amazon S3 location of the data objects and the location of the manifest file that describes the data objects.

Type: LabelingJobInputConfig (p. 1128) object

JobReferenceCode (p. 377)

A unique identifier for work done as part of a labeling job.

Type: String

Length Constraints: Minimum length of 1.

Pattern: .+

LabelAttributeName (p. 377)

The attribute used as the label in the output manifest file.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,126}$

LabelCategoryConfigS3Uri (p. 377)

The S3 location of the JSON file that defines the categories used to label data objects. Please note the following label-category limits:

- Semantic segmentation labeling jobs using automated labeling: 20 labels
- Box bounding labeling jobs (all): 10 labels

The file is a JSON structure in the following format:

```json
{
  "document-version": "2018-11-28"
  "labels": [
    {
      "label": "label 1"
    }
  ]
}
```
...}

{
  "label": "label \eta"
}
]

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)/?(.*)$

**LabelCounters (p. 377)**

Provides a breakdown of the number of data objects labeled by humans, the number of objects labeled by machine, the number of objects than couldn't be labeled, and the total number of objects labeled.

Type: **LabelCounters (p. 1119) object**

**LabelingJobAlgorithmsConfig (p. 377)**

Configuration information for automated data labeling.

Type: **LabelingJobAlgorithmsConfig (p. 1122) object**

**LabelingJobArn (p. 377)**

The Amazon Resource Name (ARN) of the labeling job.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:labeling-job/.*

**LabelingJobName (p. 377)**

The name assigned to the labeling job when it was created.

Type: String


Pattern: ^[a-zA-Z0-9-]{0,62}$

**LabelingJobOutput (p. 377)**

The location of the output produced by the labeling job.

Type: **LabelingJobOutput (p. 1129) object**
**LabelingJobStatus (p. 377)**

The processing status of the labeling job.

Type: String

Valid Values: Initializing | InProgress | Completed | Failed | Stopping | Stopped

**LastModifiedTime (p. 377)**

The date and time that the labeling job was last updated.

Type: Timestamp

**OutputConfig (p. 377)**

The location of the job's output data and the AWS Key Management Service key ID for the key used to encrypt the output data, if any.

Type: LabelingJobOutputConfig (p. 1130) object

**RoleArn (p. 377)**

The Amazon Resource Name (ARN) that SageMaker assumes to perform tasks on your behalf during data labeling.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9\-_\+\=\,\@\-\_/]+$+

**StoppingConditions (p. 377)**

A set of conditions for stopping a labeling job. If any of the conditions are met, the job is automatically stopped.

Type: LabelingJobStoppingConditions (p. 1135) object

**Tags (p. 377)**

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DescribeLineageGroup
Service: Amazon SageMaker Service

Provides a list of properties for the requested lineage group. For more information, see Cross-Account Lineage Tracking in the Amazon SageMaker Developer Guide.

Request Syntax

```json
{
   "LineageGroupName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

LineageGroupName (p. 383)

The name of the lineage group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: Yes

Response Syntax

```json
{
   "CreatedBy": {
      "DomainId": "string",
      "UserProfileArn": "string",
      "UserProfileName": "string"
   },
   "CreationTime": number,
   "Description": "string",
   "DisplayName": "string",
   "LastModifiedBy": {
      "DomainId": "string",
      "UserProfileArn": "string",
      "UserProfileName": "string"
   },
   "LastModifiedTime": number,
   "LineageGroupArn": "string",
   "LineageGroupName": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
CreatedBy (p. 383)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

CreationTime (p. 383)

The creation time of lineage group.

Type: Timestamp

Description (p. 383)

The description of the lineage group.

Type: String

Length Constraints: Maximum length of 3072.

Pattern: .*

DisplayName (p. 383)

The display name of the lineage group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,119\}

LastModifiedBy (p. 383)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

LastModifiedTime (p. 383)

The last modified time of the lineage group.

Type: Timestamp

LineageGroupArn (p. 383)

The Amazon Resource Name (ARN) of the lineage group.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-zA-Z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:lineage-group/.*

LineageGroupName (p. 383)

The name of the lineage group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,119\}
**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](https://docs.aws.amazon.com/sagemaker/latest/dg/errors-reference.html).

**ResourceNotFound**

- Resource being access is not found.

  HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeModel
Service: Amazon SageMaker Service

Describes a model that you created using the CreateModel API.

Request Syntax

```json
{
   "ModelName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ModelName (p. 386)**

The name of the model.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*

Required: Yes

Response Syntax

```json
{
   "Containers": [
      {
         "ContainerHostname": "string",
         "Environment": {
            "string": "string"
         },
         "Image": "string",
         "ImageConfig": {
            "RepositoryAccessMode": "string",
            "RepositoryAuthConfig": {
               "RepositoryCredentialsProviderArn": "string"
            }
         },
         "InferenceSpecificationName": "string",
         "Mode": "string",
         "ModelDataUrl": "string",
         "ModelPackageName": "string",
         "MultiModelConfig": {
            "ModelCacheSetting": "string"
         }
      }
   ],
   "CreationTime": number,
   "EnableNetworkIsolation": boolean,
   "ExecutionRoleArn": "string",
```
"InferenceExecutionConfig": {
  "Mode": "string"
},
"ModelArn": "string",
"ModelName": "string",
"PrimaryContainer": {
  "ContainerHostname": "string",
  "Environment": {
    "string": "string"
  },
  "Image": "string",
  "ImageConfig": {
    "RepositoryAccessMode": "string",
    "RepositoryAuthConfig": {
      "RepositoryCredentialsProviderArn": "string"
    }
  },
  "InferenceSpecificationName": "string",
  "Mode": "string",
  "ModelDataUrl": "string",
  "ModelPackageName": "string",
  "MultiModelConfig": {
    "ModelCacheSetting": "string"
  }
},
"VpcConfig": {
  "SecurityGroupIds": [ "string" ],
  "Subnets": [ "string" ]
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Containers (p. 386)

The containers in the inference pipeline.

Type: Array of ContainerDefinition (p. 956) objects

Array Members: Maximum number of 15 items.

CreationTime (p. 386)

A timestamp that shows when the model was created.

Type: Timestamp

EnableNetworkIsolation (p. 386)

If True, no inbound or outbound network calls can be made to or from the model container.

Type: Boolean

ExecutionRoleArn (p. 386)

The Amazon Resource Name (ARN) of the IAM role that you specified for the model.

Type: String

Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_/]+$  

**InferenceExecutionConfig (p. 386)**

Specifies details of how containers in a multi-container endpoint are called.

Type: `InferenceExecutionConfig (p. 1101)` object

**ModelArn (p. 386)**

The Amazon Resource Name (ARN) of the model.

Type: `String`


Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model/.*`

**ModelName (p. 386)**

Name of the SageMaker model.

Type: `String`

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9\-]*([a-zA-Z0-9\-])*$`

**PrimaryContainer (p. 386)**

The location of the primary inference code, associated artifacts, and custom environment map that the inference code uses when it is deployed in production.

Type: `ContainerDefinition (p. 956)` object

**VpcConfig (p. 386)**

A `VpcConfig (p. 1456)` object that specifies the VPC that this model has access to. For more information, see Protect Endpoints by Using an Amazon Virtual Private Cloud

Type: `VpcConfig (p. 1456)` object

**Errors**

For information about the errors that are common to all actions, see `Common Errors (p. 1472)`.

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeModelBiasJobDefinition
Service: Amazon SageMaker Service

Returns a description of a model bias job definition.

Request Syntax

```
{
    "JobDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 389)**

The name of the model bias job definition. The name must be unique within an AWS Region in the AWS account.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$

Required: Yes

Response Syntax

```
{
    "CreationTime": number,
    "JobDefinitionArn": "string",
    "JobDefinitionName": "string",
    "JobResources": {
        "ClusterConfig": {
            "InstanceCount": number,
            "InstanceType": "string",
            "VolumeKmsKeyId": "string",
            "VolumeSizeInGB": number
        }
    },
    "ModelBiasAppSpecification": {
        "ConfigUri": "string",
        "Environment": {
            "string": "string"
        },
        "ImageUri": "string"
    },
    "ModelBiasBaselineConfig": {
        "BaseliningJobName": "string",
        "ConstraintsResource": {
            "S3Uri": "string"
        }
    }
}
```

389
"ModelBiasJobInput": {
  "EndpointInput": {
    "EndpointName": "string",
    "EndTimeOffset": "string",
    "FeaturesAttribute": "string",
    "InferenceAttribute": "string",
    "LocalPath": "string",
    "ProbabilityAttribute": "string",
    "ProbabilityThresholdAttribute": number,
    "S3DataDistributionType": "string",
    "S3InputMode": "string",
    "StartTimeOffset": "string"
  },
  "GroundTruthS3Input": {
    "S3Uri": "string"
  }
},
"ModelBiasJobOutputConfig": {
  "KmsKeyId": "string",
  "MonitoringOutputs": [
    {
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
},
"NetworkConfig": {
  "EnableInterContainerTrafficEncryption": boolean,
  "EnableNetworkIsolation": boolean,
  "VpcConfig": {
    "SecurityGroupIds": [ "string" ],
    "Subnets": [ "string" ]
  }
},
"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeInSeconds": number
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 389)**

The time at which the model bias job was created.

Type: Timestamp

**JobDefinitionArn (p. 389)**

The Amazon Resource Name (ARN) of the model bias job.

Type: String

Length Constraints: Maximum length of 256. 

Pattern: . *
JobDefinitionName (p. 389)
The name of the bias job definition. The name must be unique within an AWS Region in the AWS account.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$

JobResources (p. 389)
Identifies the resources to deploy for a monitoring job.
Type: MonitoringResources (p. 1210) object

ModelBiasAppSpecification (p. 389)
Configures the model bias job to run a specified Docker container image.
Type: ModelBiasAppSpecification (p. 1150) object

ModelBiasBaselineConfig (p. 389)
The baseline configuration for a model bias job.
Type: ModelBiasBaselineConfig (p. 1151) object

ModelBiasJobInput (p. 389)
Inputs for the model bias job.
Type: ModelBiasJobInput (p. 1152) object

ModelBiasJobOutputConfig (p. 389)
The output configuration for monitoring jobs.
Type: MonitoringOutputConfig (p. 1209) object

NetworkConfig (p. 389)
Networking options for a model bias job.
Type: MonitoringNetworkConfig (p. 1207) object

RoleArn (p. 389)
The Amazon Resource Name (ARN) of the AWS Identity and Access Management (IAM) role that has read permission to the input data location and write permission to the output data location in Amazon S3.
Type: String
Pattern: ^arn:aws[a-zA\:\-]*:iam::\d{12}:role/\?{a-zA-Z_0-9+=,.@\-_/]+$*

StoppingCondition (p. 389)
A time limit for how long the monitoring job is allowed to run before stopping.
Type: MonitoringStoppingCondition (p. 1219) object

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeModelExplainabilityJobDefinition

Service: Amazon SageMaker Service

Returns a description of a model explainability job definition.

Request Syntax

```
{
    "JobDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 393)**

The name of the model explainability job definition. The name must be unique within an AWS Region in the AWS account.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$

Required: Yes

Response Syntax

```
{
    "CreationTime": number,
    "JobDefinitionArn": "string",
    "JobDefinitionName": "string",
    "JobResources": {
        "ClusterConfig": {
            "InstanceCount": number,
            "InstanceType": "string",
            "VolumeKmsKeyId": "string",
            "VolumeSizeInGB": number
        }
    },
    "ModelExplainabilityAppSpecification": {
        "ConfigUri": "string",
        "Environment": {
            "string": "string"
        },
        "ImageUri": "string"
    },
    "ModelExplainabilityBaselineConfig": {
        "BaseliningJobName": "string",
        "ConstraintsResource": {
            "S3Uri": "string"
        }
    }
}
```

393
"EndpointInput": {  
  "EndpointName": "string",  
  "EndTimeOffset": "string",  
  "FeaturesAttribute": "string",  
  "InferenceAttribute": "string",  
  "LocalPath": "string",  
  "ProbabilityAttribute": "string",  
  "ProbabilityThresholdAttribute": number,  
  "S3DataDistributionType": "string",  
  "S3InputMode": "string",  
  "StartTimeOffset": "string"
}
},
"ModelExplainabilityJobOutputConfig": {
  "KmsKeyId": "string",
  "MonitoringOutputs": [
    {
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
}
},
"NetworkConfig": {
  "EnableInterContainerTrafficEncryption": boolean,
  "EnableNetworkIsolation": boolean,
  "VpcConfig": {
    "SecurityGroupIds": [ "string" ],
    "Subnets": [ "string" ]
  }
},
"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeInSeconds": number
}
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreationTime (p. 393)

The time at which the model explainability job was created.

Type: Timestamp

JobDefinitionArn (p. 393)

The Amazon Resource Name (ARN) of the model explainability job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

JobDefinitionName (p. 393)

The name of the explainability job definition. The name must be unique within an AWS Region in the AWS account.
Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*\$

**JobResources (p. 393)**

Identifies the resources to deploy for a monitoring job.

Type: MonitoringResources (p. 1210) object

**ModelExplainabilityAppSpecification (p. 393)**

Configures the model explainability job to run a specified Docker container image.

Type: ModelExplainabilityAppSpecification (p. 1159) object

**ModelExplainabilityBaselineConfig (p. 393)**

The baseline configuration for a model explainability job.

Type: ModelExplainabilityBaselineConfig (p. 1160) object

**ModelExplainabilityJobInput (p. 393)**

Inputs for the model explainability job.

Type: ModelExplainabilityJobInput (p. 1161) object

**ModelExplainabilityJobOutputConfig (p. 393)**

The output configuration for monitoring jobs.

Type: MonitoringOutputConfig (p. 1209) object

**NetworkConfig (p. 393)**

Networking options for a model explainability job.

Type: MonitoringNetworkConfig (p. 1207) object

**RoleArn (p. 393)**

The Amazon Resource Name (ARN) of the AWS Identity and Access Management (IAM) role that has read permission to the input data location and write permission to the output data location in Amazon S3.

Type: String


Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/\?[a-zA-Z\-Z_0-9+=,.@\-_\/]++$

**StoppingCondition (p. 393)**

A time limit for how long the monitoring job is allowed to run before stopping.

Type: MonitoringStoppingCondition (p. 1219) object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeModelPackage

Service: Amazon SageMaker Service

Returns a description of the specified model package, which is used to create SageMaker models or list them on AWS Marketplace.

To create models in SageMaker, buyers can subscribe to model packages listed on AWS Marketplace.

Request Syntax

```json
{
    "ModelPackageName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ModelPackageName (p. 397)**

The name or Amazon Resource Name (ARN) of the model package to describe.

When you specify a name, the name must have 1 to 63 characters. Valid characters are a-z, A-Z, 0-9, and - (hyphen).

Type: String


Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:[a-z\-]*\/)?(a-zA-Z0-9\[a-zA-Z0-9\-\]{0,62})(?<!-)(/\[0-9\]{1,5})?${

Required: Yes

Response Syntax

```json
{
    "AdditionalInferenceSpecifications": [
        {
            "Containers": [
                {
                    "ContainerHostname": "string",
                    "Environment": {
                        "string": "string"
                    },
                    "Framework": "string",
                    "FrameworkVersion": "string",
                    "Image": "string",
                    "ImageDigest": "string",
                    "ModelDataUrl": "string",
                    "ModelInput": {
                        "DataInputConfig": "string"
                    },
                    "NearestModelName": "string"
                }
            ]
        }
    ]
}
```
"ProductId": "string"
},
"Description": "string",
"Name": "string",
"SupportedContentTypes": [ "string" ],
"SupportedRealtimeInferenceInstanceTypes": [ "string" ],
"SupportedResponseMIMETypes": [ "string" ],
"SupportedTransformInstanceTypes": [ "string" ]
}
",
"ApprovalDescription": "string",
"CertifyForMarketplace": boolean,
"CreatedBy": {
  "DomainId": "string",
  "UserProfileArn": "string",
  "UserProfileName": "string"
},
"CreationTime": number,
"CustomerMetadataProperties": {
  "string": "string"
},
"Domain": "string",
"DriftCheckBaselines": {
  "Bias": {
    "ConfigFile": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "PostTrainingConstraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "PreTrainingConstraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    }
  },
  "Explainability": {
    "ConfigFile": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "Constraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    }
  },
  "ModelDataQuality": {
    "Constraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "Statistics": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    }
  }
},
"ModelQuality": {

"Constraints": {  
    "ContentDigest": "string",  
    "ContentType": "string",  
    "S3Uri": "string"
},  
"Statistics": {  
    "ContentDigest": "string",  
    "ContentType": "string",  
    "S3Uri": "string"
}
"InferenceSpecification": {  
    "Containers": [  
        {  
            "ContainerHostname": "string",  
            "Environment": {  
                "string": "string"
            },  
            "Framework": "string",  
            "FrameworkVersion": "string",  
            "Image": "string",  
            "ImageDigest": "string",  
            "ModelDataUrl": "string",  
            "ModelInput": {  
                "DataInputConfig": "string"
            },  
            "NearestModelName": "string",  
            "ProductId": "string"
        }
    ],  
    "SupportedContentTypes": [  
        "string"
    ],  
    "SupportedRealtimeInferenceInstanceTypes": [  
        "string"
    ],  
    "SupportedResponseMIMETypes": [  
        "string"
    ],  
    "SupportedTransformInstanceTypes": [  
        "string"
    ]
},  
"LastModifiedBy": {  
    "DomainId": "string",  
    "UserProfileArn": "string",  
    "UserProfileName": "string"
},  
"LastModifiedTime": number,  
"MetadataProperties": {  
    "CommitId": "string",  
    "GeneratedBy": "string",  
    "ProjectId": "string",  
    "Repository": "string"
},  
"ModelApprovalStatus": "string",  
"ModelMetrics": {  
    "Bias": {  
        "PostTrainingReport": {  
            "ContentDigest": "string",  
            "ContentType": "string",  
            "S3Uri": "string"
        },  
        "PreTrainingReport": {  
            "ContentDigest": "string",  
            "ContentType": "string",  
            "S3Uri": "string"
        },  
        "Report": {  
            "ContentDigest": "string",  
            "ContentType": "string",  
            "S3Uri": "string"
        }
    }
}
"Explainability": {
  "Report": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  }
},
"ModelDataQuality": {
  "Constraints": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  },
  "Statistics": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  }
},
"ModelQuality": {
  "Constraints": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  },
  "Statistics": {
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  }
},
"ModelPackageArn": "string",
"ModelPackageDescription": "string",
"ModelPackageGroupName": "string",
"ModelPackageName": "string",
"ModelPackageStatus": "string",
"ModelPackageStatusDetails": {
  "ImageScanStatuses": [
    {
      "FailureReason": "string",
      "Name": "string",
      "Status": "string"
    }
  ],
  "ValidationStatuses": [
    {
      "FailureReason": "string",
      "Name": "string",
      "Status": "string"
    }
  ]
},
"ModelPackageVersion": number,
"SamplePayloadUrl": "string",
"SourceAlgorithmSpecification": {
  "SourceAlgorithms": [
    {
      "AlgorithmName": "string",
      "ModelDataUrl": "string"
    }
  ]
},
"Task": "string",
"ValidationSpecification": {


"ValidationProfiles": [
    {
        "ProfileName": "string",
        "TransformJobDefinition": {
            "BatchStrategy": "string",
            "Environment": {
                "string": "string"
            },
            "MaxConcurrentTransforms": number,
            "MaxPayloadInMB": number,
            "TransformInput": {
                "CompressionType": "string",
                "ContentType": "string",
                "DataSource": {
                    "S3DataSource": {
                        "S3DataType": "string",
                        "S3Uri": "string"
                    }
                },
                "SplitType": "string"
            },
            "TransformOutput": {
                "Accept": "string",
                "AssembleWith": "string",
                "KmsKeyId": "string",
                "S3OutputPath": "string"
            },
            "TransformResources": {
                "InstanceCount": number,
                "InstanceType": "string",
                "VolumeKmsKeyId": "string"
            }
        }
    },
    {
        "ProfileName": "string",
        "TransformJobDefinition": {
            "BatchStrategy": "string",
            "Environment": {
                "string": "string"
            },
            "MaxConcurrentTransforms": number,
            "MaxPayloadInMB": number,
            "TransformInput": {
                "CompressionType": "string",
                "ContentType": "string",
                "DataSource": {
                    "S3DataSource": {
                        "S3DataType": "string",
                        "S3Uri": "string"
                    }
                },
                "SplitType": "string"
            },
            "TransformOutput": {
                "Accept": "string",
                "AssembleWith": "string",
                "KmsKeyId": "string",
                "S3OutputPath": "string"
            },
            "TransformResources": {
                "InstanceCount": number,
                "InstanceType": "string",
                "VolumeKmsKeyId": "string"
            }
        }
    }
],
"ValidationRole": "string"
]

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AdditionalInferenceSpecifications (p. 397)

An array of additional Inference Specification objects. Each additional Inference Specification specifies artifacts based on this model package that can be used on inference endpoints. Generally used with SageMaker Neo to store the compiled artifacts.

Type: Array of AdditionalInferenceSpecificationDefinition (p. 863) objects

Array Members: Minimum number of 1 item. Maximum number of 15 items.

ApprovalDescription (p. 397)

A description provided for the model approval.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: .*
CertifyForMarketplace (p. 397)

Whether the model package is certified for listing on AWS Marketplace.

Type: Boolean

CreatedBy (p. 397)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

CreationTime (p. 397)

A timestamp specifying when the model package was created.

Type: Timestamp

CustomerMetadataProperties (p. 397)

The metadata properties associated with the model package versions.

Type: String to string map

Map Entries: Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: ^([\p{L}\p{Z}\p{N}_.:/=+-@]*):1,128

Value Length Constraints: Minimum length of 1. Maximum length of 256.

Value Pattern: ^([\p{L}\p{Z}\p{N}_.:/=+-@]*):1,256

Domain (p. 397)

The machine learning domain of the model package you specified. Common machine learning domains include computer vision and natural language processing.

Type: String

DriftCheckBaselines (p. 397)

Represents the drift check baselines that can be used when the model monitor is set using the model package. For more information, see the topic on Drift Detection against Previous Baselines in SageMaker Pipelines in the Amazon SageMaker Developer Guide.

Type: DriftCheckBaselines (p. 998) object

InferenceSpecification (p. 397)

Details about inference jobs that can be run with models based on this model package.

Type: InferenceSpecification (p. 1105) object

LastModifiedBy (p. 397)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

LastModifiedTime (p. 397)

The last time that the model package was modified.
Type: Timestamp

**MetadataProperties (p. 397)**

Metadata properties of the tracking entity, trial, or trial component.

Type: MetadataProperties (p. 1143) object

**ModelApprovalStatus (p. 397)**

The approval status of the model package.

Type: String

Valid Values: Approved | Rejected | PendingManualApproval

**ModelMetrics (p. 397)**

Metrics for the model.

Type: ModelMetrics (p. 1167) object

**ModelPackageArn (p. 397)**

The Amazon Resource Name (ARN) of the model package.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:model-package/.*`

**ModelPackageDescription (p. 397)**

A brief summary of the model package.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `[^\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*`

**ModelPackageGroupName (p. 397)**

If the model is a versioned model, the name of the model group that the versioned model belongs to.

Type: String


Pattern: `^[a-zA-Z0-9-]{0,62}`

**ModelPackageName (p. 397)**

The name of the model package being described.

Type: String


Pattern: `^[a-zA-Z0-9-]{0,62}`

**ModelPackageStatus (p. 397)**

The current status of the model package.

Type: String
Valid Values: Pending | InProgress | Completed | Failed | Deleting

ModelPackageStatusDetails (p. 397)
Details about the current status of the model package.
Type: ModelPackageStatusDetails (p. 1180) object

ModelPackageVersion (p. 397)
The version of the model package.
Type: Integer
Valid Range: Minimum value of 1.

SamplePayloadUrl (p. 397)
The Amazon Simple Storage Service (Amazon S3) path where the sample payload are stored. This path points to a single gzip compressed tar archive (.tar.gz suffix).
Type: String

SourceAlgorithmSpecification (p. 397)
Details about the algorithm that was used to create the model package.
Type: SourceAlgorithmSpecification (p. 1368) object

Task (p. 397)
The machine learning task you specified that your model package accomplishes. Common machine learning tasks include object detection and image classification.
Type: String

ValidationSpecification (p. 397)
Configurations for one or more transform jobs that SageMaker runs to test the model package.
Type: ModelPackageValidationSpecification (p. 1185) object

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeModelPackageGroup

Service: Amazon SageMaker Service

Gets a description for the specified model group.

Request Syntax

```
{
   "ModelPackageGroupName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ModelPackageGroupName (p. 405)

The name of the model group to describe.

Type: String


Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:[a-z\-]*\/)?([a-zA-Z0-9](\[a-zA-Z0-9]{0,62})(?!-))?

Required: Yes

Response Syntax

```
{
   "CreatedBy": {
      "DomainId": "string",
      "UserProfileArn": "string",
      "UserProfileName": "string"
   },
   "CreationTime": number,
   "ModelPackageGroupArn": "string",
   "ModelPackageGroupDescription": "string",
   "ModelPackageGroupName": "string",
   "ModelPackageGroupStatus": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreatedBy (p. 405)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
Type: UserContext (p. 1449) object

CreationTime (p. 405)

The time that the model group was created.

Type: Timestamp

ModelPackageGroupArn (p. 405)

The Amazon Resource Name (ARN) of the model group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:model-package-group/.*

ModelPackageGroupDescription (p. 405)

A description of the model group.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*

ModelPackageGroupName (p. 405)

The name of the model group.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\$

ModelPackageGroupStatus (p. 405)

The status of the model group.

Type: String

Valid Values: Pending | InProgress | Completed | Failed | Deleting | DeleteFailed

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DescribeModelQualityJobDefinition
Service: Amazon SageMaker Service

Returns a description of a model quality job definition.

Request Syntax

```json
{
    "JobDefinitionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**JobDefinitionName (p. 408)**

The name of the model quality job. The name must be unique within an AWS Region in the AWS account.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*$`

Required: Yes

Response Syntax

```json
{
    "CreationTime": number,
    "JobDefinitionArn": "string",
    "JobDefinitionName": "string",
    "JobResources": {
        "ClusterConfig": {
            "InstanceCount": number,
            "InstanceType": "string",
            "VolumeKmsKeyId": "string",
            "VolumeSizeInGB": number
        }
    },
    "ModelQualityAppSpecification": {
        "ContainerArguments": ["string"],
        "ContainerEntrypoint": ["string"],
        "Environment": {
            "string": "string"
        },
        "ImageUri": "string",
        "PostAnalyticsProcessorSourceUri": "string",
        "ProblemType": "string",
        "RecordPreprocessorSourceUri": "string"
    },
    "ModelQualityBaselineConfig": {
        "BaselineJobName": "string",
```
"ConstraintsResource": {"S3Uri": "string"}
},
"ModelQualityJobInput": {
  "EndpointInput": {
    "EndpointName": "string",
    "EndTimeOffset": "string",
    "FeaturesAttribute": "string",
    "InferenceAttribute": "string",
    "LocalPath": "string",
    "ProbabilityAttribute": "string",
    "ProbabilityThresholdAttribute": number,
    "S3DataDistributionType": "string",
    "S3InputMode": "string",
    "StartTimeOffset": "string"
  },
  "GroundTruthS3Input": {
    "S3Uri": "string"
  }
},
"ModelQualityJobOutputConfig": {
  "KmsKeyId": "string",
  "MonitoringOutputs": [
    {
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
},
"NetworkConfig": {
  "EnableInterContainerTrafficEncryption": boolean,
  "EnableNetworkIsolation": boolean,
  "VpcConfig": {
    "SecurityGroupIds": [ "string" ],
    "Subnets": [ "string" ]
  }
},
"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeInSeconds": number
}
}

## Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 408)**

The time at which the model quality job was created.

Type: Timestamp

**JobDefinitionArn (p. 408)**

The Amazon Resource Name (ARN) of the model quality job.

Type: String
Length Constraints: Maximum length of 256.

Pattern: .*

JobDefinitionName (p. 408)

The name of the quality job definition. The name must be unique within an AWS Region in the AWS account.

Type: String


Pattern: ^[a-zA-Z0-9][-\][a-zA-Z0-9]+$

JobResources (p. 408)

Identifies the resources to deploy for a monitoring job.

Type: MonitoringResources (p. 1210) object

ModelQualityAppSpecification (p. 408)

Configures the model quality job to run a specified Docker container image.

Type: ModelQualityAppSpecification (p. 1187) object

ModelQualityBaselineConfig (p. 408)

The baseline configuration for a model quality job.

Type: ModelQualityBaselineConfig (p. 1189) object

ModelQualityJobInput (p. 408)

Inputs for the model quality job.

Type: ModelQualityJobInput (p. 1190) object

ModelQualityJobOutputConfig (p. 408)

The output configuration for monitoring jobs.

Type: MonitoringOutputConfig (p. 1209) object

NetworkConfig (p. 408)

Networking options for a model quality job.

Type: MonitoringNetworkConfig (p. 1207) object

RoleArn (p. 408)

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-zA-Z-]+:iam::<d(12)>:role/\+[a-zA-Z0-9+=,.@-_/]+#$

StoppingCondition (p. 408)

A time limit for how long the monitoring job is allowed to run before stopping.

Type: MonitoringStoppingCondition (p. 1219) object
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeMonitoringSchedule
Service: Amazon SageMaker Service

Describes the schedule for a monitoring job.

Request Syntax

```json
{
  "MonitoringScheduleName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**MonitoringScheduleName (p. 412)**

Name of a previously created monitoring schedule.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Syntax

```json
{
  "CreationTime": number,
  "EndpointName": "string",
  "FailureReason": "string",
  "LastModifiedTime": number,
  "LastMonitoringExecutionSummary": {
    "CreationTime": number,
    "EndpointName": "string",
    "FailureReason": "string",
    "LastModifiedTime": number,
    "MonitoringExecutionStatus": "string",
    "MonitoringJobDefinitionName": "string",
    "MonitoringScheduleName": "string",
    "MonitoringType": "string",
    "ProcessingJobArn": "string",
    "ScheduledTime": number
  },
  "MonitoringScheduleArn": "string",
  "MonitoringScheduleConfig": {
    "MonitoringJobDefinition": {
      "BaselineConfig": {
        "BaseliningJobName": "string",
        "ConstraintsResource": {
          "S3Uri": "string"
        }
      },
      "StatisticsResource": {
```
"S3Uri": "string"
"Environment": {
  "string": "string"
"MonitoringAppSpecification": {
  "ContainerArguments": [ "string" ],
  "ContainerEntrypoint": [ "string" ],
  "ImageUri": "string",
  "PostAnalyticsProcessorSourceUri": "string",
  "RecordPreprocessorSourceUri": "string"
"MonitoringInputs": [
  {
    "EndpointInput": {
      "EndpointName": "string",
      "EndTimeOffset": "string",
      "FeaturesAttribute": "string",
      "InferenceAttribute": "string",
      "LocalPath": "string",
      "ProbabilityAttribute": "string",
      "ProbabilityThresholdAttribute": number,
      "S3DataDistributionType": "string",
      "S3InputMode": "string",
      "StartTimeOffset": "string"
    }
  }
"MonitoringOutputConfig": {
  "KmsKeyId": "string",
  "MonitoringOutputs": [
    {
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
"MonitoringResources": {
  "ClusterConfig": {
    "InstanceCount": number,
    "InstanceType": "string",
    "VolumeKmsKeyId": "string",
    "VolumeSizeInGB": number
  }
"NetworkConfig": {
  "EnableInterContainerTrafficEncryption": boolean,
  "EnableNetworkIsolation": boolean,
  "VpcConfig": {
    "SecurityGroupIds": [ "string" ],
    "Subnets": [ "string" ]
  }
"RoleArn": "string",
  "StoppingCondition": {
    "MaxRuntimeInSeconds": number
  }
"MonitoringJobDefinitionName": "string",
"MonitoringType": "string",
"ScheduleConfig": {
  "ScheduleExpression": "string"
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 412)**

The time at which the monitoring job was created.

Type: Timestamp

**EndpointName (p. 412)**

The name of the endpoint for the monitoring job.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$

**FailureReason (p. 412)**

A string, up to one KB in size, that contains the reason a monitoring job failed, if it failed.

Type: String

Length Constraints: Maximum length of 1024.

**LastModifiedTime (p. 412)**

The time at which the monitoring job was last modified.

Type: Timestamp

**LastMonitoringExecutionSummary (p. 412)**

Describes metadata on the last execution to run, if there was one.

Type: MonitoringExecutionSummary (p. 1199) object

**MonitoringScheduleArn (p. 412)**

The Amazon Resource Name (ARN) of the monitoring schedule.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

**MonitoringScheduleConfig (p. 412)**

The configuration object that specifies the monitoring schedule and defines the monitoring job.

Type: MonitoringScheduleConfig (p. 1215) object
MonitoringscheduleName (p. 412)

Name of the monitoring schedule.

Type: String


Pattern: ^[a-zA-Z0-9](\*\-[a-zA-Z0-9])\{0,62}$

Monitoringschedulestatus (p. 412)

The status of an monitoring job.

Type: String

Valid Values: Pending | Failed | Scheduled | Stopped

Monitorings typ e (p. 412)

The type of the monitoring job that this schedule runs. This is one of the following values.

- DATA_QUALITY - The schedule is for a data quality monitoring job.
- MODEL_QUALITY - The schedule is for a model quality monitoring job.
- MODEL_BIAS - The schedule is for a bias monitoring job.
- MODEL_EXPLAINABILITY - The schedule is for an explainability monitoring job.

Type: String

Valid Values: DataQuality | ModelQuality | ModelBias | ModelExplainability

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeNotebookInstance

Returns information about a notebook instance.

Request Syntax

```json
{
  "NotebookInstanceName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**NotebookInstanceName** (p. 416)

The name of the notebook instance that you want information about.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*  

Required: Yes

Response Syntax

```json
{
  "AcceleratorTypes": [ "string" ],
  "AdditionalCodeRepositories": [ "string" ],
  "CreationTime": number,
  "DefaultCodeRepository": "string",
  "DirectInternetAccess": "string",
  "FailureReason": "string",
  "InstanceMetadataServiceConfiguration": {
    "MinimumInstanceMetadataServiceVersion": "string"
  },
  "InstanceType": "string",
  "KmsKeyId": "string",
  "LastModifiedTime": number,
  "NetworkInterfaceId": "string",
  "NotebookInstanceArn": "string",
  "NotebookInstanceLifecycleConfigName": "string",
  "NotebookInstanceName": "string",
  "NotebookInstanceStatus": "string",
  "PlatformIdentifier": "string",
  "RoleArn": "string",
  "RootAccess": "string",
  "SecurityGroups": [ "string" ],
  "SubnetId": "string",
  "Url": "string",
  "VolumeSizeInGB": number
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AcceleratorTypes (p. 416)**

A list of the Elastic Inference (EI) instance types associated with this notebook instance. Currently only one EI instance type can be associated with a notebook instance. For more information, see Using Elastic Inference in Amazon SageMaker.

Type: Array of strings

Valid Values: ml.eia1.medium | ml.eia1.large | ml.eia1.xlarge | ml.eia2.medium | ml.eia2.large | ml.eia2.xlarge

**AdditionalCodeRepositories (p. 416)**

An array of up to three Git repositories associated with the notebook instance. These can be either the names of Git repositories stored as resources in your account, or the URL of Git repositories in AWS CodeCommit or in any other Git repository. These repositories are cloned at the same level as the default repository of your notebook instance. For more information, see Associating Git Repositories with SageMaker Notebook Instances.

Type: Array of strings

Array Members: Maximum number of 3 items.


Pattern: ^https://([^/]+)/?(.*)$|^\[a-zA-Z0-9\](-*[a-zA-Z0-9]*)*

**CreationTime (p. 416)**

A timestamp. Use this parameter to return the time when the notebook instance was created

Type: Timestamp

**DefaultCodeRepository (p. 416)**

The Git repository associated with the notebook instance as its default code repository. This can be either the name of a Git repository stored as a resource in your account, or the URL of a Git repository in AWS CodeCommit or in any other Git repository. When you open a notebook instance, it opens in the directory that contains this repository. For more information, see Associating Git Repositories with SageMaker Notebook Instances.

Type: String


Pattern: ^https://([^/]+)/?(.*)$|^\[a-zA-Z0-9\](-*[a-zA-Z0-9]*)*

**DirectInternetAccess (p. 416)**

Describes whether SageMaker provides internet access to the notebook instance. If this value is set to Disabled, the notebook instance does not have internet access, and cannot connect to SageMaker training and endpoint services.

For more information, see Notebook Instances Are Internet-Enabled by Default.

Type: String

Valid Values: Enabled | Disabled
FailureReason (p. 416)

If status is Failed, the reason it failed.

Type: String

Length Constraints: Maximum length of 1024.

InstanceMetadataServiceConfiguration (p. 416)

Information on the IMDS configuration of the notebook instance

Type: InstanceMetadataServiceConfiguration (p. 1111) object

InstanceType (p. 416)

The type of ML compute instance running on the notebook instance.

Type: String

Valid Values: ml.t2.medium | ml.t2.large | ml.t2.xlarge | ml.t2.2xlarge | ml.t3.medium | ml.t3.large | ml.t3.xlarge | ml.t3.2xlarge | ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge | ml.m4.16xlarge | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge | ml.m5.24xlarge | ml.m5d.xlarge | ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.8xlarge | ml.m5d.12xlarge | ml.m5d.16xlarge | ml.m5d.24xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge | ml.c4.8xlarge | ml.c5.xlarge | ml.c5.2xlarge | ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge | ml.c5d.xlarge | ml.c5d.2xlarge | ml.c5d.4xlarge | ml.c5d.9xlarge | ml.c5d.18xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge | ml.p3.xlarge | ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.p3dn.24xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.r5.large | ml.r5.xlarge | ml.r5.2xlarge | ml.r5.4xlarge | ml.r5.8xlarge | ml.r5.12xlarge | ml.r5.24xlarge | ml.r5.48xlarge | ml.g5.xlarge | ml.g5.2xlarge | ml.g5.4xlarge | ml.g5.8xlarge | ml.g5.16xlarge | ml.g5.12xlarge | ml.g5.24xlarge | ml.g5.48xlarge

KmsKeyId (p. 416)

The AWS KMS key ID SageMaker uses to encrypt data when storing it on the ML storage volume attached to the instance.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

LastModifiedTime (p. 416)

A timestamp. Use this parameter to retrieve the time when the notebook instance was last modified.

Type: Timestamp

NetworkInterfaceId (p. 416)

The network interface IDs that SageMaker created at the time of creating the instance.

Type: String

NotebookInstanceArn (p. 416)

The Amazon Resource Name (ARN) of the notebook instance.
DescribeNotebookInstance

Type: String

Length Constraints: Maximum length of 256.

**NotebookInstanceLifecycleConfigName (p. 416)**

Returns the name of a notebook instance lifecycle configuration.

For information about notebook instance lifestyle configurations, see Step 2.1: (Optional) Customize a Notebook Instance

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*`

**NotebookInstanceName (p. 416)**

The name of the SageMaker notebook instance.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*`

**NotebookInstanceStatus (p. 416)**

The status of the notebook instance.

Type: String

Valid Values: `Pending` | `InService` | `Stopping` | `Stopped` | `Failed` | `Deleting` | `Updating`

**PlatformIdentifier (p. 416)**

The platform identifier of the notebook instance runtime environment.

Type: String

Length Constraints: Maximum length of 15.

Pattern: `^(notebook-al1-v1|notebook-al2-v1|notebook-al2-v2)$`

**RoleArn (p. 416)**

The Amazon Resource Name (ARN) of the IAM role associated with the instance.

Type: String


Pattern: `^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9+=,.@\-_]+$/`

**RootAccess (p. 416)**

Whether root access is enabled or disabled for users of the notebook instance.

*Note*

Lifecycle configurations need root access to be able to set up a notebook instance. Because of this, lifecycle configurations associated with a notebook instance always run with root access even if you disable root access for users.

Type: String
Valid Values: Enabled | Disabled

**SecurityGroups (p. 416)**

The IDs of the VPC security groups.

Type: Array of strings

Array Members: Maximum number of 5 items.

Length Constraints: Maximum length of 32.

Pattern: [-0-9a-zA-Z]+

**SubnetId (p. 416)**

The ID of the VPC subnet.

Type: String

Length Constraints: Maximum length of 32.

Pattern: [-0-9a-zA-Z]+

**Url (p. 416)**

The URL that you use to connect to the Jupyter notebook that is running in your notebook instance.

Type: String

**VolumeSizeInGB (p. 416)**

The size, in GB, of the ML storage volume attached to the notebook instance.

Type: Integer


**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeNotebookInstanceLifecycleConfig

Service: Amazon SageMaker Service

Returns a description of a notebook instance lifecycle configuration.

For information about notebook instance lifestyle configurations, see Step 2.1: (Optional) Customize a Notebook Instance.

Request Syntax

```json
{
  "NotebookInstanceLifecycleConfigName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**NotebookInstanceLifecycleConfigName (p. 421)**

The name of the lifecycle configuration to describe.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*?

Required: Yes

Response Syntax

```json
{
  "CreationTime": number,
  "LastModifiedTime": number,
  "NotebookInstanceLifecycleConfigArn": "string",
  "NotebookInstanceLifecycleConfigName": "string",
  "OnCreate": [
    {
      "Content": "string"
    }
  ],
  "OnStart": [
    {
      "Content": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
CreationTime (p. 421)
A timestamp that tells when the lifecycle configuration was created.
Type: Timestamp

LastModifiedTime (p. 421)
A timestamp that tells when the lifecycle configuration was last modified.
Type: Timestamp

NotebookInstanceLifecycleConfigArn (p. 421)
The Amazon Resource Name (ARN) of the lifecycle configuration.
Type: String
Length Constraints: Maximum length of 256.

NotebookInstanceLifecycleConfigName (p. 421)
The name of the lifecycle configuration.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]*)*

OnCreate (p. 421)
The shell script that runs only once, when you create a notebook instance.
Type: Array of NotebookInstanceLifecycleHook (p. 1225) objects
Array Members: Maximum number of 1 item.

OnStart (p. 421)
The shell script that runs every time you start a notebook instance, including when you create the notebook instance.
Type: Array of NotebookInstanceLifecycleHook (p. 1225) objects
Array Members: Maximum number of 1 item.

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DescribePipeline
Service: Amazon SageMaker Service

Describes the details of a pipeline.

Request Syntax

```
{
  "PipelineName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**PipelineName (p. 424)**

The name of the pipeline to describe.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,255}$

Required: Yes

Response Syntax

```
{
  "CreatedBy": {
    "DomainId": "string",
    "UserProfileArn": "string",
    "UserProfileName": "string"
  },
  "CreationTime": number,
  "LastModifiedBy": {
    "DomainId": "string",
    "UserProfileArn": "string",
    "UserProfileName": "string"
  },
  "LastModifiedTime": number,
  "LastRunTime": number,
  "ParallelismConfiguration": {
    "MaxParallelExecutionSteps": number
  },
  "PipelineArn": "string",
  "PipelineDefinition": "string",
  "PipelineDescription": "string",
  "PipelineDisplayName": "string",
  "PipelineName": "string",
  "PipelineStatus": "string",
  "RoleArn": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreatedBy (p. 424)**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

**CreationTime (p. 424)**

The time when the pipeline was created.

Type: Timestamp

**LastModifiedBy (p. 424)**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

**LastModifiedTime (p. 424)**

The time when the pipeline was last modified.

Type: Timestamp

**LastRunTime (p. 424)**

The time when the pipeline was last run.

Type: Timestamp

**ParallelismConfiguration (p. 424)**

Lists the parallelism configuration applied to the pipeline.

Type: ParallelismConfiguration (p. 1247) object

**PipelineArn (p. 424)**

The Amazon Resource Name (ARN) of the pipeline.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*`

**PipelineDefinition (p. 424)**

The JSON pipeline definition.

Type: String


Pattern: `.**(?:[\r\n\t].*)*`

**PipelineDescription (p. 424)**

The description of the pipeline.
DescribePipeline

Type: String
Length Constraints: Minimum length of 0. Maximum length of 3072.
Pattern: .*

**PipelineDisplayName (p. 424)**
The display name of the pipeline.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,255}$

**PipelineName (p. 424)**
The name of the pipeline.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,255}$

**PipelineStatus (p. 424)**
The status of the pipeline execution.
Type: String
Valid Values: Active

**RoleArn (p. 424)**
The Amazon Resource Name (ARN) that the pipeline uses to execute.
Type: String
Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/?[a-zA-Z0-9=,.@\-_\//]+#$

**Errors**
For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**
Resource being access is not found.
HTTP Status Code: 400

**See Also**
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DescribePipelineDefinitionForExecution

Service: Amazon SageMaker Service

Describes the details of an execution's pipeline definition.

Request Syntax

```
{
  "PipelineExecutionArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**PipelineExecutionArn (p. 428)**

The Amazon Resource Name (ARN) of the pipeline execution.

- Type: String
- Length Constraints: Maximum length of 256.
- Pattern: ^arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:pipeline/.*\ execution/.*$
- Required: Yes

Response Syntax

```
{
  "CreationTime": number,
  "PipelineDefinition": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreationTime (p. 428)**

The time when the pipeline was created.

- Type: Timestamp

**PipelineDefinition (p. 428)**

The JSON pipeline definition.

- Type: String
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being accessed is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribePipelineExecution
Service: Amazon SageMaker Service
Describes the details of a pipeline execution.

Request Syntax

```
{
    "PipelineExecutionArn": "string"
}
```

Request Parameters
For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

PipelineExecutionArn (p. 430)
The Amazon Resource Name (ARN) of the pipeline execution.
Type: String
Length Constraints: Maximum length of 256.
Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline\./.*\//execution\./.*$
Required: Yes

Response Syntax

```
{
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "FailureReason": "string",
    "LastModifiedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "LastModifiedTime": number,
    "ParallelismConfiguration": {
        "MaxParallelExecutionSteps": number
    },
    "PipelineArn": "string",
    "PipelineExecutionArn": "string",
    "PipelineExecutionDescription": "string",
    "PipelineExecutionDisplayName": "string",
    "PipelineExecutionStatus": "string",
    "PipelineExperimentConfig": {
        "ExperimentName": "string",
        "TrialName": "string"
    }
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**CreatedBy (p. 430)**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

**CreationTime (p. 430)**

The time when the pipeline execution was created.

Type: Timestamp

**FailureReason (p. 430)**

If the execution failed, a message describing why.

Type: String

Length Constraints: Maximum length of 1300.

Pattern: . *

**LastModifiedBy (p. 430)**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

**LastModifiedTime (p. 430)**

The time when the pipeline execution was modified last.

Type: Timestamp

**ParallelismConfiguration (p. 430)**

The parallelism configuration applied to the pipeline.

Type: ParallelismConfiguration (p. 1247) object

**PipelineArn (p. 430)**

The Amazon Resource Name (ARN) of the pipeline.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]\*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*

**PipelineExecutionArn (p. 430)**

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String
Length Constraints: Maximum length of 256.
Pattern: ^arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:pipeline/.*\ crimection/.*$

**PipelineExecutionDescription (p. 430)**

The description of the pipeline execution.
Type: String
Length Constraints: Minimum length of 0. Maximum length of 3072.
Pattern: .*

**PipelineExecutionDisplayName (p. 430)**

The display name of the pipeline execution.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 82.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,81}$

**PipelineExecutionStatus (p. 430)**

The status of the pipeline execution.
Type: String
Valid Values: Executing | Stopping | Stopped | Failed | Succeeded

**PipelineExperimentConfig (p. 430)**

Specifies the names of the experiment and trial created by a pipeline.
Type: PipelineExperimentConfig (p. 1272) object

### Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFoundException**

Resource being access is not found.

HTTP Status Code: 400

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeProcessingJob
Service: Amazon SageMaker Service
Returns a description of a processing job.

Request Syntax

```json
{
    "ProcessingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ProcessingJobName (p. 434)**

The name of the processing job. The name must be unique within an AWS Region in the AWS account.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Syntax

```json
{
    "AppSpecification": {
        "ContainerArguments": [ "string" ],
        "ContainerEntrypoint": [ "string" ],
        "ImageUri": "string"
    },
    "AutoMLJobArn": "string",
    "CreationTime": number,
    "Environment": {
        "string": "string"
    },
    "ExitMessage": "string",
    "ExperimentConfig": {
        "ExperimentName": "string",
        "TrialComponentDisplayName": "string",
        "TrialName": "string"
    },
    "FailureReason": "string",
    "LastModifiedTime": number,
    "MonitoringScheduleArn": "string",
    "NetworkConfig": {
        "EnableInterContainerTrafficEncryption": boolean,
        "EnableNetworkIsolation": boolean,
        "VpcConfig": {
            "SecurityGroupId": [ "string" ],
            "Subnets": [ "string" ]
        }
    }
}
```
"ProcessingEndTime": number,
"ProcessingInputs": [
  {
    "AppManaged": boolean,
    "DatasetDefinition": {
      "AthenaDatasetDefinition": {
        "Catalog": "string",
        "Database": "string",
        "KmsKeyId": "string",
        "OutputCompression": "string",
        "OutputFormat": "string",
        "OutputS3Uri": "string",
        "QueryString": "string",
        "WorkGroup": "string"
      },
      "DataDistributionType": "string",
      "InputMode": "string",
      "LocalPath": "string",
      "RedshiftDatasetDefinition": {
        "ClusterId": "string",
        "ClusterRoleArn": "string",
        "Database": "string",
        "DbUser": "string",
        "KmsKeyId": "string",
        "OutputCompression": "string",
        "OutputFormat": "string",
        "OutputS3Uri": "string",
        "QueryString": "string"
      }
    },
    "InputName": "string",
    "S3Input": {
      "LocalPath": "string",
      "S3CompressionType": "string",
      "S3DataDistributionType": "string",
      "S3DataType": "string",
      "S3InputMode": "string",
      "S3Uri": "string"
    }
  }
],
"ProcessingJobArn": "string",
"ProcessingJobName": "string",
"ProcessingJobStatus": "string",
"ProcessingOutputConfig": {
  "KmsKeyId": "string",
  "Outputs": [
    {
      "AppManaged": boolean,
      "FeatureStoreOutput": {
        "FeatureGroupName": "string"
      },
      "OutputName": "string",
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
},
"ProcessingResources": {
  "ClusterConfig": {
    "InstanceCount": number,
    "InstanceType": "string"
  }
}
Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**AppSpecification (p. 434)**

Configures the processing job to run a specified container image.

Type: `AppSpecification (p. 891)` object

**AutoMLJobArn (p. 434)**

The ARN of an AutoML job associated with this processing job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*`

**CreationTime (p. 434)**

The time at which the processing job was created.

Type: Timestamp

**Environment (p. 434)**

The environment variables set in the Docker container.

Type: String to string map

Map Entries: Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: `[a-zA-Z_][a-zA-Z0-9_]`

Value Length Constraints: Maximum length of 256.

Value Pattern: `[\S\s]*`

**ExitMessage (p. 434)**

An optional string, up to one KB in size, that contains metadata from the processing container when the processing job exits.

Type: String

Length Constraints: Maximum length of 1024.
ExperimentConfig (p. 434)
The configuration information used to create an experiment.
Type: ExperimentConfig (p. 1029) object

FailureReason (p. 434)
A string, up to one KB in size, that contains the reason a processing job failed, if it failed.
Type: String
Length Constraints: Maximum length of 1024.

LastModifiedTime (p. 434)
The time at which the processing job was last modified.
Type: Timestamp

MonitoringScheduleArn (p. 434)
The ARN of a monitoring schedule for an endpoint associated with this processing job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: .*

NetworkConfig (p. 434)
Networking options for a processing job.
Type: NetworkConfig (p. 1223) object

ProcessingEndTime (p. 434)
The time at which the processing job completed.
Type: Timestamp

ProcessingInputs (p. 434)
The inputs for a processing job.
Type: Array of ProcessingInput (p. 1278) objects
Array Members: Minimum number of 0 items. Maximum number of 10 items.

ProcessingJobArn (p. 434)
The Amazon Resource Name (ARN) of the processing job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:processing-job/.*

ProcessingJobName (p. 434)
The name of the processing job. The name must be unique within an AWS Region in the AWS account.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

**ProcessingJobStatus (p. 434)**

Provides the status of a processing job.
Type: String
Valid Values: InProgress | Completed | Failed | Stopping | Stopped

**ProcessingOutputConfig (p. 434)**

Output configuration for the processing job.
Type: ProcessingOutputConfig (p. 1287) object

**ProcessingResources (p. 434)**

Identifies the resources, ML compute instances, and ML storage volumes to deploy for a processing job. In distributed training, you specify more than one instance.
Type: ProcessingResources (p. 1288) object

**ProcessingStartTime (p. 434)**

The time at which the processing job started.
Type: Timestamp

**RoleArn (p. 434)**

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.
Type: String
Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role//?([a-zA-Z0-9_]+,\@\-//]+\$

**StoppingCondition (p. 434)**

The time limit for how long the processing job is allowed to run.
Type: ProcessingStoppingCondition (p. 1292) object

**TrainingJobArn (p. 434)**

The ARN of a training job associated with this processing job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:training-job/.*

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeProject
Service: Amazon SageMaker Service

Describes the details of a project.

Request Syntax

```json
{
  "ProjectName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ProjectName (p. 440)**

The name of the project to describe.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}$

Required: Yes

Response Syntax

```json
{
  "CreatedBy": {
    "DomainId": "string",
    "UserProfileArn": "string",
    "UserProfileName": "string"
  },
  "CreationTime": number,
  "LastModifiedBy": {
    "DomainId": "string",
    "UserProfileArn": "string",
    "UserProfileName": "string"
  },
  "LastModifiedTime": number,
  "ProjectArn": "string",
  "ProjectDescription": "string",
  "ProjectId": "string",
  "ProjectName": "string",
  "ProjectStatus": "string",
  "ServiceCatalogProvisionedProductDetails": {
    "ProvisionedProductId": "string",
    "ProvisionedProductStatusMessage": "string"
  },
  "ServiceCatalogProvisioningDetails": {
    "PathId": "string",
    "ProductId": "string",
    "ProvisioningArtifactId": "string"
  }
}
```
"ProvisioningParameters": [
  {
    "Key": "string",
    "Value": "string"
  }
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreatedBy (p. 440)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

CreationTime (p. 440)

The time when the project was created.

Type: Timestamp

LastModifiedBy (p. 440)

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

LastModifiedTime (p. 440)

The timestamp when project was last modified.

Type: Timestamp

ProjectArn (p. 440)

The Amazon Resource Name (ARN) of the project.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:project:.*

ProjectDescription (p. 440)

The description of the project.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*

ProjectId (p. 440)

The ID of the project.

Type: String
DescribeProject

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*  

**ProjectName (p. 440)**

The name of the project.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}

**ProjectStatus (p. 440)**

The status of the project.
Type: String
Valid Values: Pending | CreateInProgress | CreateCompleted | CreateFailed | DeleteInProgress | DeleteFailed | DeleteCompleted | UpdateInProgress | UpdateCompleted | UpdateFailed

**ServiceCatalogProvisionedProductDetails (p. 440)**

Information about a provisioned service catalog product.
Type: ServiceCatalogProvisionedProductDetails (p. 1361) object

**ServiceCatalogProvisioningDetails (p. 440)**

Information used to provision a service catalog product. For information, see What is AWS Service Catalog.
Type: ServiceCatalogProvisioningDetails (p. 1362) object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeStudioLifecycleConfig
Service: Amazon SageMaker Service
Describes the Studio Lifecycle Configuration.

Request Syntax

```json
{
    "StudioLifecycleConfigName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

StudioLifecycleConfigName (p. 443)
The name of the Studio Lifecycle Configuration to describe.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}
Required: Yes

Response Syntax

```json
{
    "CreationTime": number,
    "LastModifiedTime": number,
    "StudioLifecycleConfigAppType": "string",
    "StudioLifecycleConfigArn": "string",
    "StudioLifecycleConfigContent": "string",
    "StudioLifecycleConfigName": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreationTime (p. 443)
The creation time of the Studio Lifecycle Configuration.
Type: Timestamp

LastModifiedTime (p. 443)
This value is equivalent to CreationTime because Studio Lifecycle Configurations are immutable.
Type: Timestamp

**StudioLifecycleConfigAppType (p. 443)**

The App type that the Lifecycle Configuration is attached to.

Type: String

Valid Values: JupyterServer | KernelGateway

**StudioLifecycleConfigArn (p. 443)**

The ARN of the Lifecycle Configuration to describe.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:studio-lifecycle-config/.*

**StudioLifecycleConfigContent (p. 443)**

The content of your Studio Lifecycle Configuration script.

Type: String


Pattern: [\S\s]+

**StudioLifecycleConfigName (p. 443)**

The name of the Studio Lifecycle Configuration that is described.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-]+[a-zA-Z0-9\-]{0,62}$

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
DescribeStudioLifecycleConfig

- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeSubscribedWorkteam

Service: Amazon SageMaker Service

Gets information about a work team provided by a vendor. It returns details about the subscription with a vendor in the AWS Marketplace.

Request Syntax

```json
{
   "WorkteamArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**WorkteamArn (p. 446)**

The Amazon Resource Name (ARN) of the subscribed work team to describe.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workteam/.*`

Required: Yes

Response Syntax

```json
{
   "SubscribedWorkteam": {
      "ListingId": "string",
      "MarketplaceDescription": "string",
      "MarketplaceTitle": "string",
      "SellerName": "string",
      "WorkteamArn": "string"
   }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**SubscribedWorkteam (p. 446)**

A *Workteam* instance that contains information about the work team.

Type: SubscribedWorkteam (p. 1374) object
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeTrainingJob
Service: Amazon SageMaker Service

Returns information about a training job.

Some of the attributes below only appear if the training job successfully starts. If the training job fails, TrainingJobStatus is Failed and, depending on the FailureReason, attributes like TrainingStartTime, TrainingTimeInSeconds, TrainingEndTime, and BillableTimeInSeconds may not be present in the response.

Request Syntax

```json
{
    "TrainingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

TrainingJobName (p. 448)

The name of the training job.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Syntax

```json
{
    "AlgorithmSpecification": {
        "AlgorithmName": "string",
        "EnableSageMakerMetricsTimeSeries": boolean,
        "MetricDefinitions": [
            {
                "Name": "string",
                "Regex": "string"
            }
        ],
        "TrainingImage": "string",
        "TrainingInputMode": "string"
    },
    "AutoMLJobArn": "string",
    "BillableTimeInSeconds": number,
    "CheckpointConfig": {
        "LocalPath": "string",
        "S3Uri": "string"
    },
    "CreationTime": number,
    "DebugHookConfig": {
```
"CollectionConfigurations": [
    {
        "CollectionName": "string",
        "CollectionParameters": {
            "string": "string"
        }
    },
],
"HookParameters": {
    "string": "string"
},
"LocalPath": "string",
"S3OutputPath": "string"
],
"DebugRuleConfigurations": [
    {
        "InstanceType": "string",
        "LocalPath": "string",
        "RuleConfigurationName": "string",
        "RuleEvaluatorImage": "string",
        "RuleParameters": {
            "string": "string"
        },
        "S3OutputPath": "string",
        "VolumeSizeInGB": number
    }
],
"DebugRuleEvaluationStatuses": [
    {
        "LastModifiedTime": number,
        "RuleConfigurationName": "string",
        "RuleEvaluationJobArn": "string",
        "RuleEvaluationStatus": "string",
        "StatusDetails": "string"
    }
],
"EnableInterContainerTrafficEncryption": boolean,
"EnableManagedSpotTraining": boolean,
"EnableNetworkIsolation": boolean,
"Environment": {
    "string": "string"
},
"ExperimentConfig": {
    "ExperimentName": "string",
    "TrialComponentDisplayName": "string",
    "TrialName": "string"
},
"FailureReason": "string",
"FinalMetricDataList": [
    {
        "MetricName": "string",
        "Timestamp": number,
        "Value": number
    }
],
"HyperParameters": {
    "string": "string"
},
"InputDataConfig": [
    {
        "ChannelName": "string",
        "CompressionType": "string",
        "ContentType": "string",
        "DataSource": {
            "FileSystemDataSource": {
                "DirectoryPath": "string",
            }
        }
    }
]
DescribeTrainingJob

"FileSystemAccessMode": "string",
"FileSystemId": "string",
"FileSystemType": "string"
},
"S3DataSource": {
"AttributeNames": [ "string" ],
"S3DataDistributionType": "string",
"S3DataType": "string",
"S3Uri": "string"
}
},
"InputMode": "string",
"RecordWrapperType": "string",
"ShuffleConfig": {
"Seed": number
}
],
"LabelingJobArn": "string",
"LastModifiedTime": number,
"ModelArtifacts": {
"S3ModelArtifacts": "string"
},
"OutputDataConfig": {
"KmsKeyId": "string",
"S3OutputPath": "string"
},
"ProfilerConfig": {
"ProfilingIntervalInMilliseconds": number,
"ProfilingParameters": {
"string": "string"
},
"S3OutputPath": "string"
},
"ProfilerRuleConfigurations": [
{
"InstanceType": "string",
"LocalPath": "string",
"RuleConfigurationName": "string",
"RuleEvaluatorImage": "string",
"RuleParameters": {
"string": "string"
},
"S3OutputPath": "string",
"VolumeSizeInGB": number
}
],
"ProfilerRuleEvaluationStatuses": [
{
"LastModifiedTime": number,
"RuleConfigurationName": "string",
"RuleEvaluationJobArn": "string",
"RuleEvaluationStatus": "string",
"StatusDetails": "string"
}
],
"ProfilingStatus": "string",
"ResourceConfig": {
"InstanceCount": number,
"InstanceType": "string",
"VolumeKmsKeyId": "string",
"VolumeSizeInGB": number
},
"RetryStrategy": {
"MaximumRetryAttempts": number
}
"RoleArn": "string",
"SecondaryStatus": "string",
"SecondaryStatusTransitions": [
  {
    "EndTime": number,
    "StartTime": number,
    "Status": "string",
    "StatusMessage": "string"
  }
],
"StoppingCondition": {
  "MaxRuntimeInSeconds": number,
  "MaxWaitTimeInSeconds": number
},
"TensorBoardOutputConfig": {
  "LocalPath": "string",
  "S3OutputPath": "string"
},
"TrainingEndTime": number,
"TrainingJobArn": "string",
"TrainingJobName": "string",
"TrainingJobStatus": "string",
"TrainingStartTime": number,
"TrainingTimeInSeconds": number,
"TuningJobArn": "string",
"VpcConfig": {
  "SecurityGroupIds": [ "string" ],
  "Subnets": [ "string" ]
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AlgorithmSpecification (p. 448)

Information about the algorithm used for training, and algorithm metadata.

Type: AlgorithmSpecification (p. 867) object

AutoMLJobArn (p. 448)

The Amazon Resource Name (ARN) of an AutoML job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*

BillableTimeIn_seconds (p. 448)

The billable time in seconds. Billable time refers to the absolute wall-clock time.

Multiply BillableTimeIn_seconds by the number of instances (InstanceCount) in your training cluster to get the total compute time SageMaker bills you if you run distributed training. The formula is as follows: BillableTimeIn_seconds \* InstanceCount.

You can calculate the savings from using managed spot training using the formula \((1 - \text{BillableTimeIn}_\text{seconds} / \text{TrainingTimeIn}_\text{seconds}) \* 100\). For example, if BillableTimeIn_seconds is 100 and TrainingTimeIn_seconds is 500, the savings is 80%.
Type: Integer

Valid Range: Minimum value of 1.

**CheckpointConfig (p. 448)**

Contains information about the output location for managed spot training checkpoint data.

Type: `CheckpointConfig (p. 945)` object

**CreationTime (p. 448)**

A timestamp that indicates when the training job was created.

Type: Timestamp

**DebugHookConfig (p. 448)**

Configuration information for the Debugger hook parameters, metric and tensor collections, and storage paths. To learn more about how to configure the `DebugHookConfig` parameter, see Use the SageMaker and Debugger Configuration API Operations to Create, Update, and Debug Your Training Job.

Type: `DebugHookConfig (p. 980)` object

**DebugRuleConfigurations (p. 448)**

Configuration information for Debugger rules for debugging output tensors.

Type: Array of `DebugRuleConfiguration (p. 982)` objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

**DebugRuleEvaluationStatuses (p. 448)**

Evaluation status of Debugger rules for debugging on a training job.

Type: Array of `DebugRuleEvaluationStatus (p. 984)` objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

**EnableInterContainerTrafficEncryption (p. 448)**

To encrypt all communications between ML compute instances in distributed training, choose `True`. Encryption provides greater security for distributed training, but training might take longer. How long it takes depends on the amount of communication between compute instances, especially if you use a deep learning algorithms in distributed training.

Type: Boolean

**EnableManagedSpotTraining (p. 448)**

A Boolean indicating whether managed spot training is enabled (`True`) or not (`False`).

Type: Boolean

**EnableNetworkIsolation (p. 448)**

If you want to allow inbound or outbound network calls, except for calls between peers within a training cluster for distributed training, choose `True`. If you enable network isolation for training jobs that are configured to use a VPC, SageMaker downloads and uploads customer data and model artifacts through the specified VPC, but the training container does not have network access.

Type: Boolean

**Environment (p. 448)**

The environment variables to set in the Docker container.
DescribeTrainingJob

Type: String to string map

Map Entries: Maximum number of 48 items.

Key Length Constraints: Maximum length of 512.

Key Pattern: [a-zA-Z\-\_][a-zA-Z0-9\-\_]*

Value Length Constraints: Maximum length of 512.

Value Pattern: \[\S\s\]*

ExperimentConfig (p. 448)

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:
- CreateProcessingJob (p. 162)
- CreateTrainingJob (p. 173)
- CreateTransformJob (p. 182)

Type: ExperimentConfig (p. 1029) object

FailureReason (p. 448)

If the training job failed, the reason it failed.

Type: String

Length Constraints: Maximum length of 1024.

FinalMetricDataList (p. 448)

A collection of MetricData objects that specify the names, values, and dates and times that the training algorithm emitted to Amazon CloudWatch.

Type: Array of MetricData (p. 1145) objects

Array Members: Minimum number of 0 items. Maximum number of 40 items.

HyperParameters (p. 448)

Algorithm-specific parameters.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 2500.

Value Pattern: .*

InputDataConfig (p. 448)

An array of Channel objects that describes each data input channel.

Type: Array of Channel (p. 941) objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.
LabelingJobArn (p. 448)

The Amazon Resource Name (ARN) of the SageMaker Ground Truth labeling job that created the transform or training job.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:labeling-job/.*`

LastModifiedTime (p. 448)

A timestamp that indicates when the status of the training job was last modified.

Type: Timestamp

ModelArtifacts (p. 448)

Information about the Amazon S3 location that is configured for storing model artifacts.

Type: ModelArtifacts (p. 1149) object

OutputDataConfig (p. 448)

The S3 path where model artifacts that you configured when creating the job are stored. SageMaker creates subfolders for model artifacts.

Type: OutputDataConfig (p. 1244) object

ProfilerConfig (p. 448)

Configuration information for Debugger system monitoring, framework profiling, and storage paths.

Type: ProfilerConfig (p. 1301) object

ProfilerRuleConfigurations (p. 448)

Configuration information for Debugger rules for profiling system and framework metrics.

Type: Array of ProfilerRuleConfiguration (p. 1305) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

ProfilerRuleEvaluationStatuses (p. 448)

Evaluation status of Debugger rules for profiling on a training job.

Type: Array of ProfilerRuleEvaluationStatus (p. 1307) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

ProfilingStatus (p. 448)

Profiling status of a training job.

Type: String

Valid Values: Enabled | Disabled

ResourceConfig (p. 448)

Resources, including ML compute instances and ML storage volumes, that are configured for model training.

Type: ResourceConfig (p. 1338) object
RetryStrategy (p. 448)

The number of times to retry the job when the job fails due to an InternalServerError.

Type: RetryStrategy (p. 1344) object

RoleArn (p. 448)

The AWS Identity and Access Management (IAM) role configured for the training job.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_\/]+$

SecondaryStatus (p. 448)

Provides detailed information about the state of the training job. For detailed information on the secondary status of the training job, see StatusMessage under SecondaryStatusTransition (p. 1359).

SageMaker provides primary statuses and secondary statuses that apply to each of them:

InProgress
  • Starting - Starting the training job.
  • Downloading - An optional stage for algorithms that support File training input mode. It indicates that data is being downloaded to the ML storage volumes.
  • Training - Training is in progress.
  • Interrupted - The job stopped because the managed spot training instances were interrupted.
  • Uploading - Training is complete and the model artifacts are being uploaded to the S3 location.

Completed
  • Completed - The training job has completed.

Failed
  • Failed - The training job has failed. The reason for the failure is returned in the FailureReason field of DescribeTrainingJobResponse.

Stopped
  • MaxRuntimeExceeded - The job stopped because it exceeded the maximum allowed runtime.
  • MaxWaitTimeExceeded - The job stopped because it exceeded the maximum allowed wait time.
  • Stopped - The training job has stopped.

Stopping
  • Stopping - Stopping the training job.

Important
  Valid values for SecondaryStatus are subject to change.

We no longer support the following secondary statuses:
  • LaunchingMLInstances
  • PreparingTraining
  • DownloadingTrainingImage

Type: String
Valid Values: Starting | LaunchingMLInstances | PreparingTrainingStack | Downloading | DownloadingTrainingImage | Training | Uploading | Stopping | Stopped | MaxRuntimeExceeded | Completed | Failed | Interrupted | MaxWaitTimeExceeded | Updating | Restarting

**SecondaryStatusTransitions (p. 448)**

A history of all of the secondary statuses that the training job has transitioned through.

Type: Array of SecondaryStatusTransition (p. 1359) objects

**StoppingCondition (p. 448)**

Specifies a limit to how long a model training job can run. It also specifies how long a managed Spot training job has to complete. When the job reaches the time limit, SageMaker ends the training job. Use this API to cap model training costs.

To stop a job, SageMaker sends the algorithm the SIGTERM signal, which delays job termination for 120 seconds. Algorithms can use this 120-second window to save the model artifacts, so the results of training are not lost.

Type: StoppingCondition (p. 1370) object

**TensorBoardOutputConfig (p. 448)**

Configuration of storage locations for the Debugger TensorBoard output data.

Type: TensorBoardOutputConfig (p. 1381) object

**TrainingEndTime (p. 448)**

Indicates the time when the training job ends on training instances. You are billed for the time interval between the value of TrainingStartTime and this time. For successful jobs and stopped jobs, this is the time after model artifacts are uploaded. For failed jobs, this is the time when SageMaker detects a job failure.

Type: Timestamp

**TrainingJobArn (p. 448)**

The Amazon Resource Name (ARN) of the training job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:training-job/.*

**TrainingJobName (p. 448)**

Name of the model training job.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

**TrainingJobStatus (p. 448)**

The status of the training job.

SageMaker provides the following training job statuses:

- InProgress - The training is in progress.
- Completed - The training job has completed.
DescribeTrainingJob

- Failed - The training job has failed. To see the reason for the failure, see the FailureReason field in the response to a DescribeTrainingJobResponse call.
- Stopping - The training job is stopping.
- Stopped - The training job has stopped.

For more detailed information, see SecondaryStatus.

Type: String

Valid Values:InProgress | Completed | Failed | Stopping | Stopped

TrainingStartTime (p. 448)

Indicates the time when the training job starts on training instances. You are billed for the time interval between this time and the value of TrainingEndTime. The start time in CloudWatch Logs might be later than this time. The difference is due to the time it takes to download the training data and to the size of the training container.

Type: Timestamp

TrainingTimeInSeconds (p. 448)

The training time in seconds.

Type: Integer

Valid Range: Minimum value of 1.

TuningJobArn (p. 448)

The Amazon Resource Name (ARN) of the associated hyperparameter tuning job if the training job was launched by a hyperparameter tuning job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]\{12\}:hyper-parameter-tuning-job/.*

VpcConfig (p. 448)

A VpcConfig (p. 1456) object that specifies the VPC that this training job has access to. For more information, see Protect Training Jobs by Using an Amazon Virtual Private Cloud.

Type: VpcConfig (p. 1456) object

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeTransformJob

Service: Amazon SageMaker Service

Returns information about a transform job.

Request Syntax

```
{
  "TransformJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**TransformJobName (p. 459)**

The name of the transform job that you want to view details of.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Syntax

```
{
  "AutoMLJobArn": "string",
  "BatchStrategy": "string",
  "CreationTime": number,
  "DataProcessing": {
    "InputFilter": "string",
    "JoinSource": "string",
    "OutputFilter": "string"
  },
  "Environment": {
    "string": "string"
  },
  "ExperimentConfig": {
    "ExperimentName": "string",
    "TrialComponentDisplayName": "string",
    "TrialName": "string"
  },
  "FailureReason": "string",
  "LabelingJobArn": "string",
  "MaxConcurrentTransforms": number,
  "MaxPayloadInMB": number,
  "ModelClientConfig": {
    "InvocationsMaxRetries": number,
    "InvocationsTimeoutInSeconds": number
  }
}
```
"ModelName": "string",
"TransformEndTime": number,
"TransformInput": {
    "CompressionType": "string",
    "ContentType": "string",
    "DataSource": {
        "S3DataSource": {
            "S3DataType": "string",
            "S3Uri": "string"
        }
    },
    "SplitType": "string"
},
"TransformJobArn": "string",
"TransformJobName": "string",
"TransformJobStatus": "string",
"TransformOutput": {
    "Accept": "string",
    "AssembleWith": "string",
    "KmsKeyId": "string",
    "S3OutputPath": "string"
},
"TransformResources": {
    "InstanceCount": number,
    "InstanceType": "string",
    "VolumeKmsKeyId": "string"
},
"TransformStartTime": number

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AutoMLJobArn (p. 459)**

The Amazon Resource Name (ARN) of the AutoML transform job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*

**BatchStrategy (p. 459)**

 Specifies the number of records to include in a mini-batch for an HTTP inference request. A record is a single unit of input data that inference can be made on. For example, a single line in a CSV file is a record.

To enable the batch strategy, you must set `SplitType` to `Line`, `RecordIO`, or `TFRecord`.

Type: String

Valid Values: MultiRecord | SingleRecord

**CreationTime (p. 459)**

A timestamp that shows when the transform Job was created.

Type: Timestamp
**DataProcessing (p. 459)**

The data structure used to specify the data to be used for inference in a batch transform job and to associate the data that is relevant to the prediction results in the output. The input filter provided allows you to exclude input data that is not needed for inference in a batch transform job. The output filter provided allows you to include input data relevant to interpreting the predictions in the output from the job. For more information, see *Associate Prediction Results with their Corresponding Input Records*.

Type: `DataProcessing (p. 971)` object

**Environment (p. 459)**

The environment variables to set in the Docker container. We support up to 16 key and values entries in the map.

Type: String to string map

Map Entries: Maximum number of 16 items.

Key Length Constraints: Maximum length of 1024.

Key Pattern: `[a-zA-Z_][a-zA-Z0-9_]\{0,1023\}`

Value Length Constraints: Maximum length of 10240.

Value Pattern: `[^\s]*`

**ExperimentConfig (p. 459)**

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:

- `CreateProcessingJob (p. 162)`
- `CreateTrainingJob (p. 173)`
- `CreateTransformJob (p. 182)`

Type: `ExperimentConfig (p. 1029)` object

**FailureReason (p. 459)**

If the transform job failed, `FailureReason` describes why it failed. A transform job creates a log file, which includes error messages, and stores it as an Amazon S3 object. For more information, see *Log Amazon SageMaker Events with Amazon CloudWatch*.

Type: String

Length Constraints: Maximum length of 1024.

**LabelingJobArn (p. 459)**

The Amazon Resource Name (ARN) of the Amazon SageMaker Ground Truth labeling job that created the transform or training job.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `arn:aws[a-zA-Z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:labeling-job/.*`

**MaxConcurrentTransforms (p. 459)**

The maximum number of parallel requests on each instance node that can be launched in a transform job. The default value is 1.
Type: Integer

Valid Range: Minimum value of 0.

**MaxPayloadInMB (p. 459)**

The maximum payload size, in MB, used in the transform job.

Type: Integer

Valid Range: Minimum value of 0.

**ModelClientConfig (p. 459)**

The timeout and maximum number of retries for processing a transform job invocation.

Type: ModelClientConfig (p. 1153) object

**ModelName (p. 459)**

The name of the model used in the transform job.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\[\]-\]*[a-zA-Z0-9\[\]]*

**TransformEndTime (p. 459)**

Indicates when the transform job has been completed, or has stopped or failed. You are billed for the time interval between this time and the value of TransformStartTime.

Type: Timestamp

**TransformInput (p. 459)**

Describes the dataset to be transformed and the Amazon S3 location where it is stored.

Type: TransformInput (p. 1403) object

**TransformJobArn (p. 459)**

The Amazon Resource Name (ARN) of the transform job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\[\]-]*:sagemaker:[a-zA-Z0-9\[\]-]*[0-9]\{12\}:transform-job/.*

**TransformJobName (p. 459)**

The name of the transform job.

Type: String


Pattern: ^[a-zA-Z0-9\[\]-\]*[a-zA-Z0-9\[\]]\{0,62\}

**TransformJobStatus (p. 459)**

The status of the transform job. If the transform job failed, the reason is returned in the FailureReason field.

Type: String
Valid Values: InProgress | Completed | Failed | Stopping | Stopped

**TransformOutput (p. 459)**

Identifies the Amazon S3 location where you want Amazon SageMaker to save the results from the transform job.

Type: TransformOutput (p. 1415) object

**TransformResources (p. 459)**

Describes the resources, including ML instance types and ML instance count, to use for the transform job.

Type: TransformResources (p. 1417) object

**TransformStartTime (p. 459)**

Indicates when the transform job starts on ML instances. You are billed for the time interval between this time and the value of TransformEndTime.

Type: Timestamp

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNot Found**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeTrial
Service: Amazon SageMaker Service
Provides a list of a trial's properties.

Request Syntax

```json
{
    "TrialName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**TrialName (p. 464)**

The name of the trial to describe.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: Yes

Response Syntax

```json
{
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "DisplayName": "string",
    "ExperimentName": "string",
    "LastModifiedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "LastModifiedTime": number,
    "MetadataProperties": {
        "CommitId": "string",
        "GeneratedBy": "string",
        "ProjectId": "string",
        "Repository": "string"
    },
    "Source": {
        "SourceArn": "string",
        "SourceType": "string"
    },
    "TrialArn": "string",
    "TrialName": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CreatedBy (p. 464)**
Who created the trial.
Type: UserContext (p. 1449) object

**CreationTime (p. 464)**
When the trial was created.
Type: Timestamp

**DisplayName (p. 464)**
The name of the trial as displayed. If `DisplayName` isn't specified, `TrialName` is displayed.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: `^[a-zA-Z0-9](\-*[a-zA-Z0-9])\{0,119\}`

**ExperimentName (p. 464)**
The name of the experiment the trial is part of.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: `^[a-zA-Z0-9](\-*[a-zA-Z0-9])\{0,119\}`

**LastModifiedBy (p. 464)**
Who last modified the trial.
Type: UserContext (p. 1449) object

**LastModifiedTime (p. 464)**
When the trial was last modified.
Type: Timestamp

**MetadataProperties (p. 464)**
Metadata properties of the tracking entity, trial, or trial component.
Type: MetadataProperties (p. 1143) object

**Source (p. 464)**
The Amazon Resource Name (ARN) of the source and, optionally, the job type.
Type: TrialSource (p. 1439) object

**TrialArn (p. 464)**
The Amazon Resource Name (ARN) of the trial.
Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial/.*

**TrialName (p. 464)**

The name of the trial.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

### Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

**ResourceNotFoundException**

Resource being accessed is not found.

HTTP Status Code: 400

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeTrialComponent
Service: Amazon SageMaker Service

Provides a list of a trials component's properties.

Request Syntax

```json
{
    "TrialComponentName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**TrialComponentName (p. 467)**

The name of the trial component to describe.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: (arn:aws[a-zA-Z-]*)*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:(experiment|experiment-trial|experiment-trial-component|artifact|action|context)\/(experiment-trial-component|artifact|action|context)\/)\?([a-zA-Z0-9-](-*[a-zA-Z0-9-])0,119)

Required: Yes

Response Syntax

```json
{
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "DisplayName": "string",
    "EndTime": number,
    "InputArtifacts": {
        "string": {
            "MediaType": "string",
            "Value": "string"
        }
    },
    "LastModifiedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "LastModifiedTime": number,
    "LineageGroupArn": "string",
    "MetadataProperties": {
        "CommitId": "string",
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreatedBy (p. 467)

Who created the trial component.

Type: UserContext (p. 1449) object

CreationTime (p. 467)

When the component was created.

Type: Timestamp

DisplayName (p. 467)

The name of the component as displayed. If DisplayName isn't specified, TrialComponentName is displayed.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9]\{0,119\}$

EndTime (p. 467)
When the component ended.
Type: Timestamp

InputArtifacts (p. 467)
The input artifacts of the component.
Type: String to TrialComponentArtifact (p. 1428) object map
Map Entries: Maximum number of 30 items.
Key Length Constraints: Maximum length of 64.
Key Pattern: .*

LastModifiedBy (p. 467)
Who last modified the component.
Type: UserContext (p. 1449) object

LastModifiedTime (p. 467)
When the component was last modified.
Type: Timestamp

LineageGroupArn (p. 467)
The Amazon Resource Name (ARN) of the lineage group.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]\{12\}:lineage-group/.*

MetadataProperties (p. 467)
Metadata properties of the tracking entity, trial, or trial component.
Type: MetadataProperties (p. 1143) object

Metrics (p. 467)
The metrics for the component.
Type: Array of TrialComponentMetricSummary (p. 1429) objects

OutputArtifacts (p. 467)
The output artifacts of the component.
Type: String to TrialComponentArtifact (p. 1428) object map
Map Entries: Maximum number of 30 items.
Key Length Constraints: Maximum length of 64.
Key Pattern: . *

Parameters (p. 467)

The hyperparameters of the component.

Type: String to TrialComponentParameterValue (p. 1431) object map

Map Entries: Maximum number of 150 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: . *

Source (p. 467)

The Amazon Resource Name (ARN) of the source and, optionally, the job type.

Type: TrialComponentSource (p. 1434) object

StartTime (p. 467)

When the component started.

Type: Timestamp

Status (p. 467)

The status of the component. States include:

- InProgress
- Completed
- Failed

Type: TrialComponentStatus (p. 1436) object

TrialComponentArn (p. 467)

The Amazon Resource Name (ARN) of the trial component.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-[a-zA-Z0-9]*:sagemaker:[a-z0-9\-[0-9]{12}:experiment-trial-component/.*

TrialComponentName (p. 467)

The name of the trial component.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9-]*[a-zA-Z0-9]{0,119}

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeUserProfile
Service: Amazon SageMaker Service

Describes a user profile. For more information, see CreateUserProfile.

Request Syntax

```
{
    "DomainId": "string",
    "UserProfileName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainId (p. 472)**

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**UserProfileName (p. 472)**

The user profile name. This value is not case sensitive.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Syntax

```
{
    "CreationTime": number,
    "DomainId": "string",
    "FailureReason": "string",
    "HomeEfsFileSystemUid": "string",
    "LastModifiedTime": number,
    "SingleSignOnUserIdentifier": "string",
    "SingleSignOnUserValue": "string",
    "Status": "string",
    "UserProfileArn": "string",
    "UserProfileName": "string",
    "UserSettings": {
        "ExecutionRole": "string",
        "JupyterServerAppSettings": {
            "DefaultResourceSpec": {
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
**CreationTime (p. 472)**

The creation time.

Type: Timestamp

**DomainId (p. 472)**

The ID of the domain that contains the profile.

Type: String

Length Constraints: Maximum length of 63.

**FailureReason (p. 472)**

The failure reason.

Type: String

Length Constraints: Maximum length of 1024.

**HomeEfsFileSystemUid (p. 472)**

The ID of the user's profile in the Amazon Elastic File System (EFS) volume.

Type: String

Length Constraints: Maximum length of 10.

Pattern: \d+

**LastModifiedTime (p. 472)**

The last modified time.

Type: Timestamp

**SingleSignOnUserIdentifier (p. 472)**

The SSO user identifier.

Type: String

Pattern: UserName

**SingleSignOnUserValue (p. 472)**

The SSO user value.

Type: String

Length Constraints: Maximum length of 256.

**Status (p. 472)**

The status.

Type: String

Valid Values: Deleting | Failed | InService | Pending | Updating | Update_Failed | Delete_Failed

**UserProfileArn (p. 472)**

The user profile Amazon Resource Name (ARN).

Type: String
DescribeUserProfile

Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:user-profile/.*

**UserProfileName (p. 472)**

The user profile name.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

**UserSettings (p. 472)**

A collection of settings.
Type: UserSettings (p. 1452) object

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.
HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DescribeWorkforce

Service: Amazon SageMaker Service

Lists private workforce information, including workforce name, Amazon Resource Name (ARN), and, if applicable, allowed IP address ranges (CIDRs). Allowable IP address ranges are the IP addresses that workers can use to access tasks.

**Important**
- This operation applies only to private workforces.

**Request Syntax**

```
{
   "WorkforceName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see [Common Parameters](p. 1470).

The request accepts the following data in JSON format.

**WorkforceName (p. 476)**

The name of the private workforce whose access you want to restrict. WorkforceName is automatically set to `default` when a workforce is created and cannot be modified.

Type: String


Pattern: `^[a-zA-Z0-9\-[a-zA-Z0-9\-]\{0,62\}$`

Required: Yes

**Response Syntax**

```
{
   "Workforce": {
      "CognitoConfig": {
         "ClientId": "string",
         "UserPool": "string"
      },
      "CreateDate": number,
      "LastUpdatedDate": number,
      "OidcConfig": {
         "AuthorizationEndpoint": "string",
         "ClientId": "string",
         "Issuer": "string",
         "JwksUri": "string",
         "LogoutEndpoint": "string",
         "TokenEndpoint": "string",
         "UserInfoEndpoint": "string"
      },
      "SourceIpConfig": {
         "Cidrs": [ "string" ]
      }
   }
```

476
"SubDomain": "string",
"WorkforceArn": "string",
"WorkforceName": "string"
}
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Workforce (p. 476)

A single private workforce, which is automatically created when you create your first private work
team. You can create one private workforce in each AWS Region. By default, any workforce-related
API operation used in a specific region will apply to the workforce created in that region. To learn
how to create a private workforce, see Create a Private Workforce.

Type: Workforce (p. 1457) object

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
DescribeWorkteam
Service: Amazon SageMaker Service

Gets information about a specific work team. You can see information such as the create date, the last updated date, membership information, and the work team's Amazon Resource Name (ARN).

Request Syntax

```
{
   "WorkteamName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**WorkteamName (p. 478)**

The name of the work team to return a description of.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Syntax

```
{
   "Workteam": {
      "CreateDate": number,
      "Description": "string",
      "LastUpdatedDate": number,
      "MemberDefinitions": [ 
         { "CognitoMemberDefinition": { "ClientId": "string", "UserGroup": "string", "UserPool": "string" }, "OidcMemberDefinition": { "Groups": [ "string" ] } },
         "NotificationConfiguration": { "NotificationTopicArn": "string" },
         "ProductListingIds": [ "string" ],
         "SubDomain": "string",
         "WorkforceArn": "string",
         "WorkteamArn": "string",
         "WorkteamName": "string"
   }
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**Workteam (p. 478)**

A `Workteam` instance that contains information about the work team.

Type: `Workteam (p. 1459)` object

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DisableSagemakerServicecatalogPortfolio

Service: Amazon SageMaker Service

Disables using Service Catalog in SageMaker. Service Catalog is used to create SageMaker projects.

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
DisassociateTrialComponent
Service: Amazon SageMaker Service

Disassociates a trial component from a trial. This doesn't effect other trials the component is associated with. Before you can delete a component, you must disassociate the component from all trials it is associated with. To associate a trial component with a trial, call the AssociateTrialComponent (p. 19) API.

To get a list of the trials a component is associated with, use the Search (p. 691) API. Specify ExperimentTrialComponent for the Resource parameter. The list appears in the response under Results.TrialComponent.Parents.

Request Syntax

```
{
  "TrialComponentName": "string",
  "TrialName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**TrialComponentName (p. 481)**

The name of the component to disassociate from the trial.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: Yes

**TrialName (p. 481)**

The name of the trial to disassociate from.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: Yes

Response Syntax

```
{
  "TrialArn": "string",
  "TrialComponentArn": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**TrialArn (p. 481)**

The Amazon Resource Name (ARN) of the trial.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial/.*

**TrialComponentArn (p. 481)**

The ARN of the trial component.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial-component/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
EnableSagemakerServicecatalogPortfolio
Service: Amazon SageMaker Service

Enables using Service Catalog in SageMaker. Service Catalog is used to create SageMaker projects.

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
GetDeviceFleetReport
Service: Amazon SageMaker Service

Describes a fleet.

Request Syntax

```json
{
  "DeviceFleetName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeviceFleetName (p. 484)**

The name of the fleet.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Syntax

```json
{
  "AgentVersions": [
    {"AgentCount": number,
      "Version": "string"
    }
  ],
  "Description": "string",
  "DeviceFleetArn": "string",
  "DeviceFleetName": "string",
  "DeviceStats": {
    "ConnectedDeviceCount": number,
    "RegisteredDeviceCount": number
  },
  "ModelStats": [
    {"ActiveDeviceCount": number,
      "ConnectedDeviceCount": number,
      "ModelName": "string",
      "ModelVersion": "string",
      "OfflineDeviceCount": number,
      "SamplingDeviceCount": number
    }
  ],
  "OutputConfig": {
    "KmsKeyId": "string",
    ""string": "string"
  },
  ""string": "string"
```

484
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AgentVersions (p. 484)**

The versions of Edge Manager agent deployed on the fleet.

Type: Array of AgentVersion (p. 865) objects

**Description (p. 484)**

Description of the fleet.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 800.

Pattern: \[S\s]*

**DeviceFleetArn (p. 484)**

The Amazon Resource Name (ARN) of the device.

Type: String

Pattern: ^arn:aws[a-zA-Z-]*:iam::\d{12}:device-fleet/?[a-zA-Z0-9+=,.@-_/]+$  

**DeviceFleetName (p. 484)**

The name of the fleet.

Type: String


Pattern: ^[a-zA-Z0-9]+\*\[a-zA-Z0-9]+\}\{0,63}$  

**DeviceStats (p. 484)**

Status of devices.

Type: DeviceStats (p. 991) object

**ModelStats (p. 484)**

Status of model on device.

Type: Array of EdgeModelStat (p. 1005) objects

**OutputConfig (p. 484)**

The output configuration for storing sample data collected by the fleet.

Type: EdgeOutputConfig (p. 1008) object
**ReportGenerated (p. 484)**

Timestamp of when the report was generated.

Type: Timestamp

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
GetLineageGroupPolicy

Service: Amazon SageMaker Service

The resource policy for the lineage group.

Request Syntax

```json
{
   "LineageGroupName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

LineageGroupName (p. 487)

The name or Amazon Resource Name (ARN) of the lineage group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:lineage-group/)? ([a-zA-Z0-9]{0,119})

Required: Yes

Response Syntax

```json
{
   "LineageGroupArn": "string",
   "ResourcePolicy": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

LineageGroupArn (p. 487)

The Amazon Resource Name (ARN) of the lineage group.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:lineage-group/.*

ResourcePolicy (p. 487)

The resource policy that gives access to the lineage group in another account.

487
Type: String
Length Constraints: Maximum length of 20480.
Pattern: .*(?:[ \r\n\t].*)*

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound
   Resource being access is not found.
   HTTP Status Code: 400

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
GetModelPackageGroupPolicy

Service: Amazon SageMaker Service

Gets a resource policy that manages access for a model group. For information about resource policies, see Identity-based policies and resource-based policies in the AWS Identity and Access Management User Guide.

Request Syntax

```
{
  "ModelPackageGroupName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ModelPackageGroupName (p. 489)

The name of the model group for which to get the resource policy.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Syntax

```
{
  "ResourcePolicy": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ResourcePolicy (p. 489)

The resource policy for the model group.

Type: String


Pattern: .*
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
GetSagemakerServicecatalogPortfolioStatus

Service: Amazon SageMaker Service

Gets the status of Service Catalog in SageMaker. Service Catalog is used to create SageMaker projects.

Response Syntax

{
   "Status": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Status (p. 491)

Whether Service Catalog is enabled or disabled in SageMaker.

Type: String

Valid Values: Enabled | Disabled

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
GetSearchSuggestions

An auto-complete API for the search functionality in the Amazon SageMaker console. It returns suggestions of possible matches for the property name to use in Search queries. Provides suggestions for HyperParameters, Tags, and Metrics.

Request Syntax

```json
{
   "Resource": "string",
   "SuggestionQuery": {
      "PropertyNameQuery": {
         "PropertyNameHint": "string"
      }
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**Resource** (p. 492)

The name of the Amazon SageMaker resource to search for.

Type: String

Valid Values: TrainingJob | Experiment | ExperimentTrial | ExperimentTrialComponent | Endpoint | ModelPackage | ModelPackageGroup | Pipeline | PipelineExecution | FeatureGroup | Project

Required: Yes

**SuggestionQuery** (p. 492)

Limits the property names that are included in the response.

Type: SuggestionQuery (p. 1376) object

Required: No

Response Syntax

```json
{
   "PropertyNameSuggestions": [
      {
         "PropertyName": "string"
      }
   ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**PropertyNameSuggestions (p. 492)**

A list of property names for a Resource that match a SuggestionQuery.

Type: Array of PropertyNameSuggestion (p. 1315) objects

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListActions
Service: Amazon SageMaker Service

Lists the actions in your account and their properties.

Request Syntax

```
{
  "ActionType": "string",
  "CreatedAfter": number,
  "CreatedBefore": number,
  "MaxResults": number,
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "SourceUri": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ActionType (p. 494)**

A filter that returns only actions of the specified type.

Type: String

Length Constraints: Maximum length of 256.

Required: No

**CreatedAfter (p. 494)**

A filter that returns only actions created on or after the specified time.

Type: Timestamp

Required: No

**CreatedBefore (p. 494)**

A filter that returns only actions created on or before the specified time.

Type: Timestamp

Required: No

**MaxResults (p. 494)**

The maximum number of actions to return in the response. The default value is 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No
**NextToken (p. 494)**

If the previous call to `ListActions` didn't return the full set of actions, the call returns a token for getting the next set of actions.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 494)**

The property used to sort results. The default value is `CreationTime`.

Type: String

Valid Values: Name | CreationTime

Required: No

**SortOrder (p. 494)**

The sort order. The default value is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

**SourceUri (p. 494)**

A filter that returns only actions with the specified source URI.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

### Response Syntax

```json
{
   "ActionSummaries": [
      {
         "ActionArn": "string",
         "ActionName": "string",
         "ActionType": "string",
         "CreationTime": number,
         "LastModifiedTime": number,
         "Source": {
            "SourceId": "string",
            "SourceType": "string",
            "SourceUri": "string"
         },
         "Status": "string"
      }
   ]
}
```
"NextToken": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**ActionSummaries (p. 495)**

A list of actions and their properties.

Type: Array of ActionSummary (p. 861) objects

**NextToken (p. 495)**

A token for getting the next set of actions, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
**ListAlgorithms**  
*Service: Amazon SageMaker Service*

Lists the machine learning algorithms that have been created.

**Request Syntax**

```json
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 497)**

A filter that returns only algorithms created after the specified time (timestamp).

Type: Timestamp  
Required: No

**CreationTimeBefore (p. 497)**

A filter that returns only algorithms created before the specified time (timestamp).

Type: Timestamp  
Required: No

**MaxResults (p. 497)**

The maximum number of algorithms to return in the response.

Type: Integer  
Valid Range: Minimum value of 1. Maximum value of 100.  
Required: No

**NameContains (p. 497)**

A string in the algorithm name. This filter returns only algorithms whose name contains the specified string.

Type: String  
Length Constraints: Maximum length of 63.  
Pattern: [a-zA-Z0-9-\-]+
ListAlgorithms

Required: No

NextToken (p. 497)

If the response to a previous ListAlgorithms request was truncated, the response includes a NextToken. To retrieve the next set of algorithms, use the token in the next request.

Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

SortBy (p. 497)

The parameter by which to sort the results. The default is CreationTime.

Type: String
Valid Values: Name | CreationTime
Required: No

SortOrder (p. 497)

The sort order for the results. The default is Ascending.

Type: String
Valid Values: Ascending | Descending
Required: No

Response Syntax

```json
{
  "AlgorithmSummaryList": [
    {
      "AlgorithmArn": "string",
      "AlgorithmDescription": "string",
      "AlgorithmName": "string",
      "AlgorithmStatus": "string",
      "CreationTime": number
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AlgorithmSummaryList (p. 498)

-An array of AlgorithmSummary objects, each of which lists an algorithm.

Type: Array of AlgorithmSummary (p. 872) objects
NextToken (p. 498)

If the response is truncated, SageMaker returns this token. To retrieve the next set of algorithms, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*  

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListAppImageConfigs
Service: Amazon SageMaker Service

Lists the AppImageConfigs in your account and their properties. The list can be filtered by creation time or modified time, and whether the AppImageConfig name contains a specified string.

Request Syntax

```
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "MaxResults": number,
   "ModifiedTimeAfter": number,
   "ModifiedTimeBefore": number,
   "NameContains": "string",
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 500)

A filter that returns only AppImageConfigs created on or after the specified time.

Type: Timestamp

Required: No

CreationTimeBefore (p. 500)

A filter that returns only AppImageConfigs created on or before the specified time.

Type: Timestamp

Required: No

MaxResults (p. 500)

The maximum number of AppImageConfigs to return in the response. The default value is 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

ModifiedTimeAfter (p. 500)

A filter that returns only AppImageConfigs modified on or after the specified time.

Type: Timestamp

Required: No
**ModifiedTimeBefore (p. 500)**

A filter that returns only AppImageConfigs modified on or before the specified time.

Type: Timestamp

Required: No

**NameContains (p. 500)**

A filter that returns only AppImageConfigs whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: No

**NextToken (p. 500)**

If the previous call to ListImages didn’t return the full set of AppImageConfigs, the call returns a token for getting the next set of AppImageConfigs.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 500)**

The property used to sort results. The default value is CreationTime.

Type: String

Valid Values: CreationTime | LastModifiedTime | Name

Required: No

**SortOrder (p. 500)**

The sort order. The default value is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```json
{
  "AppImageConfigs": [
    {
      "AppImageConfigArn": "string",
      "AppImageConfigName": "string",
      "CreationTime": number,
      "KernelGatewayImageConfig": {
        "FileSystemConfig": {
          ...
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AppImageConfigs (p. 501)**

A list of AppImageConfigs and their properties.

Type: Array of AppImageConfigDetails (p. 889) objects

**NextToken (p. 501)**

A token for getting the next set of AppImageConfigs, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListApps
Service: Amazon SageMaker Service

Lists apps.

Request Syntax

```json
{
    "DomainIdEquals": "string",
    "MaxResults": number,
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string",
    "UserProfileNameEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainIdEquals (p. 503)**

A parameter to search for the domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: No

**MaxResults (p. 503)**

Returns a list up to a specified limit.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NextToken (p. 503)**

If the previous response was truncated, you will receive this token. Use it in your next request to receive the next set of results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

Required: No

**SortBy (p. 503)**

The parameter by which to sort the results. The default is CreationTime.

Type: String
Valid Values: CreationTime
    Required: No
**SortOrder (p. 503)**
    The sort order for the results. The default is Ascending.
    Type: String
    Valid Values: Ascending | Descending
    Required: No
**UserProfileNameEquals (p. 503)**
    A parameter to search by user profile name.
    Type: String
    Length Constraints: Maximum length of 63.
    Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}
    Required: No

**Response Syntax**

```json
{
  "Apps": [
    {
      "AppName": "string",
      "AppType": "string",
      "CreationTime": number,
      "DomainId": "string",
      "Status": "string",
      "UserProfileName": "string"
    }
  ],
  "NextToken": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**Apps (p. 504)**

The list of apps.

Type: Array of AppDetails (p. 887) objects

**NextToken (p. 504)**

If the previous response was truncated, you will receive this token. Use it in your next request to receive the next set of results.

Type: String

Length Constraints: Maximum length of 8192.
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListArtifacts

Lists the artifacts in your account and their properties.

Request Syntax

```
{
  "ArtifactType": "string",
  "CreatedAfter": number,
  "CreatedBefore": number,
  "MaxResults": number,
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "SourceUri": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ArtifactType (p. 506)**

A filter that returns only artifacts of the specified type.

Type: String

Length Constraints: Maximum length of 256.

Required: No

**CreatedAfter (p. 506)**

A filter that returns only artifacts created on or after the specified time.

Type: Timestamp

Required: No

**CreatedBefore (p. 506)**

A filter that returns only artifacts created on or before the specified time.

Type: Timestamp

Required: No

**MaxResults (p. 506)**

The maximum number of artifacts to return in the response. The default value is 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No
NextToken (p. 506)

If the previous call to ListArtifacts didn't return the full set of artifacts, the call returns a token for getting the next set of artifacts.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

SortBy (p. 506)

The property used to sort results. The default value is CreationTime.

Type: String

Valid Values: CreationTime

Required: No

SortOrder (p. 506)

The sort order. The default value is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

SourceUri (p. 506)

A filter that returns only artifacts with the specified source URI.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

Response Syntax

```json
{
    "ArtifactSummaries": [
        {
            "ArtifactArn": "string",
            "ArtifactName": "string",
            "ArtifactType": "string",
            "CreationTime": number,
            "LastModifiedTime": number,
            "Source": {
                "SourceTypes": [
                    {
                        "SourceIdType": "string",
                        "Value": "string"
                    }
                ],
            "SourceUri": "string"
        }
    ]
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ArtifactSummaries (p. 507)**

A list of artifacts and their properties.

- Type: Array of ArtifactSummary (p. 894) objects

**NextToken (p. 507)**

A token for getting the next set of artifacts, if there are any.

- Type: String
- Length Constraints: Maximum length of 8192.
- Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListAssociations

Service: Amazon SageMaker Service

Lists the associations in your account and their properties.

Request Syntax

```
{
   "AssociationType": "string",
   "CreatedAfter": number,
   "CreatedBefore": number,
   "DestinationArn": "string",
   "DestinationType": "string",
   "MaxResults": number,
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string",
   "SourceArn": "string",
   "SourceType": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AssociationType (p. 509)**

A filter that returns only associations of the specified type.

Type: String

Valid Values: ContributedTo | AssociatedWith | DerivedFrom | Produced

Required: No

**CreatedAfter (p. 509)**

A filter that returns only associations created on or after the specified time.

Type: Timestamp

Required: No

**CreatedBefore (p. 509)**

A filter that returns only associations created on or before the specified time.

Type: Timestamp

Required: No

**DestinationArn (p. 509)**

A filter that returns only associations with the specified destination Amazon Resource Name (ARN).

Type: String

Length Constraints: Maximum length of 256.
Pattern: \texttt{arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]\{12\}:(experiment\ |experiment-trial-component|artifact|action|context)/.*}

Required: No

\textbf{DestinationType (p. 509)}

A filter that returns only associations with the specified destination type.

Type: String

Length Constraints: Maximum length of 256.

Required: No

\textbf{MaxResults (p. 509)}

The maximum number of associations to return in the response. The default value is 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

\textbf{NextToken (p. 509)}

If the previous call to \texttt{ListAssociations} didn't return the full set of associations, the call returns a token for getting the next set of associations.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

\textbf{SortBy (p. 509)}

The property used to sort results. The default value is \texttt{CreationTime}.

Type: String

Valid Values: SourceArn | DestinationArn | SourceType | DestinationType | CreationTime

Required: No

\textbf{SortOrder (p. 509)}

The sort order. The default value is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

\textbf{SourceArn (p. 509)}

A filter that returns only associations with the specified source ARN.

Type: String
ListAssociations

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

Required: No

SourceType (p. 509)
A filter that returns only associations with the specified source type.

Type: String

Length Constraints: Maximum length of 256.

Required: No

Response Syntax

```json
{
  "AssociationSummaries": [
    {
      "AssociationType": "string",
      "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
      },
      "CreationTime": number,
      "DestinationArn": "string",
      "DestinationName": "string",
      "DestinationType": "string",
      "SourceArn": "string",
      "SourceName": "string",
      "SourceType": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AssociationSummaries (p. 511)**
A list of associations and their properties.

Type: Array of AssociationSummary (p. 896) objects

**NextToken (p. 511)**
A token for getting the next set of associations, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

  Resource being access is not found.

  HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListAutoMLJobs
Service: Amazon SageMaker Service

Request a list of jobs.

Request Syntax

```
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "LastModifiedTimeAfter": number,
   "LastModifiedTimeBefore": number,
   "MaxResults": number,
   "NameContains": "string",
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string",
   "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 513)**

Request a list of jobs, using a filter for time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 513)**

Request a list of jobs, using a filter for time.

Type: Timestamp

Required: No

**LastModifiedTimeAfter (p. 513)**

Request a list of jobs, using a filter for time.

Type: Timestamp

Required: No

**LastModifiedTimeBefore (p. 513)**

Request a list of jobs, using a filter for time.

Type: Timestamp

Required: No

**MaxResults (p. 513)**

Request a list of jobs up to a specified limit.
ListAutoMLJobs

Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

NameContains (p. 513)
Request a list of jobs, using a search filter for name.
Type: String
Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9\-]+
Required: No

NextToken (p. 513)
If the previous response was truncated, you receive this token. Use it in your next request to receive the next set of results.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

SortBy (p. 513)
The parameter by which to sort the results. The default is Name.
Type: String
Valid Values: Name | CreationTime | Status
Required: No

SortOrder (p. 513)
The sort order for the results. The default is Descending.
Type: String
Valid Values: Ascending | Descending
Required: No

StatusEquals (p. 513)
Request a list of jobs, using a filter for status.
Type: String
Valid Values: Completed | InProgress | Failed | Stopped | Stopping
Required: No

Response Syntax

```json
{

```
"AutoMLJobSummaries": [ 
  { 
    "AutoMLJobArn": "string",
    "AutoMLJobName": "string",
    "AutoMLJobSecondaryStatus": "string",
    "AutoMLJobStatus": "string",
    "CreationTime": number,
    "EndTime": number,
    "FailureReason": "string",
    "LastModifiedTime": number,
    "PartialFailureReasons": [ 
      { 
        "PartialFailureMessage": "string"
      } 
    ]
  }, 
  ],
"NextToken": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AutoMLJobSummaries (p. 514)**

Returns a summary list of jobs.

Type: Array of AutoMLJobSummary (p. 919) objects

**NextToken (p. 514)**

If the previous response was truncated, you receive this token. Use it in your next request to receive the next set of results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
ListCandidatesForAutoMLJob

Service: Amazon SageMaker Service

List the candidates created for the job.

Request Syntax

```json
{
    "AutoMLJobName": "string",
    "CandidateNameEquals": "string",
    "MaxResults": number,
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string",
    "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AutoMLJobName (p. 517)**

List the candidates created for the job by providing the job's name.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}$`

Required: Yes

**CandidateNameEquals (p. 517)**

List the candidates for the job and filter by candidate name.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Required: No

**MaxResults (p. 517)**

List the job's candidates up to a specified limit.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NextToken (p. 517)**

If the previous response was truncated, you receive this token. Use it in your next request to receive the next set of results.

Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

**SortBy (p. 517)**

The parameter by which to sort the results. The default is Descending.

Type: String

Valid Values: CreationTime | Status | FinalObjectiveMetricValue

Required: No

**SortOrder (p. 517)**

The sort order for the results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No

**StatusEquals (p. 517)**

List the candidates for the job and filter by status.

Type: String

Valid Values: Completed | InProgress | Failed | Stopped | Stopping

Required: No

**Response Syntax**

```json
{
  "Candidates": [
    {
      "CandidateName": "string",
      "CandidateProperties": {
        "CandidateArtifactLocations": {
          "Explainability": "string",
          "ModelInsights": "string"
        },
        "CandidateMetrics": [
          {
            "MetricName": "string",
            "Set": "string",
            "StandardMetricName": "string",
            "Value": number
          }
        ],
      },
      "CandidateStatus": "string",
      "CandidateSteps": [
        {
          "CandidateStepArn": "string",
          "CandidateStepName": "string",
          "CandidateStepType": "string"
        }
      ],
    }
  ]
}
```
"CreationTime": number,
"EndTime": number,
"FailureReason": "string",
"FinalAutoMLJobObjectiveMetric": {
  "MetricName": "string",
  "Type": "string",
  "Value": number
},
"InferenceContainers": [
  {
    "Environment": {
      "string": "string"
    },
    "Image": "string",
    "ModelDataUrl": "string"
  }
],
"LastModifiedTime": number,
"ObjectiveStatus": "string"
],
"NextToken": "string"

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Candidates (p. 518)

Summaries about the AutoMLCandidates.

Type: Array of AutoMLCandidate (p. 904) objects

NextToken (p. 518)

If the previous response was truncated, you receive this token. Use it in your next request to receive the next set of results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
ListCodeRepositories

Service: Amazon SageMaker Service

Gets a list of the Git repositories in your account.

Request Syntax

```json
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "LastModifiedTimeAfter": number,
    "LastModifiedTimeBefore": number,
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 521)**

A filter that returns only Git repositories that were created after the specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 521)**

A filter that returns only Git repositories that were created before the specified time.

Type: Timestamp

Required: No

**LastModifiedTimeAfter (p. 521)**

A filter that returns only Git repositories that were last modified after the specified time.

Type: Timestamp

Required: No

**LastModifiedTimeBefore (p. 521)**

A filter that returns only Git repositories that were last modified before the specified time.

Type: Timestamp

Required: No

**MaxResults (p. 521)**

The maximum number of Git repositories to return in the response.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**NameContains (p. 521)**
A string in the Git repositories name. This filter returns only repositories whose name contains the specified string.

Type: String
Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9-]+
Required: No

**NextToken (p. 521)**
If the result of a ListCodeRepositoriesOutput request was truncated, the response includes a NextToken. To get the next set of Git repositories, use the token in the next request.

Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

**SortBy (p. 521)**
The field to sort results by. The default is Name.

Type: String
Valid Values: Name | CreationTime | LastModifiedTime
Required: No

**SortOrder (p. 521)**
The sort order for results. The default is Ascending.

Type: String
Valid Values: Ascending | Descending
Required: No

**Response Syntax**

```json
{
   "CodeRepositorySummaryList": [
      {
         "CodeRepositoryArn": "string",
         "CodeRepositoryName": "string",
         "CreationTime": number,
         "GitConfig": {
            "Branch": "string",
            "RepositoryUrl": "string"
         }
      }
   ]
}
```
"SecretArn": "string"
},
"LastModifiedTime": number
}
],
"NextToken": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CodeRepositorySummaryList (p. 522)**

Gets a list of summaries of the Git repositories. Each summary specifies the following values for the repository:

- Name
- Amazon Resource Name (ARN)
- Creation time
- Last modified time
- Configuration information, including the URL location of the repository and the ARN of the AWS Secrets Manager secret that contains the credentials used to access the repository.

Type: Array of `CodeRepositorySummary (p. 948)` objects

**NextToken (p. 522)**

If the result of a `ListCodeRepositoriesOutput` request was truncated, the response includes a `NextToken`. To get the next set of Git repositories, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see `Common Errors (p. 1472)`.

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListCompilationJobs

Service: Amazon SageMaker Service

Lists model compilation jobs that satisfy various filters.

To create a model compilation job, use CreateCompilationJob (p. 46). To get information about a particular model compilation job you have created, use DescribeCompilationJob (p. 308).

Request Syntax

```json
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "LastModifiedTimeAfter": number,
   "LastModifiedTimeBefore": number,
   "MaxResults": number,
   "NameContains": "string",
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string",
   "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 525)
A filter that returns the model compilation jobs that were created after a specified time.

Type: Timestamp

Required: No

CreationTimeBefore (p. 525)
A filter that returns the model compilation jobs that were created before a specified time.

Type: Timestamp

Required: No

LastModifiedTimeAfter (p. 525)
A filter that returns the model compilation jobs that were modified after a specified time.

Type: Timestamp

Required: No

LastModifiedTimeBefore (p. 525)
A filter that returns the model compilation jobs that were modified before a specified time.

Type: Timestamp

Required: No
MaxResults (p. 525)

The maximum number of model compilation jobs to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NameContains (p. 525)

A filter that returns the model compilation jobs whose name contains a specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: \[a-zA-Z0-9\-\]+

Required: No

NextToken (p. 525)

If the result of the previous ListCompilationJobs request was truncated, the response includes a NextToken. To retrieve the next set of model compilation jobs, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

SortBy (p. 525)

The field by which to sort results. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime | Status

Required: No

SortOrder (p. 525)

The sort order for results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No

StatusEquals (p. 525)

A filter that retrieves model compilation jobs with a specific DescribeCompilationJob:CompilationJobStatus (p. 309) status.

Type: String

Valid Values: INPROGRESS | COMPLETED | FAILED | STARTING | STOPPING | STOPPED

Required: No
Response Syntax

```json
{
  "CompilationJobSummaries": [
    {
      "CompilationEndTime": number,
      "CompilationJobArn": "string",
      "CompilationJobName": "string",
      "CompilationJobStatus": "string",
      "CompilationStartTime": number,
      "CompilationTargetDevice": "string",
      "CompilationTargetPlatformAccelerator": "string",
      "CompilationTargetPlatformArch": "string",
      "CompilationTargetPlatformOs": "string",
      "CreationTime": number,
      "LastModifiedTime": number
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**CompilationJobSummaries (p. 527)**

An array of CompilationJobSummary (p. 953) objects, each describing a model compilation job.

Type: Array of CompilationJobSummary (p. 953) objects

**NextToken (p. 527)**

If the response is truncated, Amazon SageMaker returns this NextToken. To retrieve the next set of model compilation jobs, use this token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript

527
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
ListContexts
Service: Amazon SageMaker Service

Lists the contexts in your account and their properties.

Request Syntax

```json
{
  "ContextType": "string",
  "CreatedAfter": number,
  "CreatedBefore": number,
  "MaxResults": number,
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "SourceUri": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ContextType (p. 529)**

A filter that returns only contexts of the specified type.

- Type: String
- Length Constraints: Maximum length of 256.
- Required: No

**CreatedAfter (p. 529)**

A filter that returns only contexts created on or after the specified time.

- Type: Timestamp
- Required: No

**CreatedBefore (p. 529)**

A filter that returns only contexts created on or before the specified time.

- Type: Timestamp
- Required: No

**MaxResults (p. 529)**

The maximum number of contexts to return in the response. The default value is 10.

- Type: Integer
- Valid Range: Minimum value of 1. Maximum value of 100.
- Required: No
**NextToken (p. 529)**

If the previous call to `ListContexts` didn't return the full set of contexts, the call returns a token for getting the next set of contexts.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 529)**

The property used to sort results. The default value is `CreationTime`.

Type: String

Valid Values: Name | CreationTime

Required: No

**SortOrder (p. 529)**

The sort order. The default value is `Descending`.

Type: String

Valid Values: Ascending | Descending

Required: No

**SourceUri (p. 529)**

A filter that returns only contexts with the specified source URI.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

**Response Syntax**

```
{
  "ContextSummaries": [
    {
      "ContextArn": "string",
      "ContextName": "string",
      "ContextType": "string",
      "CreationTime": number,
      "LastModifiedTime": number,
      "Source": {
        "SourceId": "string",
        "SourceType": "string",
        "SourceUri": "string"
      }
    }
  ],
  "NextToken": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ContextSummaries** *(p. 530)*

A list of contexts and their properties.

Type: Array of **ContextSummary** *(p. 960)* objects

**NextToken** *(p. 530)*

A token for getting the next set of contexts, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

Errors

For information about the errors that are common to all actions, see **Common Errors** *(p. 1472)*.

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListDataQualityJobDefinitions

Service: Amazon SageMaker Service

Lists the data quality job definitions in your account.

Request Syntax

```
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "EndpointName": "string",
   "MaxResults": number,
   "NameContains": "string",
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 532)

A filter that returns only data quality monitoring job definitions created after the specified time.

Type: Timestamp

Required: No

CreationTimeBefore (p. 532)

A filter that returns only data quality monitoring job definitions created before the specified time.

Type: Timestamp

Required: No

EndpointName (p. 532)

A filter that lists the data quality job definitions associated with the specified endpoint.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-\[\]*[a-zA-Z0-9\-\]]\{0,62}$

Required: No

MaxResults (p. 532)

The maximum number of data quality monitoring job definitions to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**NameContains (p. 532)**
A string in the data quality monitoring job definition name. This filter returns only data quality monitoring job definitions whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: [a-zA-Z0-9\-\_]+

Required: No

**NextToken (p. 532)**
If the result of the previous ListDataQualityJobDefinitions request was truncated, the response includes a NextToken. To retrieve the next set of transform jobs, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 532)**
The field to sort results by. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No

**SortOrder (p. 532)**
The sort order for results. The default is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```
{
  "JobDefinitionSummaries": [
    {
      "CreationTime": number,
      "EndpointName": "string",
      "MonitoringJobDefinitionArn": "string",
      "MonitoringJobDefinitionName": "string"
    }
  ],
  "NextToken": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**JobDefinitionSummaries (p. 533)**

A list of data quality monitoring job definitions.

Type: Array of MonitoringJobDefinitionSummary (p. 1205) objects

**NextToken (p. 533)**

If the result of the previous ListDataQualityJobDefinitions request was truncated, the response includes a NextToken. To retrieve the next set of data quality monitoring job definitions, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListDeviceFleets
Service: Amazon SageMaker Service

Returns a list of devices in the fleet.

Request Syntax

```
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "LastModifiedTimeAfter": number,
   "LastModifiedTimeBefore": number,
   "MaxResults": number,
   "NameContains": "string",
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 535)
Filter fleets where packaging job was created after specified time.
Type: Timestamp
Required: No

CreationTimeBefore (p. 535)
Filter fleets where the edge packaging job was created before specified time.
Type: Timestamp
Required: No

LastModifiedTimeAfter (p. 535)
Select fleets where the job was updated after X
Type: Timestamp
Required: No

LastModifiedTimeBefore (p. 535)
Select fleets where the job was updated before X
Type: Timestamp
Required: No

MaxResults (p. 535)
The maximum number of results to select.
Type: Integer
Valid Range: Maximum value of 100.
Required: No

**NameContains (p. 535)**
Filter for fleets containing this name in their fleet device name.
Type: String
Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9\-]+
Required: No

**NextToken (p. 535)**
The response from the last list when returning a list large enough to need tokening.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

**SortBy (p. 535)**
The column to sort by.
Type: String
Valid Values: NAME | CREATION_TIME | LAST_MODIFIED_TIME
Required: No

**SortOrder (p. 535)**
What direction to sort in.
Type: String
Valid Values: Ascending | Descending
Required: No

### Response Syntax

```json
{
  "DeviceFleetSummaries": [
    {
      "CreationTime": number,
      "DeviceFleetArn": "string",
      "DeviceFleetName": "string",
      "LastModifiedTime": number
    }
  ],
  "NextToken": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

DeviceFleetSummaries (p. 536)

Summary of the device fleet.
Type: Array of DeviceFleetSummary (p. 990) objects

NextToken (p. 536)

The response from the last list when returning a list large enough to need tokening.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListDevices
Service: Amazon SageMaker Service

A list of devices.

Request Syntax

```json
{
  "DeviceFleetName": "string",
  "LatestHeartbeatAfter": number,
  "MaxResults": number,
  "ModelName": "string",
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeviceFleetName (p. 538)**

Filter for fleets containing this name in their device fleet name.

Type: String


Pattern: ^[a-zA-Z0-9](-[a-zA-Z0-9]{0,62})$

Required: No

**LatestHeartbeatAfter (p. 538)**

Select fleets where the job was updated after X

Type: Timestamp

Required: No

**MaxResults (p. 538)**

Maximum number of results to select.

Type: Integer

Valid Range: Maximum value of 100.

Required: No

**ModelName (p. 538)**

A filter that searches devices that contains this name in any of their models.

Type: String

Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9]\{0,62}\$

Required: No

**NextToken (p. 538)**

The response from the last list when returning a list large enough to need tokening.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**Response Syntax**

```json
{
    "DeviceSummaries": [
        {
            "AgentVersion": "string",
            "Description": "string",
            "DeviceArn": "string",
            "DeviceFleetName": "string",
            "DeviceName": "string",
            "IotThingName": "string",
            "LatestHeartbeat": number,
            "Models": [
                {
                    "ModelName": "string",
                    "ModelVersion": "string"
                }
            ],
            "RegistrationTime": number
        }
    ],
    "NextToken": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**DeviceSummaries (p. 539)**

Summary of devices.

Type: Array of **DeviceSummary (p. 992)** objects

**NextToken (p. 539)**

The response from the last list when returning a list large enough to need tokening.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListDomains
Service: Amazon SageMaker Service

Lists the domains.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

MaxResults (p. 541)

Returns a list up to a specified limit.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken (p. 541)

If the previous response was truncated, you will receive this token. Use it in your next request to receive the next set of results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

Response Syntax

```
{
  "Domains": [
    {
      "CreationTime": number,
      "DomainArn": "string",
      "DomainId": "string",
      "DomainName": "string",
      "LastModifiedTime": number,
      "Status": "string",
      "Url": "string"
    },
    ...
  ],
  "NextToken": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Domains (p. 541)
The list of domains.

Type: Array of DomainDetails (p. 994) objects

NextToken (p. 541)
If the previous response was truncated, you will receive this token. Use it in your next request to receive the next set of results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListEdgePackagingJobs
Service: Amazon SageMaker Service

Returns a list of edge packaging jobs.

Request Syntax

```json
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "LastModifiedTimeAfter": number,
   "LastModifiedTimeBefore": number,
   "MaxResults": number,
   "ModelNameContains": "string",
   "NameContains": "string",
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string",
   "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 543)**

Select jobs where the job was created after specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 543)**

Select jobs where the job was created before specified time.

Type: Timestamp

Required: No

**LastModifiedTimeAfter (p. 543)**

Select jobs where the job was updated after specified time.

Type: Timestamp

Required: No

**LastModifiedTimeBefore (p. 543)**

Select jobs where the job was updated before specified time.

Type: Timestamp

Required: No

**MaxResults (p. 543)**

Maximum number of results to select.
ListEdgePackagingJobs

- **ModelNameContains (p. 543)**
  - Filter for jobs where the model name contains this string.
  - **Type:** String
  - **Length Constraints:** Maximum length of 63.
  - **Pattern:** [a-zA-Z0-9\-]+
  - **Required:** No

- **NameContains (p. 543)**
  - Filter for jobs containing this name in their packaging job name.
  - **Type:** String
  - **Length Constraints:** Maximum length of 63.
  - **Pattern:** [a-zA-Z0-9\-]+
  - **Required:** No

- **NextToken (p. 543)**
  - The response from the last list when returning a list large enough to need tokening.
  - **Type:** String
  - **Length Constraints:** Maximum length of 8192.
  - **Pattern:** .*
  - **Required:** No

- **SortBy (p. 543)**
  - Use to specify what column to sort by.
  - **Type:** String
  - **Valid Values:** NAME | MODEL_NAME | CREATION_TIME | LAST_MODIFIED_TIME | STATUS
  - **Required:** No

- **SortOrder (p. 543)**
  - What direction to sort by.
  - **Type:** String
  - **Valid Values:** Ascending | Descending
  - **Required:** No

- **StatusEquals (p. 543)**
  - The job status to filter for.
  - **Type:** String
Valid Values: STARTING | INPROGRESS | COMPLETED | FAILED | STOPPING | STOPPED

Required: No

Response Syntax

```
{
  "EdgePackagingJobSummaries": [
    {
      "CompilationJobName": "string",
      "CreationTime": number,
      "EdgePackagingJobArn": "string",
      "EdgePackagingJobName": "string",
      "EdgePackagingJobStatus": "string",
      "LastModifiedTime": number,
      "ModelName": "string",
      "ModelVersion": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**EdgePackagingJobSummaries (p. 545)**

Summaries of edge packaging jobs.

Type: Array of [EdgePackagingJobSummary (p. 1010)] objects

**NextToken (p. 545)**

Token to use when calling the next page of results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListEndpointConfigs
Service: Amazon SageMaker Service
Lists endpoint configurations.

Request Syntax

```json
{
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "MaxResults": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 547)**

A filter that returns only endpoint configurations with a creation time greater than or equal to the specified time (timestamp).

Type: Timestamp

Required: No

**CreationTimeBefore (p. 547)**

A filter that returns only endpoint configurations created before the specified time (timestamp).

Type: Timestamp

Required: No

**MaxResults (p. 547)**

The maximum number of training jobs to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NameContains (p. 547)**

A string in the endpoint configuration name. This filter returns only endpoint configurations whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9-]+  
Required: No

**NextToken (p. 547)**

If the result of the previous ListEndpointConfig request was truncated, the response includes a NextToken. To retrieve the next set of endpoint configurations, use the token in the next request.

Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*  
Required: No

**SortBy (p. 547)**

The field to sort results by. The default is CreationTime.

Type: String
Valid Values: Name | CreationTime  
Required: No

**SortOrder (p. 547)**

The sort order for results. The default is Descending.

Type: String
Valid Values: Ascending | Descending  
Required: No

**Response Syntax**

```
{
    "EndpointConfigs": [
    {
    "CreationTime": number,
    "EndpointConfigArn": "string",
    "EndpointConfigName": "string"
    }
    ],
    "NextToken": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**EndpointConfigs (p. 548)**

An array of endpoint configurations.

Type: Array of EndpointConfigSummary (p. 1016) objects
NextToken (p. 548)

If the response is truncated, SageMaker returns this token. To retrieve the next set of endpoint configurations, use it in the subsequent request

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListEndpoints
Service: Amazon SageMaker Service
Lists endpoints.

Request Syntax

```json
{
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "LastModifiedTimeAfter": number,
  "LastModifiedTimeBefore": number,
  "MaxResults": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 550)**

A filter that returns only endpoints with a creation time greater than or equal to the specified time (timestamp).

Type: Timestamp

Required: No

**CreationTimeBefore (p. 550)**

A filter that returns only endpoints that were created before the specified time (timestamp).

Type: Timestamp

Required: No

**LastModifiedTimeAfter (p. 550)**

A filter that returns only endpoints that were modified after the specified timestamp.

Type: Timestamp

Required: No

**LastModifiedTimeBefore (p. 550)**

A filter that returns only endpoints that were modified before the specified timestamp.

Type: Timestamp

Required: No

**MaxResults (p. 550)**

The maximum number of endpoints to return in the response. This value defaults to 10.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**NameContains (p. 550)**
A string in endpoint names. This filter returns only endpoints whose name contains the specified string.
Type: String
Length Constraints: Maximum length of 63.
Pattern: \[a-zA-Z0-9-]+\]
Required: No

**NextToken (p. 550)**
If the result of a ListEndpoints request was truncated, the response includes a NextToken. To retrieve the next set of endpoints, use the token in the next request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

**SortBy (p. 550)**
Sorts the list of results. The default is CreationTime.
Type: String
Valid Values: Name | CreationTime | Status
Required: No

**SortOrder (p. 550)**
The sort order for results. The default is Descending.
Type: String
Valid Values: Ascending | Descending
Required: No

**StatusEquals (p. 550)**
A filter that returns only endpoints with the specified status.
Type: String
Valid Values: OutOfService | Creating | Updating | SystemUpdating | RollingBack | InService | Deleting | Failed
Required: No

**Response Syntax**

```
Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**Endpoints (p. 551)**

An array or endpoint objects.

Type: Array of [EndpointSummary](p. 1023) objects

**NextToken (p. 551)**

If the response is truncated, SageMaker returns this token. To retrieve the next set of training jobs, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListExperiments
Service: Amazon SageMaker Service

Lists all the experiments in your account. The list can be filtered to show only experiments that were created in a specific time range. The list can be sorted by experiment name or creation time.

Request Syntax

```json
{
  "CreatedAfter": number,
  "CreatedBefore": number,
  "MaxResults": number,
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreatedAfter (p. 553)**

A filter that returns only experiments created after the specified time.

- Type: Timestamp
- Required: No

**CreatedBefore (p. 553)**

A filter that returns only experiments created before the specified time.

- Type: Timestamp
- Required: No

**MaxResults (p. 553)**

The maximum number of experiments to return in the response. The default value is 10.

- Type: Integer
- Valid Range: Minimum value of 1. Maximum value of 100.
- Required: No

**NextToken (p. 553)**

If the previous call to ListExperiments didn't return the full set of experiments, the call returns a token for getting the next set of experiments.

- Type: String
- Length Constraints: Maximum length of 8192.
- Pattern: . *
ListExperiments

Required: No

**SortBy (p. 553)**

The property used to sort results. The default value is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No

**SortOrder (p. 553)**

The sort order. The default value is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

### Response Syntax

```json
{
   "ExperimentSummaries": [
      {
         "CreationTime": number,
         "DisplayName": "string",
         "ExperimentArn": "string",
         "ExperimentName": "string",
         "ExperimentSource": {
            "SourceArn": "string",
            "SourceType": "string"
         },
         "LastModifiedTime": number
      }
   ],
   "NextToken": "string"
}
```

### Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**ExperimentSummaries (p. 554)**

A list of the summaries of your experiments.

Type: Array of ExperimentSummary (p. 1032) objects

**NextToken (p. 554)**

A token for getting the next set of experiments, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListFeatureGroups
Service: Amazon SageMaker Service
List FeatureGroups based on given filter and order.

Request Syntax

```json
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "FeatureGroupStatusEquals": "string",
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "OfflineStoreStatusEquals": "string",
    "SortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 556)**

Use this parameter to search for FeatureGroups created after a specific date and time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 556)**

Use this parameter to search for FeatureGroups created before a specific date and time.

Type: Timestamp

Required: No

**FeatureGroupStatusEquals (p. 556)**

A FeatureGroup status. Filters by FeatureGroup status.

Type: String

Valid Values: Creating | Created | CreateFailed | Deleting | DeleteFailed

Required: No

**MaxResults (p. 556)**

The maximum number of results returned by ListFeatureGroups.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No
NameContains (p. 556)
A string that partially matches one or more FeatureGroups names. Filters FeatureGroups by name.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 64.
Required: No

NextToken (p. 556)
A token to resume pagination of ListFeatureGroups results.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

OfflineStoreStatusEquals (p. 556)
An OfflineStore status. Filters by OfflineStore status.
Type: String
Valid Values: Active | Blocked | Disabled
Required: No

SortBy (p. 556)
The value on which the feature group list is sorted.
Type: String
Valid Values: Name | FeatureGroupStatus | OfflineStoreStatus | CreationTime
Required: No

SortOrder (p. 556)
The order in which feature groups are listed.
Type: String
Valid Values: Ascending | Descending
Required: No

Response Syntax

```json
{
  "FeatureGroupSummaries": [
    {
      "CreationTime": number,
      "FeatureGroupArn": "string",
      "FeatureGroupName": "string",
      "FeatureGroupStatus": "string",
      "OfflineStoreStatus": {}
    }
  ]
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**FeatureGroupSummaries (p. 557)**

A summary of feature groups.

Type: Array of FeatureGroupSummary (p. 1040) objects

**NextToken (p. 557)**

A token to resume pagination of ListFeatureGroups results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListFlowDefinitions

Service: Amazon SageMaker Service

Returns information about the flow definitions in your account.

Request Syntax

```json
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "MaxResults": number,
   "NextToken": "string",
   "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 559)**

A filter that returns only flow definitions with a creation time greater than or equal to the specified timestamp.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 559)**

A filter that returns only flow definitions that were created before the specified timestamp.

Type: Timestamp

Required: No

**MaxResults (p. 559)**

The total number of items to return. If the total number of available items is more than the value specified in MaxResults, then a NextToken will be provided in the output that you can use to resume pagination.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NextToken (p. 559)**

A token to resume pagination.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*
Required: No

*SortOrder (p. 559)*

An optional value that specifies whether you want the results sorted in *Ascending* or *Descending* order.

Type: String

Valid Values: *Ascending* | *Descending*

Required: No

**Response Syntax**

```json
{
    "FlowDefinitionSummaries": [
        {
            "CreationTime": "number",
            "FailureReason": "string",
            "FlowDefinitionArn": "string",
            "FlowDefinitionName": "string",
            "FlowDefinitionStatus": "string"
        }
    ],
    "NextToken": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

*FlowDefinitionSummaries (p. 560)*

An array of objects describing the flow definitions.

Type: Array of *FlowDefinitionSummary (p. 1052)* objects

*NextToken (p. 560)*

A token to resume pagination.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListHumanTaskUis
Service: Amazon SageMaker Service

Returns information about the human task user interfaces in your account.

Request Syntax

```
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "MaxResults": number,
    "NextToken": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 562)

A filter that returns only human task user interfaces with a creation time greater than or equal to the specified timestamp.

Type: Timestamp

Required: No

CreationTimeBefore (p. 562)

A filter that returns only human task user interfaces that were created before the specified timestamp.

Type: Timestamp

Required: No

MaxResults (p. 562)

The total number of items to return. If the total number of available items is more than the value specified in MaxResults, then a NextToken will be provided in the output that you can use to resume pagination.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken (p. 562)

A token to resume pagination.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*
Required: No

**SortOrder (p. 562)**

An optional value that specifies whether you want the results sorted in **Ascending** or **Descending** order.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```json
{
  "HumanTaskUiSummaries": [
    {
      "CreationTime": number,
      "HumanTaskUiArn": "string",
      "HumanTaskUiName": "string"
    }
  ],
  "NextToken": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**HumanTaskUiSummaries (p. 563)**

An array of objects describing the human task user interfaces.

Type: Array of **HumanTaskUiSummary (p. 1077)** objects

**NextToken (p. 563)**

A token to resume pagination.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**Errors**

For information about the errors that are common to all actions, see **Common Errors (p. 1472)**.

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
ListHyperParameterTuningJobs

Service: Amazon SageMaker Service

Gets a list of HyperParameterTuningJobSummary (p. 1092) objects that describe the hyperparameter tuning jobs launched in your account.

Request Syntax

```json
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "LastModifiedTimeAfter": number,
    "LastModifiedTimeBefore": number,
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string",
    "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 565)

A filter that returns only tuning jobs that were created after the specified time.

Type: Timestamp

Required: No

CreationTimeBefore (p. 565)

A filter that returns only tuning jobs that were created before the specified time.

Type: Timestamp

Required: No

LastModifiedTimeAfter (p. 565)

A filter that returns only tuning jobs that were modified after the specified time.

Type: Timestamp

Required: No

LastModifiedTimeBefore (p. 565)

A filter that returns only tuning jobs that were modified before the specified time.

Type: Timestamp

Required: No

MaxResults (p. 565)

The maximum number of tuning jobs to return. The default value is 10.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

NameContains (p. 565)
A string in the tuning job name. This filter returns only tuning jobs whose name contains the specified string.
Type: String
Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9\-]+
Required: No

NextToken (p. 565)
If the result of the previous ListHyperParameterTuningJobs request was truncated, the response includes a NextToken. To retrieve the next set of tuning jobs, use the token in the next request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

SortBy (p. 565)
The field to sort results by. The default is Name.
Type: String
Valid Values: Name | Status | CreationTime
Required: No

SortOrder (p. 565)
The sort order for results. The default is Ascending.
Type: String
Valid Values: Ascending | Descending
Required: No

StatusEquals (p. 565)
A filter that returns only tuning jobs with the specified status.
Type: String
Valid Values: Completed | InProgress | Failed | Stopped | Stopping
Required: No

Response Syntax
```json
{
```
"HyperParameterTuningJobSummaries": [  
  {  
    "CreationTime": number,  
    "HyperParameterTuningEndTime": number,  
    "HyperParameterTuningJobArn": "string",  
    "HyperParameterTuningJobName": "string",  
    "HyperParameterTuningJobStatus": "string",  
    "LastModifiedTime": number,  
    "ObjectiveStatusCounters": {  
      "Failed": number,  
      "Pending": number,  
      "Succeeded": number  
    },  
    "ResourceLimits": {  
      "MaxNumberOfTrainingJobs": number,  
      "MaxParallelTrainingJobs": number  
    },  
    "Strategy": "string",  
    "TrainingJobStatusCounters": {  
      "Completed": number,  
      "InProgress": number,  
      "NonRetryableError": number,  
      "RetryableError": number,  
      "Stopped": number  
    }  
  },  
],  
"NextToken": "string"

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

HyperParameterTuningJobSummaries (p. 566)

A list of HyperParameterTuningJobSummary (p. 1092) objects that describe the tuning jobs that the ListHyperParameterTuningJobs request returned.

Type: Array of HyperParameterTuningJobSummary (p. 1092) objects

NextToken (p. 566)

If the result of this ListHyperParameterTuningJobs request was truncated, the response includes a NextToken. To retrieve the next set of tuning jobs, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
ListImages
Service: Amazon SageMaker Service

Lists the images in your account and their properties. The list can be filtered by creation time or modified time, and whether the image name contains a specified string.

Request Syntax

```
{
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "LastModifiedTimeAfter": number,
  "LastModifiedTimeBefore": number,
  "MaxResults": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 569)**

A filter that returns only images created on or after the specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 569)**

A filter that returns only images created on or before the specified time.

Type: Timestamp

Required: No

**LastModifiedTimeAfter (p. 569)**

A filter that returns only images modified on or after the specified time.

Type: Timestamp

Required: No

**LastModifiedTimeBefore (p. 569)**

A filter that returns only images modified on or before the specified time.

Type: Timestamp

Required: No

**MaxResults (p. 569)**

The maximum number of images to return in the response. The default value is 10.
Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NameContains (p. 569)**

A filter that returns only images whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-\.]+$

Required: No

**NextToken (p. 569)**

If the previous call to `ListImages` didn't return the full set of images, the call returns a token for getting the next set of images.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 569)**

The property used to sort results. The default value is `CREATION_TIME`.

Type: String

Valid Values: `CREATION_TIME` | `LAST_MODIFIED_TIME` | `IMAGE_NAME`

Required: No

**SortOrder (p. 569)**

The sort order. The default value is `DESCENDING`.

Type: String

Valid Values: `ASCENDING` | `DESCENDING`

Required: No

**Response Syntax**

```
{
  "Images": [ 
    { 
      "CreationTime": number, 
      "Description": "string", 
      "DisplayName": "string", 
      "FailureReason": "string", 
      "ImageArn": "string", 
      "ImageName": "string", 
      "ImageStatus": "string", 
    }, 
  ]
}
```
### Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**Images (p. 570)**

A list of images and their properties.

- Type: Array of [Image (p. 1096) objects](#)

**NextToken (p. 570)**

A token for getting the next set of images, if there are any.

- Type: String
  - Length Constraints: Maximum length of 8192.
  - Pattern: .*

### Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

```json
"LastModifiedTime": number
}
],
"NextToken": "string"
}
```
ListImageVersions
Service: Amazon SageMaker Service

Lists the versions of a specified image and their properties. The list can be filtered by creation time or modified time.

Request Syntax

```json
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "ImageName":  "string",
   "LastModifiedTimeAfter": number,
   "LastModifiedTimeBefore": number,
   "MaxResults": number,
   "NextToken":  "string",
   "SortBy": "string",
   "SortOrder":  "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 572)**

A filter that returns only versions created on or after the specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 572)**

A filter that returns only versions created on or before the specified time.

Type: Timestamp

Required: No

**ImageName (p. 572)**

The name of the image to list the versions of.

Type: String


Pattern: `^[a-zA-Z0-9](\-[a-zA-Z0-9])\{0,62}$`

Required: Yes

**LastModifiedTimeAfter (p. 572)**

A filter that returns only versions modified on or after the specified time.

Type: Timestamp

Required: No
LastModifiedTimeBefore (p. 572)
A filter that returns only versions modified on or before the specified time.
Type: Timestamp
Required: No

MaxResults (p. 572)
The maximum number of versions to return in the response. The default value is 10.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

NextToken (p. 572)
If the previous call to ListImageVersions didn't return the full set of versions, the call returns a
token for getting the next set of versions.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

SortBy (p. 572)
The property used to sort results. The default value is CREATION_TIME.
Type: String
Valid Values: CREATION_TIME | LAST_MODIFIED_TIME | VERSION
Required: No

SortOrder (p. 572)
The sort order. The default value is DESCENDING.
Type: String
Valid Values: ASCENDING | DESCENDING
Required: No

Response Syntax

{
  "ImageVersions": [
    {
      "CreationTime": number,
      "FailureReason": "string",
      "ImageArn": "string",
      "ImageVersionArn": "string",
      "ImageVersionStatus": "string",
      "LastModifiedTime": number,
      "Version": number
    }
  ]
}
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ImageVersions (p. 573)**

A list of versions and their properties.

Type: Array of ImageVersion (p. 1099) objects

**NextToken (p. 573)**

A token for getting the next set of versions, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFoundException**

Resource being accessed is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListInferenceRecommendationsJobs
Service: Amazon SageMaker Service
Lists recommendation jobs that satisfy various filters.

Request Syntax

```
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "LastModifiedTimeAfter": number,
    "LastModifiedTimeBefore": number,
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string",
    "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 575)
A filter that returns only jobs created after the specified time (timestamp).
Type: Timestamp
Required: No

CreationTimeBefore (p. 575)
A filter that returns only jobs created before the specified time (timestamp).
Type: Timestamp
Required: No

LastModifiedTimeAfter (p. 575)
A filter that returns only jobs that were last modified after the specified time (timestamp).
Type: Timestamp
Required: No

LastModifiedTimeBefore (p. 575)
A filter that returns only jobs that were last modified before the specified time (timestamp).
Type: Timestamp
Required: No

MaxResults (p. 575)
The maximum number of recommendations to return in the response.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NameContains (p. 575)**

A string in the job name. This filter returns only recommendations whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: [a-zA-Z0-9\-]+

Required: No

**NextToken (p. 575)**

If the response to a previous `ListInferenceRecommendationsJobsRequest` request was truncated, the response includes a NextToken. To retrieve the next set of recommendations, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 575)**

The parameter by which to sort the results.

Type: String

Valid Values: Name | CreationTime | Status

Required: No

**SortOrder (p. 575)**

The sort order for the results.

Type: String

Valid Values: Ascending | Descending

Required: No

**StatusEquals (p. 575)**

A filter that retrieves only inference recommendations jobs with a specific status.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | FAILED | STOPPING | STOPPED

Required: No

**Response Syntax**

```json
{
```
"InferenceRecommendationsJobs": [ 
  
  "CompletionTime": number,
  "CreationTime": number,
  "FailureReason": "string",
  "JobArn": "string",
  "JobDescription": "string",
  "JobName": "string",
  "JobType": "string",
  "LastModifiedTime": number,
  "RoleArn": "string",
  "Status": "string"
  
  
  ],
  "NextToken": "string"
}
ListLabelingJobs
Service: Amazon SageMaker Service

Gets a list of labeling jobs.

Request Syntax

```
{
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "LastModifiedTimeAfter": number,
  "LastModifiedTimeBefore": number,
  "MaxResults": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 578)**

A filter that returns only labeling jobs created after the specified time (timestamp).

- Type: Timestamp
- Required: No

**CreationTimeBefore (p. 578)**

A filter that returns only labeling jobs created before the specified time (timestamp).

- Type: Timestamp
- Required: No

**LastModifiedTimeAfter (p. 578)**

A filter that returns only labeling jobs modified after the specified time (timestamp).

- Type: Timestamp
- Required: No

**LastModifiedTimeBefore (p. 578)**

A filter that returns only labeling jobs modified before the specified time (timestamp).

- Type: Timestamp
- Required: No

**MaxResults (p. 578)**

The maximum number of labeling jobs to return in each page of the response.
**ListLabelingJobs**

**Type:** Integer

**Valid Range:** Minimum value of 1. Maximum value of 100.

**Required:** No

**NameContains (p. 578)**

A string in the labeling job name. This filter returns only labeling jobs whose name contains the specified string.

**Type:** String

**Length Constraints:** Maximum length of 63.

**Pattern:** [a-zA-Z0-9\-]+

**Required:** No

**NextToken (p. 578)**

If the result of the previous `ListLabelingJobs` request was truncated, the response includes a `NextToken`. To retrieve the next set of labeling jobs, use the token in the next request.

**Type:** String

**Length Constraints:** Maximum length of 8192.

**Pattern:** .*

**Required:** No

**SortBy (p. 578)**

The field to sort results by. The default is `CreationTime`.

**Type:** String

**Valid Values:** Name | CreationTime | Status

**Required:** No

**SortOrder (p. 578)**

The sort order for results. The default is `Ascending`.

**Type:** String

**Valid Values:** Ascending | Descending

**Required:** No

**StatusEquals (p. 578)**

A filter that retrieves only labeling jobs with a specific status.

**Type:** String

**Valid Values:** Initializing | InProgress | Completed | Failed | Stopping | Stopped

**Required:** No

579
Response Syntax

```json
{
  "LabelingJobSummaryList": [
    {
      "AnnotationConsolidationLambdaArn": "string",
      "CreationTime": number,
      "FailureReason": "string",
      "InputConfig": {
        "DataAttributes": {
          "ContentClassifiers": [ "string" ]
        },
        "DataSource": {
          "S3DataSource": {
            "ManifestS3Uri": "string"
          },
          "SnsDataSource": {
            "SnsTopicArn": "string"
          }
        }
      },
      "LabelCounters": {
        "FailedNonRetryableError": number,
        "HumanLabeled": number,
        "MachineLabeled": number,
        "TotalLabeled": number,
        "Unlabeled": number
      },
      "LabelingJobArn": "string",
      "LabelingJobName": "string",
      "LabelingJobOutput": {
        "FinalActiveLearningModelArn": "string",
        "OutputDatasetS3Uri": "string"
      },
      "LabelingJobStatus": "string",
      "LastModifiedTime": number,
      "PreHumanTaskLambdaArn": "string",
      "WorkteamArn": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**LabelingJobSummaryList (p. 580)**

An array of `LabelingJobSummary` objects, each describing a labeling job.

Type: Array of `LabelingJobSummary` (p. 1136) objects

**NextToken (p. 580)**

If the response is truncated, SageMaker returns this token. To retrieve the next set of labeling jobs, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.
Pattern: . *

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListLabelingJobsForWorkteam

Service: Amazon SageMaker Service

Gets a list of labeling jobs assigned to a specified work team.

Request Syntax

```json
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "JobReferenceCodeContains": "string",
   "MaxResults": number,
   "NextToken": string,
   "SortBy": "string",
   "SortOrder": "string",
   "WorkteamArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 582)**

A filter that returns only labeling jobs created after the specified time (timestamp).

Type: Timestamp

Required: No

**CreationTimeBefore (p. 582)**

A filter that returns only labeling jobs created before the specified time (timestamp).

Type: Timestamp

Required: No

**JobReferenceCodeContains (p. 582)**

A filter the limits jobs to only the ones whose job reference code contains the specified string.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .+

Required: No

**MaxResults (p. 582)**

The maximum number of labeling jobs to return in each page of the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No
NextToken (p. 582)
If the result of the previous ListLabelingJobsForWorkteam request was truncated, the response includes a NextToken. To retrieve the next set of labeling jobs, use the token in the next request.

Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

SortBy (p. 582)
The field to sort results by. The default is CreationTime.

Type: String
Valid Values: CreationTime
Required: No

SortOrder (p. 582)
The sort order for results. The default is Ascending.

Type: String
Valid Values: Ascending | Descending
Required: No

WorkteamArn (p. 582)
The Amazon Resource Name (ARN) of the work team for which you want to see labeling jobs for.

Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workteam/.*
Required: Yes

Response Syntax

```json
{
  "LabelingJobSummaryList": [
    {
      "CreationTime": number,
      "JobReferenceCode": "string",
      "LabelCounters": {
        "HumanLabeled": number,
        "PendingHuman": number,
        "Total": number
      },
      "LabelingJobName": "string",
      "NumberOfHumanWorkersPerDataObject": number,
      "WorkRequesterAccountId": "string"
    }
  ],
  "NextToken": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**LabelingJobSummaryList (p. 583)**

An array of `LabelingJobSummary` objects, each describing a labeling job.

Type: Array of `LabelingJobForWorkteamSummary (p. 1126)` objects

**NextToken (p. 583)**

If the response is truncated, SageMaker returns this token. To retrieve the next set of labeling jobs, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListLineageGroups
Service: Amazon SageMaker Service

A list of lineage groups shared with your AWS account. For more information, see Cross-Account Lineage Tracking in the Amazon SageMaker Developer Guide.

Request Syntax

```
{
   "CreatedAfter": number,
   "CreatedBefore": number,
   "MaxResults": number,
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreatedAfter (p. 585)

A timestamp to filter against lineage groups created after a certain point in time.

Type: Timestamp

Required: No

CreatedBefore (p. 585)

A timestamp to filter against lineage groups created before a certain point in time.

Type: Timestamp

Required: No

MaxResults (p. 585)

The maximum number of endpoints to return in the response. This value defaults to 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken (p. 585)

If the response is truncated, SageMaker returns this token. To retrieve the next set of algorithms, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No
SortBy (p. 585)

The parameter by which to sort the results. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No

SortOrder (p. 585)

The sort order for the results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No

Response Syntax

```
{
  "LineageGroupSummaries": [
    {
      "CreationTime": number,
      "DisplayName": "string",
      "LastModifiedTime": number,
      "LineageGroupArn": "string",
      "LineageGroupName": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

LineageGroupSummaries (p. 586)

A list of lineage groups and their properties.

Type: Array of LineageGroupSummary (p. 1140) objects

NextToken (p. 586)

If the response is truncated, SageMaker returns this token. To retrieve the next set of algorithms, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListModelBiasJobDefinitions

Service: Amazon SageMaker Service

Lists model bias jobs definitions that satisfy various filters.

Request Syntax

```
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "EndpointName": "string",
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 588)**

A filter that returns only model bias jobs created after a specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 588)**

A filter that returns only model bias jobs created before a specified time.

Type: Timestamp

Required: No

**EndpointName (p. 588)**

Name of the endpoint to monitor for model bias.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9]{0,62}

Required: No

**MaxResults (p. 588)**

The maximum number of model bias jobs to return in the response. The default value is 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No
NameContains (p. 588)
Filter for model bias jobs whose name contains a specified string.
Type: String
Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9-]+
Required: No

NextToken (p. 588)
The token returned if the response is truncated. To retrieve the next set of job executions, use it in the next request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

SortBy (p. 588)
Whether to sort results by the Name or CreationTime field. The default is CreationTime.
Type: String
Valid Values: Name | CreationTime
Required: No

SortOrder (p. 588)
Whether to sort the results in Ascending or Descending order. The default is Descending.
Type: String
Valid Values: Ascending | Descending
Required: No

Response Syntax
```
{
  "JobDefinitionSummaries": [
    {
      "CreationTime": number,
      "EndpointName": "string",
      "MonitoringJobDefinitionArn": "string",
      "MonitoringJobDefinitionName": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements
If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**JobDefinitionSummaries (p. 589)**

A JSON array in which each element is a summary for a model bias jobs.

Type: Array of [MonitoringJobDefinitionSummary (p. 1205)] objects

**NextToken (p. 589)**

If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of jobs, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListModelExplainabilityJobDefinitions

Service: Amazon SageMaker Service

Lists model explainability job definitions that satisfy various filters.

Request Syntax

```json
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "EndpointName": "string",
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "sortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 591)**

A filter that returns only model explainability jobs created after a specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 591)**

A filter that returns only model explainability jobs created before a specified time.

Type: Timestamp

Required: No

**EndpointName (p. 591)**

Name of the endpoint to monitor for model explainability.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\$

Required: No

**MaxResults (p. 591)**

The maximum number of jobs to return in the response. The default value is 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No
NameContains (p. 591)

Filter for model explainability jobs whose name contains a specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: \[a-zA-Z0-9\-\]+

Required: No

NextToken (p. 591)

The token returned if the response is truncated. To retrieve the next set of job executions, use it in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

SortBy (p. 591)

Whether to sort results by the Name or CreationTime field. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No

SortOrder (p. 591)

Whether to sort the results in Ascending or Descending order. The default is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

Response Syntax

```
{
   "JobDefinitionSummaries": [
   {
      "CreationTime": number,
      "EndpointName": "string",
      "MonitoringJobDefinitionArn": "string",
      "MonitoringJobDefinitionName": "string"
   }
   ],
   "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**JobDefinitionSummaries (p. 592)**

A JSON array in which each element is a summary for an explainability bias job.

Type: Array of `MonitoringJobDefinitionSummary (p. 1205)` objects

**NextToken (p. 592)**

If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of jobs, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListModelMetadata

Service: Amazon SageMaker Service

Lists the domain, framework, task, and model name of standard machine learning models found in common model zoos.

Request Syntax

```json
{
    "MaxResults": number,
    "NextToken": "string",
    "SearchExpression": {
        "Filters": [
            {
                "Name": "string",
                "Value": "string"
            }
        ]
    }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

MaxResults (p. 594)

The maximum number of models to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken (p. 594)

If the response to a previous ListModelMetadataResponse request was truncated, the response includes a NextToken. To retrieve the next set of model metadata, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

SearchExpression (p. 594)

One or more filters that searches for the specified resource or resources in a search. All resource objects that satisfy the expression's condition are included in the search results. Specify the Framework, FrameworkVersion, Domain or Task to filter supported. Filter names and values are case-sensitive.

Type: ModelMetadataSearchExpression (p. 1165) object

Required: No
Response Syntax

```
{
  "ModelMetadataSummaries": [
    {
      "Domain": "string",
      "Framework": "string",
      "FrameworkVersion": "string",
      "Model": "string",
      "Task": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ModelMetadataSummaries (p. 595)**

A structure that holds model metadata.

Type: Array of ModelMetadataSummary (p. 1166) objects

**NextToken (p. 595)**

A token for getting the next set of recommendations, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListModelPackageGroups
Service: Amazon SageMaker Service

Gets a list of the model groups in your AWS account.

Request Syntax

```json
{
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "MaxResults": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 596)**

A filter that returns only model groups created after the specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 596)**

A filter that returns only model groups created before the specified time.

Type: Timestamp

Required: No

**MaxResults (p. 596)**

The maximum number of results to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NameContains (p. 596)**

A string in the model group name. This filter returns only model groups whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: [a-zA-Z0-9\-]+
**NextToken (p. 596)**

If the result of the previous `ListModelPackageGroups` request was truncated, the response includes a `NextToken`. To retrieve the next set of model groups, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 596)**

The field to sort results by. The default is `CreationTime`.

Type: String

Valid Values: Name | CreationTime

Required: No

**SortOrder (p. 596)**

The sort order for results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```
{
  "ModelPackageGroupSummaryList": [
    {
      "CreationTime": number,
      "ModelPackageGroupArn": "string",
      "ModelPackageGroupDescription": "string",
      "ModelPackageGroupName": "string",
      "ModelPackageGroupStatus": "string"
    }
  ],
  "NextToken": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ModelPackageGroupSummaryList (p. 597)**

A list of summaries of the model groups in your AWS account.

Type: Array of `ModelPackageGroupSummary (p. 1178)` objects
NextToken (p. 597)

If the response is truncated, SageMaker returns this token. To retrieve the next set of model groups, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListModelPackages
Service: Amazon SageMaker Service

Lists the model packages that have been created.

Request Syntax

```json
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "MaxResults": number,
    "ModelApprovalStatus": "string",
    "ModelPackageGroupName": "string",
    "ModelPackageType": "string",
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 599)**

A filter that returns only model packages created after the specified time (timestamp).

Type: Timestamp

Required: No

**CreationTimeBefore (p. 599)**

A filter that returns only model packages created before the specified time (timestamp).

Type: Timestamp

Required: No

**MaxResults (p. 599)**

The maximum number of model packages to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**ModelApprovalStatus (p. 599)**

A filter that returns only the model packages with the specified approval status.

Type: String

Valid Values: Approved | Rejected | PendingManualApproval

Required: No
**ModelPackageName (p. 599)**

A filter that returns only model versions that belong to the specified model group.

Type: String


Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:[a-z\-]*\/)?(a-zA-Z0-9)([a-zA-Z0-9-]{0,62})(?<!-)

Required: No

**ModelPackageType (p. 599)**

A filter that returns only the model packages of the specified type. This can be one of the following values.

- UNVERSIONED - List only unversioned models. This is the default value if no ModelPackageType is specified.
- VERSIONED - List only versioned models.
- BOTH - List both versioned and unversioned models.

Type: String

Valid Values: Versioned | Unversioned | Both

Required: No

**NameContains (p. 599)**

A string in the model package name. This filter returns only model packages whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: [a-zA-Z0-9-]+

Required: No

**NextToken (p. 599)**

If the response to a previous ListModelPackages request was truncated, the response includes a NextToken. To retrieve the next set of model packages, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 599)**

The parameter by which to sort the results. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No
SortOrder (p. 599)
The sort order for the results. The default is Ascending.
Type: String
Valid Values: Ascending | Descending
Required: No

Response Syntax

```json
{
    "ModelPackageSummaryList": [
        {
            "CreationTime": number,
            "ModelApprovalStatus": "string",
            "ModelPackageArn": "string",
            "ModelPackageDescription": "string",
            "ModelPackageGroupName": "string",
            "ModelPackageName": "string",
            "ModelPackageStatus": "string",
            "ModelPackageVersion": number
        }
    ],
    "NextToken": "string"
}
```

Response Elements
If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

ModelPackageSummaryList (p. 601)
An array of ModelPackageSummary objects, each of which lists a model package.
Type: Array of ModelPackageSummary (p. 1182) objects

NextToken (p. 601)
If the response is truncated, SageMaker returns this token. To retrieve the next set of model packages, use it in the subsequent request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
ListModelQualityJobDefinitions
Service: Amazon SageMaker Service

Gets a list of model quality monitoring job definitions in your account.

Request Syntax

```
{
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "EndpointName": "string",
  "MaxResults": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 603)**

A filter that returns only model quality monitoring job definitions created after the specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 603)**

A filter that returns only model quality monitoring job definitions created before the specified time.

Type: Timestamp

Required: No

**EndpointName (p. 603)**

A filter that returns only model quality monitoring job definitions that are associated with the specified endpoint.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9-]*[a-zA-Z0-9]{0,62}$

Required: No

**MaxResults (p. 603)**

The maximum number of results to return in a call to ListModelQualityJobDefinitions.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**NameContains (p. 603)**

A string in the transform job name. This filter returns only model quality monitoring job definitions whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: [a-zA-Z0-9\-]+

Required: No

**NextToken (p. 603)**

If the result of the previous ListModelQualityJobDefinitions request was truncated, the response includes a NextToken. To retrieve the next set of model quality monitoring job definitions, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 603)**

The field to sort results by. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No

**SortOrder (p. 603)**

The sort order for results. The default is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```json
{
  "JobDefinitionSummaries": [
    {
      "CreationTime": number,
      "EndpointName": "string",
      "MonitoringJobDefinitionArn": "string",
      "MonitoringJobDefinitionName": "string"
    }
  ],
  "NextToken": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**JobDefinitionSummaries (p. 604)**

A list of summaries of model quality monitoring job definitions.

Type: Array of MonitoringJobDefinitionSummary (p. 1205) objects

**NextToken (p. 604)**

If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of model quality monitoring job definitions, use it in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListModels
Service: Amazon SageMaker Service
Lists models created with the CreateModel API.

Request Syntax
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string"
}

Request Parameters
For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 606)
A filter that returns only models with a creation time greater than or equal to the specified time (timestamp).
Type: Timestamp
Required: No

CreationTimeBefore (p. 606)
A filter that returns only models created before the specified time (timestamp).
Type: Timestamp
Required: No

MaxResults (p. 606)
The maximum number of models to return in the response.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

NameContains (p. 606)
A string in the model name. This filter returns only models whose name contains the specified string.
Type: String
Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9-]+
Required: No
NextToken (p. 606)

If the response to a previous ListModels request was truncated, the response includes a NextToken. To retrieve the next set of models, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

SortBy (p. 606)

Sorts the list of results. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No

SortOrder (p. 606)

The sort order for results. The default is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

Response Syntax

```json
{
  "Models": [
    {
      "CreationTime": number,
      "ModelArn": "string",
      "ModelName": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Models (p. 607)

An array of ModelSummary objects, each of which lists a model.

Type: Array of ModelSummary (p. 1192) objects

NextToken (p. 607)

If the response is truncated, SageMaker returns this token. To retrieve the next set of models, use it in the subsequent request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListMonitoringExecutions
Service: Amazon SageMaker Service
Returns list of all monitoring job executions.

Request Syntax

```
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "EndpointName": "string",
    "LastModifiedTimeAfter": number,
    "LastModifiedTimeBefore": number,
    "MaxResults": number,
    "MonitoringJobDefinitionName": "string",
    "MonitoringScheduleName": "string",
    "MonitoringTypeEquals": "string",
    "NextToken": "string",
    "ScheduledTimeAfter": number,
    "ScheduledTimeBefore": number,
    "SortBy": "string",
    "SortOrder": "string",
    "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 609)**

A filter that returns only jobs created after a specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 609)**

A filter that returns only jobs created before a specified time.

Type: Timestamp

Required: No

**EndpointName (p. 609)**

Name of a specific endpoint to fetch jobs for.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9]([-][a-zA-Z0-9]{0,62})

Required: No

**LastModifiedTimeAfter (p. 609)**

A filter that returns only jobs modified before a specified time.
ListMonitoringExecutions

Type: Timestamp
Required: No

LastModifiedTimeBefore (p. 609)
A filter that returns only jobs modified after a specified time.
Type: Timestamp
Required: No

MaxResults (p. 609)
The maximum number of jobs to return in the response. The default value is 10.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

MonitoringJobDefinitionName (p. 609)
Gets a list of the monitoring job runs of the specified monitoring job definitions.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$
Required: No

MonitoringScheduleName (p. 609)
Name of a specific schedule to fetch jobs for.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$
Required: No

MonitoringTypeEquals (p. 609)
A filter that returns only the monitoring job runs of the specified monitoring type.
Type: String
Valid Values: DataQuality | ModelQuality | ModelBias | ModelExplainability
Required: No

NextToken (p. 609)
The token returned if the response is truncated. To retrieve the next set of job executions, use it in
the next request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
ListMonitoringExecutions

Required: No

ScheduledTimeAfter (p. 609)
Filter for jobs scheduled after a specified time.
Type: Timestamp
Required: No

ScheduledTimeBefore (p. 609)
Filter for jobs scheduled before a specified time.
Type: Timestamp
Required: No

SortBy (p. 609)
Whether to sort results by Status, CreationTime, ScheduledTime field. The default is CreationTime.
Type: String
Valid Values: CreationTime | ScheduledTime | Status
Required: No

SortOrder (p. 609)
Whether to sort the results in Ascending or Descending order. The default is Descending.
Type: String
Valid Values: Ascending | Descending
Required: No

StatusEquals (p. 609)
A filter that retrieves only jobs with a specific status.
Type: String
Valid Values: Pending | Completed | CompletedWithViolations | InProgress | Failed | Stopping | Stopped
Required: No

Response Syntax

```
{
   "MonitoringExecutionSummaries": [ 
   {
      "CreationTime": number,
      "EndpointName": "string",
      "FailureReason": "string",
      "LastModifiedTime": number,
      "MonitoringExecutionStatus": "string",
      "MonitoringJobDefinitionName": "string",
      "MonitoringScheduleName": "string",
      "MonitoringType": "string",
      "ProcessingJobArn": "string",
   }]
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**MonitoringExecutionSummaries (p. 611)**

- A JSON array in which each element is a summary for a monitoring execution.

  Type: Array of [MonitoringExecutionSummary (p. 1199)] objects

**NextToken (p. 611)**

- If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of jobs, use it in the subsequent request.

  Type: String

  Length Constraints: Maximum length of 8192.

  Pattern: .*

Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListMonitoringSchedules
Service: Amazon SageMaker Service

Returns list of all monitoring schedules.

Request Syntax

```
{
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "EndpointName": "string",
  "LastModifiedTimeAfter": number,
  "LastModifiedTimeBefore": number,
  "MaxResults": number,
  "MonitoringJobDefinitionName": "string",
  "MonitoringTypeEquals": "string",
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 613)**

A filter that returns only monitoring schedules created after a specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 613)**

A filter that returns only monitoring schedules created before a specified time.

Type: Timestamp

Required: No

**EndpointName (p. 613)**

Name of a specific endpoint to fetch schedules for.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-\*]*[a-zA-Z0-9]\{0,62\}

Required: No

**LastModifiedTimeAfter (p. 613)**

A filter that returns only monitoring schedules modified after a specified time.
Type: Timestamp
Required: No

**LastModifiedTimeBefore (p. 613)**
A filter that returns only monitoring schedules modified before a specified time.
Type: Timestamp
Required: No

**MaxResults (p. 613)**
The maximum number of jobs to return in the response. The default value is 10.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**MonitoringJobDefinitionName (p. 613)**
Gets a list of the monitoring schedules for the specified monitoring job definition.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$
Required: No

**MonitoringTypeEquals (p. 613)**
A filter that returns only the monitoring schedules for the specified monitoring type.
Type: String
Valid Values: DataQuality | ModelQuality | ModelBias | ModelExplainability
Required: No

**NameContains (p. 613)**
Filter for monitoring schedules whose name contains a specified string.
Type: String
Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9\-]+
Required: No

**NextToken (p. 613)**
The token returned if the response is truncated. To retrieve the next set of job executions, use it in the next request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
**SortBy (p. 613)**

Whether to sort results by Status, CreationTime, ScheduledTime field. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime | Status

Required: No

**SortOrder (p. 613)**

Whether to sort the results in Ascending or Descending order. The default is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

**StatusEquals (p. 613)**

A filter that returns only monitoring schedules modified before a specified time.

Type: String

Valid Values: Pending | Failed | Scheduled | Stopped

Required: No

**Response Syntax**

```json
{
    "MonitoringScheduleSummaries": [
        {
            "CreationTime": number,
            "EndpointName": "string",
            "LastModifiedTime": number,
            "MonitoringJobDefinitionName": "string",
            "MonitoringScheduleArn": "string",
            "MonitoringScheduleName": "string",
            "MonitoringScheduleStatus": "string",
            "MonitoringType": "string"
        }
    ],
    "NextToken": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**MonitoringScheduleSummaries (p. 615)**

A JSON array in which each element is a summary for a monitoring schedule.

Type: Array of MonitoringScheduleSummary (p. 1216) objects
NextToken (p. 615)

If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of jobs, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListNotebookInstanceLifecycleConfigs

Service: Amazon SageMaker Service

Lists notebook instance lifestyle configurations created with the CreateNotebookInstanceLifecycleConfig (p. 150) API.

Request Syntax

```json
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "LastModifiedTimeAfter": number,
    "LastModifiedTimeBefore": number,
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreationTimeAfter (p. 617)

A filter that returns only lifecycle configurations that were created after the specified time (timestamp).

Type: Timestamp

Required: No

CreationTimeBefore (p. 617)

A filter that returns only lifecycle configurations that were created before the specified time (timestamp).

Type: Timestamp

Required: No

LastModifiedTimeAfter (p. 617)

A filter that returns only lifecycle configurations that were modified after the specified time (timestamp).

Type: Timestamp

Required: No

LastModifiedTimeBefore (p. 617)

A filter that returns only lifecycle configurations that were modified before the specified time (timestamp).

Type: Timestamp

Required: No
**MaxResults (p. 617)**

The maximum number of lifecycle configurations to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NameContains (p. 617)**

A string in the lifecycle configuration name. This filter returns only lifecycle configurations whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: [a-zA-Z0-9-]+

Required: No

**NextToken (p. 617)**

If the result of a ListNotebookInstanceLifecycleConfigs request was truncated, the response includes a NextToken. To get the next set of lifecycle configurations, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 617)**

Sorts the list of results. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime | LastModifiedTime

Required: No

**SortOrder (p. 617)**

The sort order for results.

Type: String

Valid Values: Ascending | Descending

Required: No

### Response Syntax

```json
{
    "NextToken": "string",
}
```
"NotebookInstanceLifecycleConfigs": [ 
  { 
    "CreationTime": number,
    "LastModifiedTime": number,
    "NotebookInstanceLifecycleConfigArn": "string",
    "NotebookInstanceLifecycleConfigName": "string"
  }
]

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken (p. 618)

If the response is truncated, SageMaker returns this token. To get the next set of lifecycle configurations, use it in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

NotebookInstanceLifecycleConfigs (p. 618)

An array of NotebookInstanceLifecycleConfiguration objects, each listing a lifecycle configuration.

Type: Array of NotebookInstanceLifecycleConfigSummary (p. 1224) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListNotebookInstances

Service: Amazon SageMaker Service

Returns a list of the SageMaker notebook instances in the requester's account in an AWS Region.

Request Syntax

```json
{
   "AdditionalCodeRepositoryEquals": "string",
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "DefaultCodeRepositoryContains": "string",
   "LastModifiedTimeAfter": number,
   "LastModifiedTimeBefore": number,
   "MaxResults": number,
   "NameContains": "string",
   "NextToken": "string",
   "NotebookInstanceLifecycleConfigNameContains": "string",
   "SortBy": "string",
   "SortOrder": "string",
   "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AdditionalCodeRepositoryEquals (p. 620)**

A filter that returns only notebook instances with associated with the specified git repository.

Type: String


Pattern: ^https://([^/]+)/?(.*)$|^[a-zA-Z0-9]*[^a-zA-Z0-9]*$

Required: No

**CreationTimeAfter (p. 620)**

A filter that returns only notebook instances that were created after the specified time (timestamp).

Type: Timestamp

Required: No

**CreationTimeBefore (p. 620)**

A filter that returns only notebook instances that were created before the specified time (timestamp).

Type: Timestamp

Required: No
DefaultCodeRepositoryContains (p. 620)

A string in the name or URL of a Git repository associated with this notebook instance. This filter returns only notebook instances associated with a git repository with a name that contains the specified string.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: \[a-zA-Z0-9-]+\]

Required: No

LastModifiedTimeAfter (p. 620)

A filter that returns only notebook instances that were modified after the specified time (timestamp).

Type: Timestamp

Required: No

LastModifiedTimeBefore (p. 620)

A filter that returns only notebook instances that were modified before the specified time (timestamp).

Type: Timestamp

Required: No

MaxResults (p. 620)

The maximum number of notebook instances to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NameContains (p. 620)

A string in the notebook instances' name. This filter returns only notebook instances whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: \[a-zA-Z0-9-]+\]

Required: No

NextToken (p. 620)

If the previous call to the ListNotebookInstances is truncated, the response includes a NextToken. You can use this token in your subsequent ListNotebookInstances request to fetch the next set of notebook instances.

Note

You might specify a filter or a sort order in your request. When response is truncated, you must use the same values for the filter and sort order in the next request.

Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

**NotebookInstanceLifecycleConfigNameContains (p. 620)**

A string in the name of a notebook instances lifecycle configuration associated with this notebook instance. This filter returns only notebook instances associated with a lifecycle configuration with a name that contains the specified string.

Type: String

Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9\-\[\]*\-\[a-zA-Z0-9\-\]]*
Required: No

**SortBy (p. 620)**

The field to sort results by. The default is Name.

Type: String

Valid Values: Name | CreationTime | Status

Required: No

**SortOrder (p. 620)**

The sort order for results.

Type: String

Valid Values: Ascending | Descending

Required: No

**StatusEquals (p. 620)**

A filter that returns only notebook instances with the specified status.

Type: String

Valid Values: Pending | InService | Stopping | Stopped | Failed | Deleting | Updating

Required: No

### Response Syntax

```json
{
"NextToken": "string",
"NotebookInstances": [ 
  {
    "AdditionalCodeRepositories": [ "string" ],
    "CreationTime": number,
    "DefaultCodeRepository": "string",
    "InstanceType": "string",
    "LastModifiedTime": number,
    "NotebookInstanceArn": "string",
  }
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 622)**

If the response to the previous ListNotebookInstances request was truncated, SageMaker returns this token. To retrieve the next set of notebook instances, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**NotebookInstances (p. 622)**

An array of NotebookInstanceSummary objects, one for each notebook instance.

Type: Array of NotebookInstanceSummary (p. 1226) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListPipelineExecutions

Service: Amazon SageMaker Service

Gets a list of the pipeline executions.

Request Syntax

```json
{
    "CreatedAfter": number,
    "CreatedBefore": number,
    "MaxResults": number,
    "NextToken": "string",
    "PipelineName": "string",
    "SortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreatedAfter (p. 624)**

A filter that returns the pipeline executions that were created after a specified time.

Type: Timestamp

Required: No

**CreatedBefore (p. 624)**

A filter that returns the pipeline executions that were created before a specified time.

Type: Timestamp

Required: No

**MaxResults (p. 624)**

The maximum number of pipeline executions to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NextToken (p. 624)**

If the result of the previous ListPipelineExecutions request was truncated, the response includes a NextToken. To retrieve the next set of pipeline executions, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*
Required: No

**PipelineName (p. 624)**

The name of the pipeline.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: ^[a-zA-Z0-9\-\[a-zA-Z0-9\]\{0,255\}

Required: Yes

**SortBy (p. 624)**

The field by which to sort results. The default is CreatedTime.

Type: String

Valid Values: CreationTime | PipelineExecutionArn

Required: No

**SortOrder (p. 624)**

The sort order for results.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```json
{
    "NextToken": "string",
    "PipelineExecutionSummaries": [
        {
            "PipelineExecutionArn": "string",
            "PipelineExecutionDescription": "string",
            "PipelineExecutionDisplayName": "string",
            "PipelineExecutionFailureReason": "string",
            "PipelineExecutionStatus": "string",
            "StartTime": number
        }
    ]
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 625)**

If the result of the previous ListPipelineExecutions request was truncated, the response includes a NextToken. To retrieve the next set of pipeline executions, use the token in the next request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: . *

PipelineExecutionSummaries (p. 625)
Contains a sorted list of pipeline execution summary objects matching the specified filters. Each run summary includes the Amazon Resource Name (ARN) of the pipeline execution, the run date, and the status. This list can be empty.

Type: Array of PipelineExecutionSummary (p. 1270) objects
Array Members: Minimum number of 0 items. Maximum number of 100 items.

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFoundException
Resource being access is not found.
HTTP Status Code: 400

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListPipelineExecutionSteps
Service: Amazon SageMaker Service

Gets a list of PipelineExecutionStep objects.

Request Syntax

{
    "MaxResults": number,
    "NextToken": "string",
    "PipelineExecutionArn": "string",
    "SortOrder": "string"
}

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

MaxResults (p. 627)

The maximum number of pipeline execution steps to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken (p. 627)

If the result of the previous ListPipelineExecutionSteps request was truncated, the response includes a NextToken. To retrieve the next set of pipeline execution steps, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

PipelineExecutionArn (p. 627)

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:pipeline\/.*/\execution/\.*$*

Required: No

SortOrder (p. 627)

The field by which to sort results. The default is CreatedTime.
Type: String

Valid Values: Ascending | Descending

Required: No

Response Syntax

```json
{
    "NextToken": "string",
    "PipelineExecutionSteps": [
        {
            "AttemptCount": number,
            "CacheHitResult": {
                "SourcePipelineExecutionArn": "string"
            },
            "EndTime": number,
            "FailureReason": "string",
            "Metadata": {
                "Callback": {
                    "CallbackToken": "string",
                    "OutputParameters": [
                        {
                            "Name": "string",
                            "Value": "string"
                        }
                    ],
                    "SqsQueueUrl": "string"
                },
                "ClarifyCheck": {
                    "BaselineUsedForDriftCheckConstraints": "string",
                    "CalculatedBaselineConstraints": "string",
                    "CheckJobArn": "string",
                    "CheckType": "string",
                    "ModelPackageGroupName": "string",
                    "RegisterNewBaseline": boolean,
                    "SkipCheck": boolean,
                    "ViolationReport": "string"
                },
                "Condition": {
                    "Outcome": "string"
                },
                "EMR": {
                    "ClusterId": "string",
                    "LogFilePath": "string",
                    "StepId": "string",
                    "StepName": "string"
                },
                "Fail": {
                    "ErrorMessage": "string"
                },
                "Lambda": {
                    "Arn": "string",
                    "OutputParameters": [
                        {
                            "Name": "string",
                            "Value": "string"
                        }
                    ]
                },
                "Model": {
                    "Arn": "string"
                }
            }
        }
    ]
}
```
"ProcessingJob": {
  "Arn": "string"
},
"QualityCheck": {
  "BaselineUsedForDriftCheckConstraints": "string",
  "BaselineUsedForDriftCheckStatistics": "string",
  "CalculatedBaselineConstraints": "string",
  "CalculatedBaselineStatistics": "string",
  "CheckJobArn": "string",
  "CheckType": "string",
  "ModelPackageGroupName": "string",
  "RegisterNewBaseline": boolean,
  "SkipCheck": boolean,
  "ViolationReport": "string"
},
"RegisterModel": {
  "Arn": "string"
},
"TrainingJob": {
  "Arn": "string"
},
"TransformJob": {
  "Arn": "string"
},
"TuningJob": {
  "Arn": "string"
},
"StartTime": number,
"StepDescription": "string",
"StepDisplayName": "string",
"StepName": "string",
"StepStatus": "string"
]

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 628)**

If the result of the previous `ListPipelineExecutionSteps` request was truncated, the response includes a `NextToken`. To retrieve the next set of pipeline execution steps, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**PipelineExecutionSteps (p. 628)**

A list of `PipelineExecutionStep` objects. Each `PipelineExecutionStep` consists of StepName, StartTime, EndTime, StepStatus, and Metadata. Metadata is an object with properties for each job that contains relevant information about the job created by the step.

Type: Array of `PipelineExecutionStep (p. 1265)` objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

  Resource being access is not found.

  HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListPipelineParametersForExecution

Service: Amazon SageMaker Service

Gets a list of parameters for a pipeline execution.

Request Syntax

```json
{
    "MaxResults": number,
    "NextToken": "string",
    "PipelineExecutionArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

MaxResults (p. 631)

The maximum number of parameters to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken (p. 631)

If the result of the previous ListPipelineParametersForExecution request was truncated, the response includes a NextToken. To retrieve the next set of parameters, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

PipelineExecutionArn (p. 631)

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:pipeline/.*\|execution/.*\$*

Required: Yes

Response Syntax

```json
{
    
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 631)**

If the result of the previous ListPipelineParametersForExecution request was truncated, the response includes a NextToken. To retrieve the next set of parameters, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**PipelineParameters (p. 631)**

Contains a list of pipeline parameters. This list can be empty.

Type: Array of Parameter (p. 1248) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNot Found**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
ListPipelines
Service: Amazon SageMaker Service

Gets a list of pipelines.

Request Syntax

```
{
  "CreatedAfter": number,
  "CreatedBefore": number,
  "MaxResults": number,
  "NextToken": "string",
  "PipelineNamePrefix": "string",
  "SortBy": "string",
  "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreatedAfter (p. 634)

A filter that returns the pipelines that were created after a specified time.

Type: Timestamp

Required: No

CreatedBefore (p. 634)

A filter that returns the pipelines that were created before a specified time.

Type: Timestamp

Required: No

MaxResults (p. 634)

The maximum number of pipelines to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken (p. 634)

If the result of the previous ListPipelines request was truncated, the response includes a NextToken. To retrieve the next set of pipelines, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*
Required: No

**PipelineNamePrefix (p. 634)**

The prefix of the pipeline name.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: ^[a-zA-Z0-9]-*[a-zA-Z0-9]\{0,255\}

Required: No

**SortBy (p. 634)**

The field by which to sort results. The default is CreatedTime.

Type: String

Valid Values: Name | CreationTime

Required: No

**SortOrder (p. 634)**

The sort order for results.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```json
{
  "NextToken": "string",
  "PipelineSummaries": [
    {
      "CreationTime": number,
      "LastExecutionTime": number,
      "LastModifiedTime": number,
      "PipelineArn": "string",
      "PipelineDescription": "string",
      "PipelineDisplayName": "string",
      "PipelineName": "string",
      "RoleArn": "string"
    }
  ]
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 635)**

If the result of the previous ListPipelines request was truncated, the response includes a NextToken. To retrieve the next set of pipelines, use the token in the next request.
Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**PipelineSummaries (p. 635)**

Contains a sorted list of PipelineSummary objects matching the specified filters. Each PipelineSummary consists of PipelineArn, PipelineName, ExperimentName, PipelineDescription, CreationTime, LastModifiedTime, LastRunTime, and RoleArn. This list can be empty.

Type: Array of PipelineSummary (p. 1273) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListProcessingJobs
Service: Amazon SageMaker Service

Lists processing jobs that satisfy various filters.

Request Syntax

```
{
    "CreationTimeAfter": number,
    "CreationTimeBefore": number,
    "LastModifiedTimeAfter": number,
    "LastModifiedTimeBefore": number,
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string",
    "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 637)**

A filter that returns only processing jobs created after the specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 637)**

A filter that returns only processing jobs created after the specified time.

Type: Timestamp

Required: No

**LastModifiedTimeAfter (p. 637)**

A filter that returns only processing jobs modified after the specified time.

Type: Timestamp

Required: No

**LastModifiedTimeBefore (p. 637)**

A filter that returns only processing jobs modified before the specified time.

Type: Timestamp

Required: No

**MaxResults (p. 637)**

The maximum number of processing jobs to return in the response.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**NameContains (p. 637)**
A string in the processing job name. This filter returns only processing jobs whose name contains the specified string.
Type: String
Required: No

**NextToken (p. 637)**
If the result of the previous ListProcessingJobs request was truncated, the response includes a NextToken. To retrieve the next set of processing jobs, use the token in the next request.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

**SortBy (p. 637)**
The field to sort results by. The default is CreationTime.
Type: String
Valid Values: Name | CreationTime | Status
Required: No

**SortOrder (p. 637)**
The sort order for results. The default is Ascending.
Type: String
Valid Values: Ascending | Descending
Required: No

**StatusEquals (p. 637)**
A filter that retrieves only processing jobs with a specific status.
Type: String
Valid Values: InProgress | Completed | Failed | Stopping | Stopped
Required: No

**Response Syntax**

```
{
  "NextToken": "string",
  "ProcessingJobSummaries": [
  ]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken (p. 638)

If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of processing jobs, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

ProcessingJobSummaries (p. 638)

An array of ProcessingJobSummary objects, each listing a processing job.

Type: Array of ProcessingJobSummary (p. 1284) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListProjects
Service: Amazon SageMaker Service

Gets a list of the projects in an AWS account.

Request Syntax

```json
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "MaxResults": number,
   "NameContains": "string",
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 640)**

A filter that returns the projects that were created after a specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 640)**

A filter that returns the projects that were created before a specified time.

Type: Timestamp

Required: No

**MaxResults (p. 640)**

The maximum number of projects to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NameContains (p. 640)**

A filter that returns the projects whose name contains a specified string.

Type: String


Pattern: ^[a-zA-Z0-9][^a-zA-Z0-9]{0,31}$

Required: No
NextToken (p. 640)

If the result of the previous ListProjects request was truncated, the response includes a NextToken. To retrieve the next set of projects, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

SortBy (p. 640)

The field by which to sort results. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No

SortOrder (p. 640)

The sort order for results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No

Response Syntax

```json
{
  "NextToken": "string",
  "ProjectSummaryList": [
    {
      "CreationTime": number,
      "ProjectArn": "string",
      "ProjectDescription": "string",
      "ProjectId": "string",
      "ProjectName": "string",
      "ProjectStatus": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken (p. 641)

If the result of the previous ListCompilationJobs request was truncated, the response includes a NextToken. To retrieve the next set of model compilation jobs, use the token in the next request.

Type: String
ListProjects

Length Constraints: Maximum length of 8192.

Pattern: . *

ProjectSummaryList (p. 641)

A list of summaries of projects.

Type: Array of ProjectSummary (p. 1312) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListStudioLifecycleConfigs

Service: Amazon SageMaker Service

Lists the Studio Lifecycle Configurations in your AWS Account.

Request Syntax

```
{
  "AppTypeEquals": "string",
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "ModifiedTimeAfter": number,
  "ModifiedTimeBefore": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AppTypeEquals (p. 643)**

A parameter to search for the App Type to which the Lifecycle Configuration is attached.

- Type: String
- Valid Values: JupyterServer | KernelGateway
- Required: No

**CreationTimeAfter (p. 643)**

A filter that returns only Lifecycle Configurations created on or after the specified time.

- Type: Timestamp
- Required: No

**CreationTimeBefore (p. 643)**

A filter that returns only Lifecycle Configurations created on or before the specified time.

- Type: Timestamp
- Required: No

**MaxResults (p. 643)**

The maximum number of Studio Lifecycle Configurations to return in the response. The default value is 10.

- Type: Integer
- Valid Range: Minimum value of 1. Maximum value of 100.
- Required: No
**ModifiedTimeAfter (p. 643)**

A filter that returns only Lifecycle Configurations modified after the specified time.

Type: Timestamp

Required: No

**ModifiedTimeBefore (p. 643)**

A filter that returns only Lifecycle Configurations modified before the specified time.

Type: Timestamp

Required: No

**NameContains (p. 643)**

A string in the Lifecycle Configuration name. This filter returns only Lifecycle Configurations whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$

Required: No

**NextToken (p. 643)**

If the previous call to ListStudioLifecycleConfigs didn't return the full set of Lifecycle Configurations, the call returns a token for getting the next set of Lifecycle Configurations.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 643)**

The property used to sort results. The default value is CreationTime.

Type: String

Valid Values: CreationTime | LastModifiedTime | Name

Required: No

**SortOrder (p. 643)**

The sort order. The default value is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```
"NextToken": "string",
"StudioLifecycleConfigs": [
  {
    "CreationTime": number,
    "LastModifiedTime": number,
    "StudioLifecycleConfigAppType": "string",
    "StudioLifecycleConfigArn": "string",
    "StudioLifecycleConfigName": "string"
  }
]

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken (p. 644)

A token for getting the next set of actions, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

StudioLifecycleConfigs (p. 644)

A list of Lifecycle Configurations and their properties.

Type: Array of StudioLifecycleConfigDetails (p. 1372) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceInUse

Resource being accessed is in use.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
• AWS SDK for Ruby V3
ListSubscribedWorkteams
Service: Amazon SageMaker Service

Gets a list of the work teams that you are subscribed to in the AWS Marketplace. The list may be empty if no work team satisfies the filter specified in the NameContains parameter.

Request Syntax

```json
{
   "MaxResults": number,
   "NameContains": "string",
   "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

MaxResults (p. 647)

The maximum number of work teams to return in each page of the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NameContains (p. 647)

A string in the work team name. This filter returns only work teams whose name contains the specified string.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

Required: No

NextToken (p. 647)

If the result of the previous ListSubscribedWorkteams request was truncated, the response includes a NextToken. To retrieve the next set of labeling jobs, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

Response Syntax

```json
{
}
```
"NextToken": "string",
"SubscribedWorkteams": [
  {
    "ListingId": "string",
    "MarketplaceDescription": "string",
    "MarketplaceTitle": "string",
    "SellerName": "string",
    "WorkteamArn": "string"
  }
]

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 647)**

If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of work teams, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

**SubscribedWorkteams (p. 647)**

An array of Workteam objects, each describing a work team.

Type: Array of SubscribedWorkteam (p. 1374) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListTags
Service: Amazon SageMaker Service

Returns the tags for the specified SageMaker resource.

Request Syntax

```json
{
   "MaxResults": number,
   "NextToken": "string",
   "ResourceArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**MaxResults (p. 649)**

- Maximum number of tags to return.
- Type: Integer
- Valid Range: Minimum value of 50.
- Required: No

**NextToken (p. 649)**

- If the response to the previous ListTags request is truncated, SageMaker returns this token. To retrieve the next set of tags, use it in the subsequent request.
- Type: String
- Length Constraints: Maximum length of 8192.
- Pattern: .*
- Required: No

**ResourceArn (p. 649)**

- The Amazon Resource Name (ARN) of the resource whose tags you want to retrieve.
- Type: String
- Length Constraints: Maximum length of 256.
- Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:.*
- Required: Yes

Response Syntax

```json
{
}
```
"NextToken": "string",
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
]

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 649)**

If response is truncated, SageMaker includes a token in the response. You can use this token in your subsequent request to fetch next set of tokens.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: . *

**Tags (p. 649)**

An array of Tag objects, each with a tag key and a value.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListTrainingJobs

Service: Amazon SageMaker Service

Lists training jobs.

**Note**
When `StatusEquals` and `MaxResults` are set at the same time, the `MaxResults` number of training jobs are first retrieved ignoring the `StatusEquals` parameter and then they are filtered by the `StatusEquals` parameter, which is returned as a response.

For example, if `ListTrainingJobs` is invoked with the following parameters:

```
{ ... MaxResults: 100, StatusEquals: InProgress ... }
```

First, 100 trainings jobs with any status, including those other than `InProgress`, are selected (sorted according to the creation time, from the most current to the oldest). Next, those with a status of `InProgress` are returned.

You can quickly test the API using the following AWS CLI code.

```bash
aws sagemaker list-training-jobs --max-results 100 --status-equals InProgress
```

**Request Syntax**

```
{
  "CreationTimeAfter": number,
  "CreationTimeBefore": number,
  "LastModifiedTimeAfter": number,
  "LastModifiedTimeBefore": number,
  "MaxResults": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "StatusEquals": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 651)**

A filter that returns only training jobs created after the specified time (timestamp).

Type: Timestamp

Required: No

**CreationTimeBefore (p. 651)**

A filter that returns only training jobs created before the specified time (timestamp).

Type: Timestamp

Required: No

**LastModifiedTimeAfter (p. 651)**

A filter that returns only training jobs modified after the specified time (timestamp).

Type: Timestamp

Required: No
ListTrainingJobs

Required: No

**LastModifiedTimeBefore (p. 651)**

A filter that returns only training jobs modified before the specified time (timestamp).

Type: Timestamp

Required: No

**MaxResults (p. 651)**

The maximum number of training jobs to return in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NameContains (p. 651)**

A string in the training job name. This filter returns only training jobs whose name contains the specified string.

Type: String

Length Constraints: Maximum length of 63.

Pattern: [a-zA-Z0-9\-]+

Required: No

**NextToken (p. 651)**

If the result of the previous ListTrainingJobs request was truncated, the response includes a NextToken. To retrieve the next set of training jobs, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 651)**

The field to sort results by. The default is CreationTime.

Type: String

Valid Values: Name | CreationTime | Status

Required: No

**SortOrder (p. 651)**

The sort order for results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No
ListTrainingJobs

StatusEquals (p. 651)

A filter that retrieves only training jobs with a specific status.

Type: String

Valid Values: InProgress | Completed | Failed | Stopping | Stopped

Required: No

Response Syntax

```json
{
  "NextToken": "string",
  "TrainingJobSummaries": [
    {
      "CreationTime": number,
      "LastModifiedTime": number,
      "TrainingEndTime": number,
      "TrainingJobArn": "string",
      "TrainingJobName": "string",
      "TrainingJobStatus": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken (p. 653)

If the response is truncated, SageMaker returns this token. To retrieve the next set of training jobs, use it in the subsequent request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

TrainingJobSummaries (p. 653)

An array of TrainingJobSummary objects, each listing a training job.

Type: Array of TrainingJobSummary (p. 1398) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListTrainingJobsForHyperParameterTuningJob

Service: Amazon SageMaker Service

Gets a list of TrainingJobSummary objects that describe the training jobs that a hyperparameter tuning job launched.

Request Syntax

```json
{
   "HyperParameterTuningJobName": "string",
   "MaxResults": number,
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string",
   "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**HyperParameterTuningJobName (p. 655)**

The name of the tuning job whose training jobs you want to list.

- Type: String
- Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}$

- Required: Yes

**MaxResults (p. 655)**

The maximum number of training jobs to return. The default value is 10.

- Type: Integer
- Valid Range: Minimum value of 1. Maximum value of 100.

- Required: No

**NextToken (p. 655)**

If the result of the previous ListTrainingJobsForHyperParameterTuningJob request was truncated, the response includes a NextToken. To retrieve the next set of training jobs, use the token in the next request.

- Type: String
- Length Constraints: Maximum length of 8192.
- Pattern: .*

- Required: No

**SortBy (p. 655)**

The field to sort results by. The default is Name.
If the value of this field is `FinalObjectiveMetricValue`, any training jobs that did not return an objective metric are not listed.

Type: String

Valid Values: Name | CreationTime | Status | FinalObjectiveMetricValue

Required: No

**SortOrder (p. 655)**

The sort order for results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No

**StatusEquals (p. 655)**

A filter that returns only training jobs with the specified status.

Type: String

Valid Values: InProgress | Completed | Failed | Stopping | Stopped

Required: No

**Response Syntax**

```json
{
    "NextToken": "string",
    "TrainingJobSummaries": [
        {
            "CreationTime": number,
            "FailureReason": "string",
            "FinalHyperParameterTuningJobObjectiveMetric": {
                "MetricName": "string",
                "Type": "string",
                "Value": number
            },
            "ObjectiveStatus": "string",
            "TrainingEndTime": number,
            "TrainingJobArn": "string",
            "TrainingJobDefinitionName": "string",
            "TrainingJobName": "string",
            "TrainingJobStatus": "string",
            "TrainingStartTime": number,
            "TunedHyperParameters": {
                "string": "string"
            },
            "TuningJobName": "string"
        }
    ]
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
NextToken (p. 656)

If the result of this ListTrainingJobsForHyperParameterTuningJob request was truncated, the response includes a NextToken. To retrieve the next set of training jobs, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

TrainingJobSummaries (p. 656)

A list of TrainingJobSummary (p. 1398) objects that describe the training jobs that the ListTrainingJobsForHyperParameterTuningJob request returned.

Type: Array of HyperParameterTrainingJobSummary (p. 1086) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListTransformJobs
Service: Amazon SageMaker Service

Lists transform jobs.

Request Syntax

```
{
   "CreationTimeAfter": number,
   "CreationTimeBefore": number,
   "LastModifiedTimeAfter": number,
   "LastModifiedTimeBefore": number,
   "MaxResults": number,
   "NameContains": "string",
   "NextToken": "string",
   "SortBy": "string",
   "SortOrder": "string",
   "StatusEquals": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreationTimeAfter (p. 658)**

A filter that returns only transform jobs created after the specified time.

Type: Timestamp

Required: No

**CreationTimeBefore (p. 658)**

A filter that returns only transform jobs created before the specified time.

Type: Timestamp

Required: No

**LastModifiedTimeAfter (p. 658)**

A filter that returns only transform jobs modified after the specified time.

Type: Timestamp

Required: No

**LastModifiedTimeBefore (p. 658)**

A filter that returns only transform jobs modified before the specified time.

Type: Timestamp

Required: No

**MaxResults (p. 658)**

The maximum number of transform jobs to return in the response. The default value is 10.
Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**NameContains (p. 658)**
A string in the transform job name. This filter returns only transform jobs whose name contains the specified string.

Type: String
Length Constraints: Maximum length of 63.
Pattern: [a-zA-Z0-9\-]+
Required: No

**NextToken (p. 658)**
If the result of the previous ListTransformJobs request was truncated, the response includes a NextToken. To retrieve the next set of transform jobs, use the token in the next request.

Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

**SortBy (p. 658)**
The field to sort results by. The default is CreationTime.

Type: String
Valid Values: Name | CreationTime | Status
Required: No

**SortOrder (p. 658)**
The sort order for results. The default is Descending.

Type: String
Valid Values: Ascending | Descending
Required: No

**StatusEquals (p. 658)**
A filter that retrieves only transform jobs with a specific status.

Type: String
Valid Values: InProgress | Completed | Failed | Stopping | Stopped
Required: No

**Response Syntax**

```json
{
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 659)**

If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of transform jobs, use it in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**TransformJobSummaries (p. 659)**

An array of TransformJobSummary objects.

Type: Array of TransformJobSummary (p. 1413) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListTrialComponents
Service: Amazon SageMaker Service

Lists the trial components in your account. You can sort the list by trial component name or creation time. You can filter the list to show only components that were created in a specific time range. You can also filter on one of the following:

- ExperimentName
- SourceArn
- TrialName

Request Syntax

```json
{
  "CreatedAfter": number,
  "CreatedBefore": number,
  "ExperimentName": "string",
  "MaxResults": number,
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "SourceArn": "string",
  "TrialName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CreatedAfter (p. 661)

A filter that returns only components created after the specified time.

Type: Timestamp

Required: No

CreatedBefore (p. 661)

A filter that returns only components created before the specified time.

Type: Timestamp

Required: No

ExperimentName (p. 661)

A filter that returns only components that are part of the specified experiment. If you specify ExperimentName, you can't filter by SourceArn or TrialName.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9])\{0,119\}
**MaxResults (p. 661)**

The maximum number of components to return in the response. The default value is 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NextToken (p. 661)**

If the previous call to ListTrialComponents didn't return the full set of components, the call returns a token for getting the next set of components.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 661)**

The property used to sort results. The default value is CreationTime.

Type: String

Valid Values: Name | CreationTime

Required: No

**SortOrder (p. 661)**

The sort order. The default value is Descending.

Type: String

Valid Values: Ascending | Descending

Required: No

**SourceArn (p. 661)**

A filter that returns only components that have the specified source Amazon Resource Name (ARN). If you specify SourceArn, you can't filter by ExperimentName or TrialName.

Type: String

Length Constraints: Maximum length of 256.

Required: No

**TrialName (p. 661)**

A filter that returns only components that are part of the specified trial. If you specify TrialName, you can't filter by ExperimentName or SourceArn.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,119}$

Required: No
Response Syntax

```json
{
"NextToken": "string",
"TrialComponentSummaries": [ 
{
"CreatedBy": {
"DomainId": "string",
"UserProfileArn": "string",
"UserProfileName": "string"
},
"CreationTime": number,
"DisplayName": "string",
"EndTime": number,
"LastModifiedBy": {
"DomainId": "string",
"UserProfileArn": "string",
"UserProfileName": "string"
},
"LastModifiedTime": number,
"StartTime": number,
"Status": {
"Message": "string",
"PrimaryStatus": "string"
},
"TrialComponentArn": "string",
"TrialComponentName": "string",
"TrialComponentSource": {
"SourceArn": "string",
"SourceType": "string"
}
}
]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 663)**

A token for getting the next set of components, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**TrialComponentSummaries (p. 663)**

A list of the summaries of your trial components.

Type: Array of TrialComponentSummary (p. 1437) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListTrials

Service: Amazon SageMaker Service

Lists the trials in your account. Specify an experiment name to limit the list to the trials that are part of that experiment. Specify a trial component name to limit the list to the trials that associated with that trial component. The list can be filtered to show only trials that were created in a specific time range. The list can be sorted by trial name or creation time.

Request Syntax

```json
{
  "CreatedAfter": number,
  "CreatedBefore": number,
  "ExperimentName": "string",
  "MaxResults": number,
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string",
  "TrialComponentName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CreatedAfter (p. 665)**

A filter that returns only trials created after the specified time.

Type: Timestamp

Required: No

**CreatedBefore (p. 665)**

A filter that returns only trials created before the specified time.

Type: Timestamp

Required: No

**ExperimentName (p. 665)**

A filter that returns only trials that are part of the specified experiment.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$`

Required: No

**MaxResults (p. 665)**

The maximum number of trials to return in the response. The default value is 10.

Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**NextToken (p. 665)**
If the previous call to ListTrials didn't return the full set of trials, the call returns a token for getting the next set of trials.
Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

**SortBy (p. 665)**
The property used to sort results. The default value is CreationTime.
Type: String
Valid Values: Name | CreationTime
Required: No

**SortOrder (p. 665)**
The sort order. The default value is Descending.
Type: String
Valid Values: Ascending | Descending
Required: No

**TrialComponentName (p. 665)**
A filter that returns only trials that are associated with the specified trial component.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$
Required: No

**Response Syntax**

```json
{
   "NextToken": "string",
   "TrialSummaries": [
      {
         "CreationTime": number,
         "DisplayName": "string",
         "LastModifiedTime": number,
         "TrialArn": "string",
         "TrialName": "string",
         "TrialSource": {
            "SourceArn": "string",
            "SourceType": "string"
         }
      }
   ]
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**NextToken (p. 666)**

A token for getting the next set of trials, if there are any.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**TrialSummaries (p. 666)**

A list of the summaries of your trials.

Type: Array of TrialSummary (p. 1440) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListUserProfiles
Service: Amazon SageMaker Service
Lists user profiles.

Request Syntax

```json
{
    "DomainIdEquals": "string",
    "MaxResults": number,
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string",
    "UserProfileNameContains": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainIdEquals (p. 668)**

A parameter by which to filter the results.

Type: String

Length Constraints: Maximum length of 63.

Required: No

**MaxResults (p. 668)**

Returns a list up to a specified limit.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NextToken (p. 668)**

If the previous response was truncated, you will receive this token. Use it in your next request to receive the next set of results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

**SortBy (p. 668)**

The parameter by which to sort the results. The default is CreationTime.

Type: String
Valid Values: CreationTime | LastModifiedTime

Required: No

**SortOrder (p. 668)**

The sort order for the results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No

**UserProfileNameContains (p. 668)**

A parameter by which to filter the results.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

Required: No

### Response Syntax

```json
{
  "NextToken": "string",
  "UserProfiles": [
    {
      "CreationTime": number,
      "DomainId": "string",
      "LastModifiedTime": number,
      "Status": "string",
      "UserProfileName": "string"
    }
  ]
}
```

### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 669)**

If the previous response was truncated, you will receive this token. Use it in your next request to receive the next set of results.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

**UserProfiles (p. 669)**

The list of user profiles.
Type: Array of UserProfileDetails (p. 1450) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListWorkforces
Service: Amazon SageMaker Service

Use this operation to list all private and vendor workforces in an AWS Region. Note that you can only have one private workforce per AWS Region.

Request Syntax

```json
{
  "MaxResults": number,
  "NameContains": "string",
  "NextToken": "string",
  "SortBy": "string",
  "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**MaxResults (p. 671)**

The maximum number of workforces returned in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

**NameContains (p. 671)**

A filter you can use to search for workforces using part of the workforce name.

Type: String


Pattern: `^[a-zA-Z0-9](\[a-zA-Z0-9\-\]){0,62}$`

Required: No

**NextToken (p. 671)**

A token to resume pagination.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: `.*`

Required: No

**SortBy (p. 671)**

Sort workforces using the workforce name or creation date.
Type: String

Valid Values: Name | CreateDate

Required: No

**SortOrder (p. 671)**

Sort workforces in ascending or descending order.

Type: String

Valid Values: Ascending | Descending

Required: No

**Response Syntax**

```json
{
   "NextToken": "string",
   "Workforces": [
      {
         "CognitoConfig": {
            "ClientId": "string",
            "UserPool": "string"
         },
         "CreateDate": number,
         "LastUpdatedDate": number,
         "OidcConfig": {
            "AuthorizationEndpoint": "string",
            "ClientId": "string",
            "Issuer": "string",
            "JwksUri": "string",
            "LogoutEndpoint": "string",
            "TokenEndpoint": "string",
            "UserInfoEndpoint": "string"
         },
         "SourceIpConfig": {
            "Cidrs": [ "string" ]
         },
         "SubDomain": "string",
         "WorkforceArn": "string",
         "WorkforceName": "string"
      }
   ]
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 672)**

A token to resume pagination.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*
Workforces (p. 672)

A list containing information about your workforce.

Type: Array of Workforce (p. 1457) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
ListWorkteams
Service: Amazon SageMaker Service

Gets a list of private work teams that you have defined in a region. The list may be empty if no work team satisfies the filter specified in the NameContains parameter.

Request Syntax

```json
{
    "MaxResults": number,
    "NameContains": "string",
    "NextToken": "string",
    "SortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

MaxResults (p. 674)

The maximum number of work teams to return in each page of the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NameContains (p. 674)

A string in the work team's name. This filter returns only work teams whose name contains the specified string.

Type: String


Pattern: ^[a-zA-Z0-9]([-][a-zA-Z0-9]){0,62}$

Required: No

NextToken (p. 674)

If the result of the previous ListWorkteams request was truncated, the response includes a NextToken. To retrieve the next set of labeling jobs, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.

Pattern: .*

Required: No

SortBy (p. 674)

The field to sort results by. The default is CreationTime.
Type: String

Valid Values: Name | CreateDate

Required: No

**SortOrder (p. 674)**

The sort order for results. The default is Ascending.

Type: String

Valid Values: Ascending | Descending

Required: No

### Response Syntax

```json
{
    "NextToken": "string",
    "Workteams": [
        {
            "CreateDate": number,
            "Description": "string",
            "LastUpdatedDate": number,
            "MemberDefinitions": [
                {
                    "CognitoMemberDefinition": {
                        "ClientId": "string",
                        "UserGroup": "string",
                        "UserPool": "string"
                    },
                    "OidcMemberDefinition": {
                        "Groups": [ "string" ]
                    }
                }
            ],
            "NotificationConfiguration": {
                "NotificationTopicArn": "string"
            },
            "ProductListingIds": [ "string" ],
            "SubDomain": "string",
            "WorkforceArn": "string",
            "WorkteamArn": "string",
            "WorkteamName": "string"
        }
    ]
}
```

### Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 675)**

If the response is truncated, Amazon SageMaker returns this token. To retrieve the next set of work teams, use it in the subsequent request.

Type: String
Length Constraints: Maximum length of 8192.

Pattern: .*

Workteams (p. 675)

An array of Workteam objects, each describing a work team.

Type: Array of Workteam (p. 1459) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
PutModelPackageGroupPolicy
Service: Amazon SageMaker Service

Adds a resource policy to control access to a model group. For information about resource policies, see Identity-based policies and resource-based policies in the AWS Identity and Access Management User Guide.

Request Syntax

```
{
  "ModelPackageGroupName": "string",
  "ResourcePolicy": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ModelPackageGroupName (p. 677)**

The name of the model group to add a resource policy to.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

**ResourcePolicy (p. 677)**

The resource policy for the model group.

Type: String


Pattern: .*

Required: Yes

Response Syntax

```
{
  "ModelPackageGroupArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
ModelPackageGroupArn (p. 677)

The Amazon Resource Name (ARN) of the model package group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package-group/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
QueryLineage

Service: Amazon SageMaker Service

Use this action to inspect your lineage and discover relationships between entities. For more information, see Querying Lineage Entities in the Amazon SageMaker Developer Guide.

Request Syntax

```
{
  "Direction": "string",
  "Filters": {
    "CreatedAfter": number,
    "CreatedBefore": number,
    "LineageTypes": [ "string" ],
    "ModifiedAfter": number,
    "ModifiedBefore": number,
    "Properties": {
      "string": "string"
    },
    "Types": [ "string" ]
  },
  "IncludeEdges": boolean,
  "MaxDepth": number,
  "MaxResults": number,
  "NextToken": "string",
  "StartArns": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**Direction (p. 679)**

Associations between lineage entities have a direction. This parameter determines the direction from the StartArn(s) that the query traverses.

Type: String

Valid Values: Both | Ascendants | Descendants

Required: No

**Filters (p. 679)**

A set of filtering parameters that allow you to specify which entities should be returned.

- Properties - Key-value pairs to match on the lineage entities' properties.
- LineageTypes - A set of lineage entity types to match on. For example: TrialComponent, Artifact, or Context.
- CreatedBefore - Filter entities created before this date.
- ModifiedBefore - Filter entities modified before this date.
- ModifiedAfter - Filter entities modified after this date.

Type: QueryFilters (p. 1322) object

Required: No
IncludeEdges (p. 679)

Setting this value to True retrieves not only the entities of interest but also the Associations and lineage entities on the path. Set to False to only return lineage entities that match your query.

Type: Boolean

Required: No

MaxDepth (p. 679)

The maximum depth in lineage relationships from the StartArns that are traversed. Depth is a measure of the number of Associations from the StartArn entity to the matched results.

Type: Integer

Valid Range: Maximum value of 10.

Required: No

MaxResults (p. 679)

Limits the number of vertices in the results. Use the NextToken in a response to to retrieve the next page of results.

Type: Integer

Valid Range: Maximum value of 50.

Required: No

NextToken (p. 679)

Limits the number of vertices in the request. Use the NextToken in a response to to retrieve the next page of results.

Type: String

Length Constraints: Maximum length of 8192.

Required: No

StartArns (p. 679)

A list of resource Amazon Resource Name (ARN) that represent the starting point for your lineage query.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

Required: Yes

Response Syntax

```json
{
    "Edges": [
        {
        
```
"AssociationType": "string",
"DestinationArn": "string",
"SourceArn": "string"
]

"NextToken": "string",
"Vertices": [
  {
    "Arn": "string",
    "LineageType": "string",
    "Type": "string"
  }
]

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Edges (p. 680)

A list of edges that connect vertices in the response.

Type: Array of Edge (p. 1003) objects

NextToken (p. 680)

Limits the number of vertices in the response. Use the NextToken in a response to retrieve the next page of results.

Type: String

Length Constraints: Maximum length of 8192.

Vertices (p. 680)

A list of vertices connected to the start entity(ies) in the lineage graph.

Type: Array of Vertex (p. 1455) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFoundException

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
RegisterDevices
Service: Amazon SageMaker Service

Register devices.

Request Syntax

```json
{
  "DeviceFleetName": "string",
  "Devices": [
    {
      "Description": "string",
      "DeviceName": "string",
      "IotThingName": "string"
    }
  ],
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeviceFleetName (p. 683)**

The name of the fleet.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

**Devices (p. 683)**

A list of devices to register with SageMaker Edge Manager.

Type: Array of Device (p. 989) objects

Required: Yes

**Tags (p. 683)**

The tags associated with devices.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No
Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
RenderUiTemplate
Service: Amazon SageMaker Service

Renders the UI template so that you can preview the worker's experience.

Request Syntax

```json
{
   "HumanTaskUiArn": "string",
   "RoleArn": "string",
   "Task": {
      "Input": "string"
   },
   "UiTemplate": {
      "Content": "string"
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**HumanTaskUiArn (p. 685)**

The HumanTaskUiArn of the worker UI that you want to render. Do not provide a HumanTaskUiArn if you use the UiTemplate parameter.

See a list of available Human Ui Amazon Resource Names (ARNs) in UiConfig (p. 1444).

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:human-task-ui/.*`

Required: No

**RoleArn (p. 685)**

The Amazon Resource Name (ARN) that has access to the S3 objects that are used by the template.

Type: String


Pattern: `^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9_\-+\=,\.\@\#\$\%^&\*\(\)\[\]\|\-\\_\\^\%\\\]\"\:\\]+$`

Required: Yes

**Task (p. 685)**

A RenderableTask object containing a representative task to render.

Type: `RenderableTask (p. 1334)` object

Required: Yes
UiTemplate (p. 685)

A Template object containing the worker UI template to render.

Type: UiTemplate (p. 1446) object

Required: No

Response Syntax

```json
{
    "Errors": [  
        {  
            "Code": "string",
            "Message": "string"
        }
    ],
    "RenderedContent": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors (p. 686)

A list of one or more RenderingError objects if any were encountered while rendering the template. If there were no errors, the list is empty.

Type: Array of RenderingError (p. 1335) objects

RenderedContent (p. 686)

A Liquid template that renders the HTML for the worker UI.

Type: String

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
RetryPipelineExecution
Service: Amazon SageMaker Service
Retry the execution of the pipeline.

Request Syntax

```json
{
   "ClientRequestToken": "string",
   "ParallelismConfiguration": {
      "MaxParallelExecutionSteps": number
   },
   "PipelineExecutionArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ClientRequestToken (p. 688)**

A unique, case-sensitive identifier that you provide to ensure the idempotency of the operation. An idempotent operation completes no more than once.

Type: String


Required: Yes

**ParallelismConfiguration (p. 688)**

This configuration, if specified, overrides the parallelism configuration of the parent pipeline.

Type: ParallelismConfiguration (p. 1247) object

Required: No

**PipelineExecutionArn (p. 688)**

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*\execution/.*$.

Required: Yes

Response Syntax

```json
{
   "PipelineExecutionArn": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**PipelineExecutionArn (p. 688)**

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline\/.*/.*$\

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ConflictException**

There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
Search
Service: Amazon SageMaker Service

Finds Amazon SageMaker resources that match a search query. Matching resources are returned as a list of SearchRecord objects in the response. You can sort the search results by any resource property in an ascending or descending order.

You can query against the following value types: numeric, text, Boolean, and timestamp.

Request Syntax

```json
{
    "MaxResults": number,
    "NextToken": "string",
    "Resource": "string",
    "SearchExpression": {
        "Filters": [
            {
                "Name": "string",
                "Operator": "string",
                "Value": "string"
            }
        ],
        "NestedFilters": [
            {
                "Filters": [
                    {
                        "Name": "string",
                        "Operator": "string",
                        "Value": "string"
                    }
                ],
                "NestedPropertyName": "string"
            }
        ],
        "Operator": "string",
        "SubExpressions": [ "SearchExpression"
        ]
    },
    "SortBy": "string",
    "SortOrder": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**MaxResults (p. 691)**

The maximum number of results to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No
NextToken (p. 691)

If more than MaxResults resources match the specified SearchExpression, the response includes a NextToken. The NextToken can be passed to the next SearchRequest to continue retrieving results.

Type: String
Length Constraints: Maximum length of 8192.
Pattern: .*
Required: No

Resource (p. 691)

The name of the Amazon SageMaker resource to search for.

Type: String
Valid Values: TrainingJob | Experiment | ExperimentTrial | ExperimentTrialComponent | Endpoint | ModelPackage | ModelPackageGroup | Pipeline | PipelineExecution | FeatureGroup | Project
Required: Yes

SearchExpression (p. 691)

A Boolean conditional statement. Resources must satisfy this condition to be included in search results. You must provide at least one subexpression, filter, or nested filter. The maximum number of recursive SubExpressions, NestedFilters, and Filters that can be included in a SearchExpression object is 50.

Type: SearchExpression (p. 1355) object
Required: No

SortBy (p. 691)

The name of the resource property used to sort the SearchResults. The default is LastModifiedTime.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 255.
Pattern: .+
Required: No

SortOrder (p. 691)

How SearchResults are ordered. Valid values are Ascending or Descending. The default is Descending.

Type: String
Valid Values: Ascending | Descending
Required: No

Response Syntax

```json
{
    "NextToken": "string",
}
```
"Results": [
  {
    "Endpoint": {
      "CreationTime": number,
      "DataCaptureConfig": {
        "CaptureStatus": "string",
        "CurrentSamplingPercentage": number,
        "DestinationS3Uri": "string",
        "EnableCapture": boolean,
        "KmsKeyId": "string"
      },
      "EndpointArn": "string",
      "EndpointConfigName": "string",
      "EndpointName": "string",
      "EndpointStatus": "string",
      "FailureReason": "string",
      "LastModifiedTime": number,
      "MonitoringSchedules": [
        {
          "CreationTime": number,
          "EndpointName": "string",
          "FailureReason": "string",
          "LastModifiedTime": number,
          "LastMonitoringExecutionSummary": {
            "CreationTime": number,
            "EndpointName": "string",
            "FailureReason": "string",
            "LastModifiedTime": number,
            "MonitoringExecutionStatus": "string",
            "MonitoringJobDefinitionName": "string",
            "MonitoringScheduleName": "string",
            "MonitoringType": "string",
            "ProcessingJobArn": "string",
            "ScheduledTime": number
          },
          "MonitoringScheduleArn": "string",
          "MonitoringScheduleConfig": {
            "MonitoringJobDefinition": {
              "BaselineConfig": {
                "BaseliningJobName": "string",
                "ConstraintsResource": {
                  "S3Uri": "string"
                },
                "StatisticsResource": {
                  "S3Uri": "string"
                }
              },
              "Environment": {
                "string": "string"
              },
              "MonitoringAppSpecification": {
                "ContainerArguments": [ "string" ],
                "ContainerEntrypoint": [ "string" ],
                "ImageUri": "string",
                "PostAnalyticsProcessorSourceUri": "string",
                "RecordPreprocessorSourceUri": "string"
              },
              "MonitoringInputs": [
                {
                  "EndpointInput": {
                    "EndpointName": "string",
                    "EndTimeOffset": "string",
                    "FeaturesAttribute": "string",
                    "InferenceAttribute": "string",
                    "LocalPath": "string",
                    "ProbabilityAttribute": "string"
                  }
                }
              ]
            }
          }
        }
      ]
    }
  }
]
"ProbabilityThresholdAttribute": number,
"S3DataDistributionType": "string",
"S3InputMode": "string",
"StartTimeOffset": "string"
}
],
"MonitoringOutputConfig": {
  "KmsKeyId": "string",
  "MonitoringOutputs": [
    {
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
},
"MonitoringResources": {
  "ClusterConfig": {
    "InstanceCount": number,
    "InstanceType": "string",
    "VolumeKmsKeyId": "string",
    "VolumeSizeInGB": number
  }
},
"NetworkConfig": {
  "EnableInterContainerTrafficEncryption": boolean,
  "EnableNetworkIsolation": boolean,
  "VpcConfig": {
    "SecurityGroupId": [ "string" ],
    "Subnets": [ "string" ]
  }
},
"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeInSeconds": number
}
],
"MonitoringJobDefinitionName": "string",
"MonitoringType": "string",
"ScheduleConfig": {
  "ScheduleExpression": "string"
}
],
"MonitoringScheduleName": "string",
"MonitoringScheduleStatus": "string",
"MonitoringType": "string",
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
]
],
"ProductionVariants": [
  {
    "CurrentInstanceCount": number,
    "CurrentServerlessConfig": {
      "MaxConcurrency": number,
      "MemorySizeInMB": number
    },
    "CurrentWeight": number,
    "DeployedImages": [
Search

{
    "ResolutionTime": number,
    "ResolvedImage": "string",
    "SpecifiedImage": "string"
}

"DesiredInstanceCount": number,
"DesiredServerlessConfig": {
    "MaxConcurrency": number,
    "MemorySizeInMB": number
},
"DesiredWeight": number,
"VariantName": "string",
"VariantStatus": [
    {
        "StartTime": number,
        "Status": "string",
        "StatusMessage": "string"
    }
]

"Tags": [
    {
        "Key": "string",
        "Value": "string"
    }
]

"Experiment": {
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "Description": "string",
    "DisplayName": "string",
    "ExperimentArn": "string",
    "ExperimentName": "string",
    "LastModifiedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "LastModifiedTime": number,
    "Source": {
        "SourceArn": "string",
        "SourceType": "string"
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ]
},
"FeatureGroup": {
    "CreationTime": number,
    "Description": "string",
    "EventTimeFeatureName": "string",
    "FailureReason": "string",
    "FeatureDefinitions": [
        {
            "FeatureName": "string",
            "FeatureType": "string"
        }
    ]
}
Amazon SageMaker Amazon SageMaker API Reference

Search

```
{
  "FeatureGroupArn": "string",
  "FeatureGroupName": "string",
  "FeatureGroupStatus": "string",
  "OfflineStoreConfig": {
    "DataCatalogConfig": {
      "Catalog": "string",
      "Database": "string",
      "TableName": "string"
    },
    "DisableGlueTableCreation": boolean,
    "S3StorageConfig": {
      "KmsKeyId": "string",
      "ResolvedOutputS3Uri": "string",
      "S3Uri": "string"
    }
  },
  "OfflineStoreStatus": {
    "BlockedReason": "string",
    "Status": "string"
  },
  "OnlineStoreConfig": {
    "EnableOnlineStore": boolean,
    "SecurityConfig": {
      "KmsKeyId": "string"
    }
  },
  "RecordIdentifierFeatureName": "string",
  "RoleArn": "string",
  "Tags": [
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "ModelPackage": {
    "AdditionalInferenceSpecifications": [
      {
        "Containers": [
          {
            "ContainerHostname": "string",
            "Environment": {
              "string": "string"
            },
            "Framework": "string",
            "FrameworkVersion": "string",
            "Image": "string",
            "ImageDigest": "string",
            "ModelDataUrl": "string",
            "ModelInput": {
              "DataInputConfig": "string"
            },
            "NearestModelName": "string",
            "ProductId": "string"
          }
        ],
        "Description": "string",
        "Name": "string",
        "SupportedContentTypes": [ "string" ],
        "SupportedRealtimeInferenceInstanceTypes": [ "string" ],
        "SupportedResponseMIMETypes": [ "string" ],
        "SupportedTransformInstanceTypes": [ "string" ]
      }
    ]
  }
}
```
"ApprovalDescription": "string",
"CertifyForMarketplace": boolean,
"CreatedBy": {
  "DomainId": "string",
  "UserProfileArn": "string",
  "UserProfileName": "string"
},
"CreationTime": number,
"CustomerMetadataProperties": {
  "string": "string"
},
"Domain": "string",
"DriftCheckBaselines": {
  "Bias": {
    "ConfigFile": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "PostTrainingConstraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "PreTrainingConstraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    }
  },
  "Explainability": {
    "ConfigFile": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "Constraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    }
  },
  "ModelDataQuality": {
    "Constraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "Statistics": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    }
  },
  "ModelQuality": {
    "Constraints": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    },
    "Statistics": {
      "ContentDigest": "string",
      "ContentType": "string",
      "S3Uri": "string"
    }
  }
}
"InferenceSpecification": {
    "Containers": [
        {
            "ContainerHostname": "string",
            "Environment": {
                "string": "string"
            },
            "Framework": "string",
            "FrameworkVersion": "string",
            "Image": "string",
            "ImageDigest": "string",
            "ModelDataUrl": "string",
            "ModelInput": {
                "DataInputConfig": "string"
            },
            "NearestModelName": "string",
            "ProductId": "string"
        }
    ],
    "SupportedContentTypes": [ "string" ],
    "SupportedRealtimeInferenceInstanceTypes": [ "string" ],
    "SupportedResponseMIMETypes": [ "string" ],
    "SupportedTransformInstanceTypes": [ "string" ]
},
"LastModifiedBy": {
    "DomainId": "string",
    "UserProfileArn": "string",
    "UserProfileName": "string"
},
"LastModifiedTime": number,
"MetadataProperties": {
    "CommitId": "string",
    "GeneratedBy": "string",
    "ProjectId": "string",
    "Repository": "string"
},
"ModelApprovalStatus": "string",
"ModelMetrics": {
    "Bias": {
        "PostTrainingReport": {
            "ContentDigest": "string",
            "ContentType": "string",
            "S3Uri": "string"
        },
        "PreTrainingReport": {
            "ContentDigest": "string",
            "ContentType": "string",
            "S3Uri": "string"
        },
        "Report": {
            "ContentDigest": "string",
            "ContentType": "string",
            "S3Uri": "string"
        }
    },
    "Explainability": {
        "Report": {
            "ContentDigest": "string",
            "ContentType": "string",
            "S3Uri": "string"
        }
    },
    "ModelDataQuality": {
        "Constraints": {
            "ContentDigest": "string",
            "ContentType": "string",
            "S3Uri": "string"
        }
    }
}
"ContentType": "string",
"S3Uri": "string"
},
"Statistics": { 
  "ContentDigest": "string",
  "ContentType": "string",
  "S3Uri": "string"
}
},
"ModelQuality": { 
  "Constraints": { 
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  },
  "Statistics": { 
    "ContentDigest": "string",
    "ContentType": "string",
    "S3Uri": "string"
  }
},
"ModelPackageArn": "string",
"ModelPackageDescription": "string",
"ModelPackageGroupName": "string",
"ModelPackageName": "string",
"ModelPackageStatus": "string",
"ModelPackageStatusDetails": { 
  "ImageScanStatuses": [ 
    { 
      "FailureReason": "string",
      "Name": "string",
      "Status": "string"
    }
  ]
},
"ValidationStatuses": [ 
  { 
    "FailureReason": "string",
    "Name": "string",
    "Status": "string"
  }
]
},
"ModelPackageVersion": number,
"SamplePayloadUrl": "string",
"SourceAlgorithmSpecification": { 
  "SourceAlgorithms": [ 
    { 
      "AlgorithmName": "string",
      "ModelDataUrl": "string"
    }
  ]
},
"Tags": [ 
  { 
    "Key": "string",
    "Value": "string"
  }
],
"Task": "string",
"ValidationSpecification": { 
  "ValidationProfiles": [ 
    { 
      "ProfileName": "string",
      "TransformJobDefinition": { 
        "BatchStrategy": "string",

"Environment": {
    "string": "string"
},
"MaxConcurrentTransforms": number,
"MaxPayloadInMB": number,
"TransformInput": {
    "CompressionType": "string",
    "ContentType": "string",
    "DataSource": {
        "S3DataSource": {
            "S3DataType": "string",
            "S3Uri": "string"
        }
    },
    "SplitType": "string"
},
"TransformOutput": {
    "Accept": "string",
    "AssembleWith": "string",
    "KmsKeyId": "string",
    "S3OutputPath": "string"
},
"TransformResources": {
    "InstanceCount": number,
    "InstanceType": "string",
    "VolumeKmsKeyId": "string"
}
},
"ValidationRole": "string"
},
"ModelPackageGroup": {
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "ModelPackageGroupArn": "string",
    "ModelPackageGroupDescription": "string",
    "ModelPackageGroupName": "string",
    "ModelPackageGroupStatus": "string",
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ]
},
"Pipeline": {
    "CreatedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "CreationTime": number,
    "LastModifiedBy": {
        "DomainId": "string",
        "UserProfileArn": "string",
        "UserProfileName": "string"
    },
    "LastModifiedTime": number,
    "LastRunTime": number,
    "ParallelismConfiguration": {"Environment": {"string": "string"},"MaxConcurrentTransforms": number,"MaxPayloadInMB": number,"TransformInput": {
    "CompressionType": "string",
    "ContentType": "string",
    "DataSource": {
        "S3DataSource": {
            "S3DataType": "string",
            "S3Uri": "string"
        }
    },
    "SplitType": "string"
},
"TransformOutput": {
    "Accept": "string",
    "AssembleWith": "string",
    "KmsKeyId": "string",
    "S3OutputPath": "string"
},
"TransformResources": {
    "InstanceCount": number,
    "InstanceType": "string",
    "VolumeKmsKeyId": "string"
}
},
"ValidationRole": "string"}
"MaxParallelExecutionSteps": number
}
,"PipelineArn": "string",
"PipelineDescription": "string",
"PipelineDisplayName": "string",
"PipelineName": "string",
"PipelineStatus": "string",
"RoleArn": "string",
"Tags": [
{
"Key": "string",
"Value": "string"
}
]
,"PipelineExecution": {
"CreatedBy": {
"DomainId": "string",
"UserProfileArn": "string",
"UserProfileName": "string"
},
"CreationTime": number,
"FailureReason": "string",
"LastModifiedBy": {
"DomainId": "string",
"UserProfileArn": "string",
"UserProfileName": "string"
},
"LastModifiedTime": number,
"ParallelismConfiguration": {
"MaxParallelExecutionSteps": number
},
"PipelineArn": "string",
"PipelineExecutionArn": "string",
"PipelineExecutionDescription": "string",
"PipelineExecutionDisplayName": "string",
"PipelineExecutionStatus": "string",
"PipelineExperimentConfig": {
"ExperimentName": "string",
"TrialName": "string"
},
"PipelineParameters": [
{
"Name": "string",
"Value": "string"
}
],
"Project": {
"CreatedBy": {
"DomainId": "string",
"UserProfileArn": "string",
"UserProfileName": "string"
},
"CreationTime": number,
"LastModifiedBy": {
"DomainId": "string",
"UserProfileArn": "string",
"UserProfileName": "string"
},
"LastModifiedTime": number,
"ProjectArn": "string",
"ProjectDescription": "string",
"ProjectId": "string",
"ProjectName": "string",
"ProjectStatus": "string"}
<table>
<thead>
<tr>
<th>Key</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProvisionedProductId</td>
<td>string</td>
</tr>
<tr>
<td>ProvisionedProductStatusMessage</td>
<td>string</td>
</tr>
<tr>
<td>PathId</td>
<td>string</td>
</tr>
<tr>
<td>ProductId</td>
<td>string</td>
</tr>
<tr>
<td>ProvisioningArtifactId</td>
<td>string</td>
</tr>
<tr>
<td>ProvisioningParameters</td>
<td>[</td>
</tr>
<tr>
<td></td>
<td>{</td>
</tr>
<tr>
<td></td>
<td>Key: string,</td>
</tr>
<tr>
<td></td>
<td>Value: string</td>
</tr>
<tr>
<td></td>
<td>}</td>
</tr>
<tr>
<td></td>
<td>]</td>
</tr>
<tr>
<td>Tags</td>
<td>[</td>
</tr>
<tr>
<td></td>
<td>{</td>
</tr>
<tr>
<td></td>
<td>Key: string,</td>
</tr>
<tr>
<td></td>
<td>Value: string</td>
</tr>
<tr>
<td></td>
<td>}</td>
</tr>
<tr>
<td></td>
<td>]</td>
</tr>
<tr>
<td>TrainingJob</td>
<td></td>
</tr>
<tr>
<td>AlgorithmSpecification</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AlgorithmName: string</td>
</tr>
<tr>
<td></td>
<td>EnableSageMakerMetricsTimeSeries: boolean</td>
</tr>
<tr>
<td></td>
<td>MetricDefinitions: [</td>
</tr>
<tr>
<td></td>
<td>Name: string,</td>
</tr>
<tr>
<td></td>
<td>Regex: string</td>
</tr>
<tr>
<td></td>
<td>]</td>
</tr>
<tr>
<td></td>
<td>TrainingImage: string</td>
</tr>
<tr>
<td></td>
<td>TrainingInputMode: string</td>
</tr>
<tr>
<td></td>
<td>AutoMLJobArn: string</td>
</tr>
<tr>
<td></td>
<td>BillableTimeInSeconds: number</td>
</tr>
<tr>
<td>CheckpointConfig</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LocalPath: string</td>
</tr>
<tr>
<td></td>
<td>S3Uri: string</td>
</tr>
<tr>
<td></td>
<td>CreationTime: number</td>
</tr>
<tr>
<td>DebugHookConfig</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CollectionConfigurations: [</td>
</tr>
<tr>
<td></td>
<td>CollectionName: string,</td>
</tr>
<tr>
<td></td>
<td>CollectionParameters: {</td>
</tr>
<tr>
<td></td>
<td>string: string</td>
</tr>
<tr>
<td></td>
<td>}</td>
</tr>
<tr>
<td></td>
<td>HookParameters: [</td>
</tr>
<tr>
<td></td>
<td>string: string</td>
</tr>
<tr>
<td></td>
<td>]</td>
</tr>
<tr>
<td></td>
<td>LocalPath: string</td>
</tr>
<tr>
<td></td>
<td>S3OutputPath: string</td>
</tr>
<tr>
<td>DebugRuleConfigurations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[</td>
</tr>
<tr>
<td></td>
<td>InstanceType: string,</td>
</tr>
<tr>
<td></td>
<td>LocalPath: string,</td>
</tr>
<tr>
<td></td>
<td>RuleConfigurationName:</td>
</tr>
<tr>
<td></td>
<td>RuleEvaluatorImage:</td>
</tr>
<tr>
<td></td>
<td>RuleParameters: [</td>
</tr>
<tr>
<td></td>
<td>string: string</td>
</tr>
<tr>
<td></td>
<td>]</td>
</tr>
<tr>
<td></td>
<td>]</td>
</tr>
</tbody>
</table>
"OutputDataConfig": {
  "KmsKeyId": "string",
  "S3OutputPath": "string"
},
"ResourceConfig": {
  "InstanceCount": number,
  "InstanceType": "string",
  "VolumeKmsKeyId": "string",
  "VolumeSizeInGB": number
},
"RetryStrategy": {
  "MaximumRetryAttempts": number
},
"RoleArn": "string",
"SecondaryStatus": "string",
"SecondaryStatusTransitions": [
  {
    "EndTime": number,
    "StartTime": number,
    "Status": "string",
    "StatusMessage": "string"
  }
],
"StoppingCondition": {
  "MaxRuntimeInSeconds": number,
  "MaxWaitTimeInSeconds": number
},
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
],
"TensorBoardOutputConfig": {
  "LocalPath": "string",
  "S3OutputPath": "string"
},
"TrainingEndTime": number,
"TrainingJobArn": "string",
"TrainingJobName": "string",
"TrainingJobStatus": "string",
"TrainingStartTime": number,
"TrainingTimeInSeconds": number,
"TuningJobArn": "string",
"VpcConfig": {
  "SecurityGroupIds": [ "string" ],
  "Subnets": [ "string" ]
}
},
"Trial": {
  "CreatedBy": {
    "DomainId": "string",
    "UserProfileArn": "string",
    "UserProfileName": "string"
  },
  "CreationTime": number,
  "DisplayName": "string",
  "ExperimentName": "string",
  "LastModifiedBy": {
    "DomainId": "string",
    "UserProfileArn": "string",
    "UserProfileName": "string"
  },
  "LastModifiedTime": number,
  "MetadataProperties": {
    "CommitId": "string",
    "Key": "string",
    "Value": "string"
  }
}
"CreatedBy": { "DomainId": "string", "UserProfileArn": "string", "UserProfileName": "string" },
"CreationTime": number,
"DisplayName": "string",
"EndTime": number,
"InputArtifacts": { "string": { "MediaType": "string", "Value": "string" } },
"LastModifiedBy": { "DomainId": "string", "UserProfileArn": "string", "UserProfileName": "string" },
"LastModifiedTime": number,
"LineageGroupArn": "string",
"MetadataProperties": { "CommitId": "string", "GeneratedBy": "string", "ProjectId": "string", "Repository": "string" },
"Metrics": [
{ "Avg": number,
"Count": number,
"Last": number,
"705"}]}
"Max": number,
"MetricName": "string",
"Min": number,
"SourceArn": "string",
"StdDev": number,
"TimeStamp": number
},
"OutputArtifacts": {
  "string": {
    "MediaType": "string",
    "Value": "string"
  }
},
"Parameters": {
  "string": {
    "NumberValue": number,
    "StringValue": "string"
  }
},
"Parents": [
  {
    "ExperimentName": "string",
    "TrialName": "string"
  }
],
"Source": {
  "SourceArn": "string",
  "SourceType": "string"
},
"SourceDetail": {
  "ProcessingJob": {
    "AppSpecification": {
      "ContainerArguments": [ "string" ],
      "ContainerEntrypoint": [ "string" ],
      "ImageUri": "string"
    },
    "AutoMLJobArn": "string",
    "CreationTime": number,
    "Environment": {
      "string": "string"
    },
    "ExitMessage": "string",
    "ExperimentConfig": {
      "ExperimentName": "string",
      "TrialComponentDisplayName": "string",
      "TrialName": "string"
    },
    "FailureReason": "string",
    "LastModifiedTime": number,
    "MonitoringScheduleArn": "string",
    "NetworkConfig": {
      "EnableInterContainerTrafficEncryption": boolean,
      "EnableNetworkIsolation": boolean,
      "VpcConfig": {
        "SecurityGroupIds": [ "string" ],
        "Subnets": [ "string" ]
      }
    },
    "ProcessingEndTime": number,
    "ProcessingInputs": [
      {
        "AppManaged": boolean,
        "DatasetDefinition": {
          "AthenaDatasetDefinition": {
            "Catalog": "string",
            "Database": "string",
            "TableName": "string"
          }
        }
      }
    ]
  }
}
"Database": "string",
"KmsKeyId": "string",
"OutputCompression": "string",
"OutputFormat": "string",
"OutputS3Uri": "string",
"QueryString": "string",
"WorkGroup": "string"
},
"DataDistributionType": "string",
"InputMode": "string",
"LocalPath": "string",
"RedshiftDatasetDefinition": {
  "ClusterId": "string",
  "ClusterRoleArn": "string",
  "Database": "string",
  "DbUser": "string",
  "KmsKeyId": "string",
  "OutputCompression": "string",
  "OutputFormat": "string",
  "OutputS3Uri": "string",
  "QueryString": "string"
}
},
"InputName": "string",
"S3Input": {
  "LocalPath": "string",
  "S3CompressionType": "string",
  "S3DataDistributionType": "string",
  "S3DataType": "string",
  "S3InputMode": "string",
  "S3Uri": "string"
}
],
"ProcessingJobArn": "string",
"ProcessingJobName": "string",
"ProcessingJobStatus": "string",
"ProcessingOutputConfig": {
  "KmsKeyId": "string",
  "Outputs": [
    {
      "AppManaged": boolean,
      "FeatureStoreOutput": {
        "FeatureGroupName": "string"
      },
      "OutputName": "string",
      "S3Output": {
        "LocalPath": "string",
        "S3UploadMode": "string",
        "S3Uri": "string"
      }
    }
  ]
},
"ProcessingResources": {
  "ClusterConfig": {
    "InstanceCount": number,
    "InstanceType": "string",
    "VolumeKmsKeyId": "string",
    "VolumeSizeInGB": number
  }
},
"ProcessingStartTime": number,
"RoleArn": "string",
"StoppingCondition": {
  "MaxRuntimeIn Seconds": number
"EnableInterContainerTrafficEncryption": boolean,
"EnableManagedSpotTraining": boolean,
"EnableNetworkIsolation": boolean,
"Environment": {
  "string": "string"
},
"ExperimentConfig": {
  "ExperimentName": "string",
  "TrialComponentDisplayName": "string",
  "TrialName": "string"
},
"FailureReason": "string",
"FinalMetricDataList": [
  {
    "MetricName": "string",
    "Timestamp": number,
    "Value": number
  }
],
"HyperParameters": {
  "string": "string"
},
"InputDataConfig": [
  {
    "ChannelName": "string",
    "CompressionType": "string",
    "ContentType": "string",
    "DataSource": {
      "FileSystemDataSource": {
        "DirectoryPath": "string",
        "FileSystemAccessMode": "string",
        "FileSystemId": "string",
        "FileSystemType": "string"
      },
      "S3DataSource": {
        "AttributeNames": [ "string" ],
        "S3DataDistributionType": "string",
        "S3DataType": "string",
        "S3Uri": "string"
      }
    },
    "InputMode": "string",
    "RecordWrapperType": "string",
    "ShuffleConfig": {
      "Seed": number
    }
  }
],
"LabelingJobArn": "string",
"LastModifiedTime": number,
"ModelArtifacts": {
  "S3ModelArtifacts": "string"
},
"OutputDataConfig": {
  "KmsKeyId": "string",
  "S3OutputPath": "string"
},
"ResourceConfig": {
  "InstanceCount": number,
  "InstanceType": "string",
  "VolumeKmsKeyId": "string",
  "VolumeSizeInGB": number
},
"RetryStrategy": {
  "MaximumRetryAttempts": number
Amazon SageMaker API Reference

Search

{},
"RoleArn": "string",
"SecondaryStatus": "string",
"SecondaryStatusTransitions": [
  {
    "EndTime": number,
    "StartTime": number,
    "Status": "string",
    "StatusMessage": "string"
  }
],
"StoppingCondition": {
  "MaxRuntimeInSeconds": number,
  "MaxWaitTimeInSeconds": number
},
"Tags": [
  {
    "Key": "string",
    "Value": "string"
  }
],
"TensorBoardOutputConfig": {
  "LocalPath": "string",
  "S3OutputPath": "string"
},
"TrainingEndTime": number,
"TrainingJobArn": "string",
"TrainingJobName": "string",
"TrainingJobStatus": "string",
"TrainingStartTime": number,
"TrainingTimeInSeconds": number,
"TuningJobArn": "string",
"VpcConfig": {
  "SecurityGroupIds": [ "string" ],
  "Subnets": [ "string" ]
},
"TransformJob": {
  "AutoMLJobArn": "string",
  "BatchStrategy": "string",
  "CreationTime": number,
  "DataProcessing": {
    "InputFilter": "string",
    "JoinSource": "string",
    "OutputFilter": "string"
  },
  "Environment": {
    "string": "string"
  },
  "ExperimentConfig": {
    "ExperimentName": "string",
    "TrialComponentDisplayName": "string",
    "TrialName": "string"
  },
  "FailureReason": "string",
  "LabelingJobArn": "string",
  "MaxConcurrentTransforms": number,
  "MaxPayloadInMB": number,
  "ModelClientConfig": {
    "InvocationsMaxRetries": number,
    "InvocationsTimeoutInSeconds": number
  },
  "ModelName": "string",
  "Tags": [
    {
      "Key": "string",
    }
  ]
}
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**NextToken (p. 692)**

If the result of the previous Search request was truncated, the response includes a NextToken. To retrieve the next set of results, use the token in the next request.

Type: String

Length Constraints: Maximum length of 8192.
Pattern: .*

**Results (p. 692)**

A list of `SearchRecord` objects.

Type: Array of `SearchRecord (p. 1357)` objects

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
**SendPipelineExecutionStepFailure**  
*Service: Amazon SageMaker Service*

Notifies the pipeline that the execution of a callback step failed, along with a message describing why. When a callback step is run, the pipeline generates a callback token and includes the token in a message sent to Amazon Simple Queue Service (Amazon SQS).

**Request Syntax**

```json
{
   "CallbackToken": "string",
   "ClientRequestToken": "string",
   "FailureReason": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CallbackToken (p. 713)**

The pipeline generated token from the Amazon SQS queue.

Type: String  
Length Constraints: Fixed length of 10.  
Pattern: ^[a-zA-Z0-9]+$  
Required: Yes

**ClientRequestToken (p. 713)**

A unique, case-sensitive identifier that you provide to ensure the idempotency of the operation. An idempotent operation completes no more than one time.

Type: String  
Required: No

**FailureReason (p. 713)**

A message describing why the step failed.

Type: String  
Length Constraints: Maximum length of 256.  
Required: No

**Response Syntax**

```json
{
}
```
"PipelineExecutionArn": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**PipelineExecutionArn (p. 713)**

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline\./.*\ execution\./.*$

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
SendPipelineExecutionStepSuccess

Service: Amazon SageMaker Service

Notifies the pipeline that the execution of a callback step succeeded and provides a list of the step's output parameters. When a callback step is run, the pipeline generates a callback token and includes the token in a message sent to Amazon Simple Queue Service (Amazon SQS).

Request Syntax

```json
{
    "CallbackToken": "string",
    "ClientRequestToken": "string",
    "OutputParameters": [
        {
            "Name": "string",
            "Value": "string"
        }
    ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CallbackToken (p. 715)**

The pipeline generated token from the Amazon SQS queue.

Type: String

Length Constraints: Fixed length of 10.

Pattern: ^[a-zA-Z0-9]+$

Required: Yes

**ClientRequestToken (p. 715)**

A unique, case-sensitive identifier that you provide to ensure the idempotency of the operation. An idempotent operation completes no more than one time.

Type: String


Required: No

**OutputParameters (p. 715)**

A list of the output parameters of the callback step.

Type: Array of OutputParameter (p. 1246) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No
Response Syntax

```
{
  "PipelineExecutionArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**PipelineExecutionArn (p. 716)**

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:pipeline\/.*/.*\execution\/.*$

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFoundException**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
StartMonitoringSchedule

Service: Amazon SageMaker Service

Starts a previously stopped monitoring schedule.

**Note**

By default, when you successfully create a new schedule, the status of a monitoring schedule is scheduled.

**Request Syntax**

```json
{
   "MonitoringScheduleName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

**MonitoringScheduleName** (p. 717)

The name of the schedule to start.

Type: String


Pattern: ^[a-zA-Z0-9\-]*(a-zA-Z0-9\-\d)\{0,62\}$

Required: Yes

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see [Common Errors](#).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
**StartNotebookInstance**

Service: Amazon SageMaker Service

Launches an ML compute instance with the latest version of the libraries and attaches your ML storage volume. After configuring the notebook instance, SageMaker sets the notebook instance status to *InService*. A notebook instance's status must be *InService* before you can connect to your Jupyter notebook.

**Request Syntax**

```
{
   "NotebookInstanceName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see [Common Parameters (p. 1470)](#).

The request accepts the following data in JSON format.

**NotebookInstanceName (p. 719)**

The name of the notebook instance to start.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*`

Required: Yes

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
StartPipelineExecution

Service: Amazon SageMaker Service

Starts a pipeline execution.

Request Syntax

```
{
    "ClientRequestToken": "string",
    "ParallelismConfiguration": {
        "MaxParallelExecutionSteps": number
    },
    "PipelineExecutionDescription": "string",
    "PipelineExecutionDisplayName": "string",
    "PipelineName": "string",
    "PipelineParameters": [
        {
            "Name": "string",
            "Value": "string"
        }
    ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ClientRequestToken (p. 721)**

A unique, case-sensitive identifier that you provide to ensure the idempotency of the operation. An idempotent operation completes no more than once.

Type: String


Required: Yes

**ParallelismConfiguration (p. 721)**

This configuration, if specified, overrides the parallelism configuration of the parent pipeline for this specific run.

Type: ParallelismConfiguration (p. 1247) object

Required: No

**PipelineExecutionDescription (p. 721)**

The description of the pipeline execution.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 3072.

Pattern: .*

Required: No
PipelineExecutionDisplayName (p. 721)

The display name of the pipeline execution.
Type: String

Length Constraints: Minimum length of 1. Maximum length of 82.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,81\}

Required: No

PipelineName (p. 721)

The name of the pipeline.
Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,255\}

Required: Yes

PipelineParameters (p. 721)

Contains a list of pipeline parameters. This list can be empty.
Type: Array of Parameter (p. 1248) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

Response Syntax

```json
{
   "PipelineExecutionArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

PipelineExecutionArn (p. 722)

The Amazon Resource Name (ARN) of the pipeline execution.
Type: String

Length Constraints: Maximum length of 256.
Pattern: ^arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:pipeline\./.*\execution\./.*$

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).
ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
StopAutoMLJob
Service: Amazon SageMaker Service

A method for forcing the termination of a running job.

Request Syntax

```
{
    "AutoMLJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**AutoMLJobName (p. 724)**

The name of the object you are requesting.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
StopCompilationJob
Service: Amazon SageMaker Service

Stops a model compilation job.

To stop a job, Amazon SageMaker sends the algorithm the SIGTERM signal. This gracefully shuts the job down. If the job hasn't stopped, it sends the SIGKILL signal.

When it receives a StopCompilationJob request, Amazon SageMaker changes the CompilationJobSummary:CompilationJobStatus (p. 953) of the job to Stopping. After Amazon SageMaker stops the job, it sets the CompilationJobSummary:CompilationJobStatus (p. 953) to Stopped.

Request Syntax

```
{
    "CompilationJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

CompilationJobName (p. 726)

The name of the model compilation job to stop.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

- Resource being access is not found.

  HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
StopEdgePackagingJob
Service: Amazon SageMaker Service

Request to stop an edge packaging job.

Request Syntax

```
{
   "EdgePackagingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**EdgePackagingJobName (p. 728)**

The name of the edge packaging job.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
StopHyperParameterTuningJob

Service: Amazon SageMaker Service

Stops a running hyperparameter tuning job and all running training jobs that the tuning job launched.

All model artifacts output from the training jobs are stored in Amazon Simple Storage Service (Amazon S3). All data that the training jobs write to Amazon CloudWatch Logs are still available in CloudWatch. After the tuning job moves to the Stopped state, it releases all reserved resources for the tuning job.

Request Syntax

```
{
  "HyperParameterTuningJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

HyperParameterTuningJobName (p. 729)

The name of the tuning job to stop.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNot Found

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
StopInferenceRecommendationsJob
Service: Amazon SageMaker Service

Stops an Inference Recommender job.

Request Syntax

```json
{
   "JobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

JobName (p. 731)

The name of the job you want to stop.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,63}$

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
StopLabelingJob
Service: Amazon SageMaker Service

Stops a running labeling job. A job that is stopped cannot be restarted. Any results obtained before the job is stopped are placed in the Amazon S3 output bucket.

Request Syntax

```json
{
   "LabelingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**LabelingJobName (p. 733)**

The name of the labeling job to stop.

- Type: String
- Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}`
- Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

- Resource being access is not found.
- HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2

733
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
StopMonitoringSchedule
Service: Amazon SageMaker Service

Stops a previously started monitoring schedule.

Request Syntax

```
{
    "MonitoringScheduleName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**MonitoringScheduleName (p. 735)**

  - The name of the schedule to stop.
  - Type: String
  - Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$`
  - Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFoundException**

  - Resource being access is not found.
  - HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
StopNotebookInstance
Service: Amazon SageMaker Service

Terminates the ML compute instance. Before terminating the instance, SageMaker disconnects the ML storage volume from it. SageMaker preserves the ML storage volume. SageMaker stops charging you for the ML compute instance when you call StopNotebookInstance.

To access data on the ML storage volume for a notebook instance that has been terminated, call the StartNotebookInstance API. StartNotebookInstance launches another ML compute instance, configures it, and attaches the preserved ML storage volume so you can continue your work.

Request Syntax

```json
{
    "NotebookInstanceName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**NotebookInstanceName (p. 737)**

The name of the notebook instance to terminate.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
StopPipelineExecution
Service: Amazon SageMaker Service

Stops a pipeline execution.

Callback Step

A pipeline execution won't stop while a callback step is running. When you call
StopPipelineExecution on a pipeline execution with a running callback step, SageMaker Pipelines
sends an additional Amazon SQS message to the specified SQS queue. The body of the SQS message
contains a "Status" field which is set to "Stopping".

You should add logic to your Amazon SQS message consumer to take any needed
action (for example, resource cleanup) upon receipt of the message followed by a call to
SendPipelineExecutionStepSuccess or SendPipelineExecutionStepFailure.

Only when SageMaker Pipelines receives one of these calls will it stop the pipeline execution.

Lambda Step

A pipeline execution can't be stopped while a lambda step is running because the Lambda function
invoked by the lambda step can't be stopped. If you attempt to stop the execution while the Lambda
function is running, the pipeline waits for the Lambda function to finish or until the timeout is hit,
whichever occurs first, and then stops. If the Lambda function finishes, the pipeline execution status is
Stopped. If the timeout is hit the pipeline execution status is Failed.

Request Syntax

{
    "ClientRequestToken": "string",
    "PipelineExecutionArn": "string"
}

Request Parameters

For information about the parameters that are common to all actions, see Common
Parameters (p. 1470).

The request accepts the following data in JSON format.

ClientRequestToken (p. 739)

A unique, case-sensitive identifier that you provide to ensure the idempotency of the operation. An
idempotent operation completes no more than once.

Type: String


Required: Yes

PipelineExecutionArn (p. 739)

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.
StopPipelineExecution

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline\./.*\/execution\./.*$

Required: Yes

Response Syntax

```json
{
  "PipelineExecutionArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

PipelineExecutionArn (p. 740)

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline\./.*\/execution\./.*$

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
StopProcessingJob

Service: Amazon SageMaker Service

Stops a processing job.

Request Syntax

```json
{
    "ProcessingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ProcessingJobName (p. 741)**

The name of the processing job to stop.

- Type: String
- Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`
- Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

- Resource being access is not found.
- HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
StopTrainingJob
Service: Amazon SageMaker Service

Stops a training job. To stop a job, SageMaker sends the algorithm the \texttt{SIGTERM} signal, which delays job termination for 120 seconds. Algorithms might use this 120-second window to save the model artifacts, so the results of the training is not lost.

When it receives a \texttt{StopTrainingJob} request, SageMaker changes the status of the job to \texttt{Stopping}. After SageMaker stops the job, it sets the status to \texttt{Stopped}.

Request Syntax

```json
{
   "TrainingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

\textbf{TrainingJobName (p. 743)}

The name of the training job to stop.

Type: String


Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9]\(\-*[a-zA-Z0-9]\)\{0,62\}

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

\textbf{ResourceNotFound}

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
**StopTransformJob**

Service: Amazon SageMaker Service

Stops a batch transform job.

When Amazon SageMaker receives a `StopTransformJob` request, the status of the job changes to `Stopping`. After Amazon SageMaker stops the job, the status is set to `Stopped`. When you stop a batch transform job before it is completed, Amazon SageMaker doesn't store the job's output in Amazon S3.

**Request Syntax**

```json
{
    "TransformJobName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**TransformJobName (p. 745)**

The name of the batch transform job to stop.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

Required: Yes

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
UpdateAction
Service: Amazon SageMaker Service
Updates an action.

Request Syntax

```
{
  "ActionName": "string",
  "Description": "string",
  "Properties": {
    "string": "string"
  },
  "PropertiesToRemove": [ "string" ],
  "Status": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ActionName (p. 747)**

The name of the action to update.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: Yes

**Description (p. 747)**

The new description for the action.

Type: String

Length Constraints: Maximum length of 3072.

Pattern: .*

Required: No

**Properties (p. 747)**

The new list of properties. Overwrites the current property list.

Type: String to string map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 256.
Value Pattern: .*  
Required: No

**PropertiesToRemove (p. 747)**  
A list of properties to remove.  
Type: Array of strings  
Length Constraints: Maximum length of 256.  
Pattern: .*  
Required: No

**Status (p. 747)**  
The new status for the action.  
Type: String  
Valid Values: Unknown | InProgress | Completed | Failed | Stopping | Stopped  
Required: No

**Response Syntax**

```
{
   "ActionArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ActionArn (p. 748)**  
The Amazon Resource Name (ARN) of the action.  
Type: String  
Length Constraints: Maximum length of 256.  
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:action/.*

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

**ConflictException**  
There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.

HTTP Status Code: 400
ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateAppImageConfig
Service: Amazon SageMaker Service

Updates the properties of an AppImageConfig.

Request Syntax

```json
{
  "AppImageConfigName": "string",
  "KernelGatewayImageConfig": {
    "FileSystemConfig": {
      "DefaultGid": number,
      "DefaultUid": number,
      "MountPath": "string"
    },
    "KernelSpecs": [
      {
        "DisplayName": "string",
        "Name": "string"
      }
    ]
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AppImageConfigName (p. 750)

The name of the AppImageConfig to update.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]{0,62})

Required: Yes

KernelGatewayImageConfig (p. 750)

The new KernelGateway app to run on the image.

Type: KernelGatewayImageConfig (p. 1117) object

Required: No

Response Syntax

```json
{
  "AppImageConfigArn": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**AppImageConfigArn (p. 750)**

The Amazon Resource Name (ARN) for the AppImageConfig.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:app-image-config/.*`

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFoundException**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
**UpdateArtifact**

Service: Amazon SageMaker Service

Updates an artifact.

**Request Syntax**

```json
{
   "ArtifactArn": "string",
   "ArtifactName": "string",
   "Properties": {
      "string": "string"
   },
   "PropertiesToRemove": [ "string" ]
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ArtifactArn (p. 752)**

The Amazon Resource Name (ARN) of the artifact to update.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:artifact/.*

Required: Yes

**ArtifactName (p. 752)**

The new name for the artifact.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9\-][\-\*[a-zA-Z0-9\-]]{0,119}

Required: No

**Properties (p. 752)**

The new list of properties. Overwrites the current property list.

Type: String to string map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 256.
Value Pattern: . *
Required: No

PropertiesToRemove (p. 752)
A list of properties to remove.
Type: Array of strings
Length Constraints: Maximum length of 256.
Pattern: . *
Required: No

Response Syntax

```json
{
    "ArtifactArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

ArtifactArn (p. 753)
The Amazon Resource Name (ARN) of the artifact.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:artifact/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ConflictException
There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.
HTTP Status Code: 400

ResourceNotFoundException
Resource being access is not found.
HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateCodeRepository

Service: Amazon SageMaker Service

Updates the specified Git repository with the specified values.

Request Syntax

```json
{
   "CodeRepositoryName": "string",
   "GitConfig": {
      "SecretArn": "string"
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**CodeRepositoryName (p. 755)**

The name of the Git repository to update.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

**GitConfig (p. 755)**

The configuration of the git repository, including the URL and the Amazon Resource Name (ARN) of the AWS Secrets Manager secret that contains the credentials used to access the repository. The secret must have a staging label of AWSCURRENT and must be in the following format:

```json
{"username": UserName, "password": Password}
```

Type: GitConfigForUpdate (p. 1055) object

Required: No

Response Syntax

```json
{
   "CodeRepositoryArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
**CodeRepositoryArn (p. 755)**

The ARN of the Git repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:code-repository/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateContext
Service: Amazon SageMaker Service
Updates a context.

Request Syntax

```json
{
    "ContextName": "string",
    "Description": "string",
    "Properties": {
        "string": "string"
    },
    "PropertiesToRemove": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ContextName (p. 757)**

The name of the context to update.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: Yes

**Description (p. 757)**

The new description for the context.

Type: String

Length Constraints: Maximum length of 3072.

Pattern: .*

Required: No

**Properties (p. 757)**

The new list of properties. Overwrites the current property list.

Type: String to string map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 256.
Value Pattern: .*
Required: No

PropertiesToRemove (p. 757)
A list of properties to remove.
Type: Array of strings
Length Constraints: Maximum length of 256.
Pattern: .*
Required: No

Response Syntax

```json
{
   "ContextArn": "string"
}
```

Response Elements
If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

ContextArn (p. 758)
The Amazon Resource Name (ARN) of the context.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:context/.*

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

ConflictException
There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.
HTTP Status Code: 400

ForResourceNotFound
Resource being access is not found.
HTTP Status Code: 400

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
UpdateDeviceFleet

Service: Amazon SageMaker Service

Updates a fleet of devices.

Request Syntax

```json
{
   "Description": "string",
   "DeviceFleetName": "string",
   "EnableIotRoleAlias": boolean,
   "OutputConfig": {
      "KmsKeyId": "string",
      "PresetDeploymentConfig": "string",
      "PresetDeploymentType": "string",
      "S3OutputLocation": "string"
   },
   "RoleArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**Description (p. 760)**

Description of the fleet.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 800.

Pattern: [\S\s]+

Required: No

**DeviceFleetName (p. 760)**

The name of the fleet.

Type: String


Pattern: ^[a-zA-Z0-9]{0,62}$

Required: Yes

**EnableIotRoleAlias (p. 760)**

Whether to create an AWS IoT Role Alias during device fleet creation. The name of the role alias generated will match this pattern: "SageMakerEdge-{DeviceFleetName}".

For example, if your device fleet is called "demo-fleet", the name of the role alias will be "SageMakerEdge-demo-fleet".

Type: Boolean
**UpdateDeviceFleet**

**OutputConfig** (p. 760)

Output configuration for storing sample data collected by the fleet.

Type: EdgeOutputConfig (p. 1008) object

Required: Yes

**RoleArn** (p. 760)

The Amazon Resource Name (ARN) of the device.

Type: String


Pattern: ^arn:aws[a-zA-Z-]+:iam::\d{12}:role/\?[a-zA-Z0-9+=,.@-_/%]+$

Required: No

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateDevices
Service: Amazon SageMaker Service

Updates one or more devices in a fleet.

Request Syntax

```json
{
  "DeviceFleetName": "string",
  "Devices": [
    {
      "Description": "string",
      "DeviceName": "string",
      "IotThingName": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

DeviceFleetName (p. 762)

The name of the fleet the devices belong to.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

Devices (p. 762)

List of devices to register with Edge Manager agent.

Type: Array of Device (p. 989) objects

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
UpdateDomain
Service: Amazon SageMaker Service

Updates the default settings for new user profiles in the domain.

Request Syntax

```json
{
  "DefaultUserSettings": {
    "ExecutionRole": "string",
    "JupyterServerAppSettings": {
      "DefaultResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      },
      "LifecycleConfigArns": [ "string" ]
    },
    "KernelGatewayAppSettings": {
      "CustomImages": [ {
        "AppImageConfigName": "string",
        "ImageName": "string",
        "ImageVersionNumber": number
      } ],
      "DefaultResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      },
      "LifecycleConfigArns": [ "string" ]
    },
    "RSessionAppSettings": {
      "CustomImages": [ {
        "AppImageConfigName": "string",
        "ImageName": "string",
        "ImageVersionNumber": number
      } ],
      "DefaultResourceSpec": {
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      }
    },
    "RStudioServerProAppSettings": {
      "AccessStatus": "string",
      "UserGroup": "string"
    },
    "SecurityGroups": [ "string" ],
    "SharingSettings": { 
      "NotebookOutputOption": "string",
      "S3KmsKeyId": "string",
      "S3OutputPath": "string"
    },
    "TensorBoardAppSettings": { 
      "DefaultResourceSpec": {
        "InstanceType": "string",
      }
    }
  }
}
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DefaultUserSettings (p. 764)**

A collection of settings.

Type: UserSettings (p. 1452) object

Required: No

**DomainId (p. 764)**

The ID of the domain to be updated.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**DomainSettingsForUpdate (p. 764)**

A collection of DomainSettings configuration values to update.

Type: DomainSettingsForUpdate (p. 997) object

Required: No

Response Syntax

```json
{
  "DomainArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**DomainArn (p. 765)**

The Amazon Resource Name (ARN) of the domain.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:domain/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateEndpoint
Service: Amazon SageMaker Service

Deploys the new EndpointConfig specified in the request, switches to using newly created endpoint, and then deletes resources provisioned for the endpoint using the previous EndpointConfig (there is no availability loss).

When SageMaker receives the request, it sets the endpoint status to Updating. After updating the endpoint, it sets the status to InService. To check the status of an endpoint, use the DescribeEndpoint (p. 338) API.

Note
You must not delete an EndpointConfig in use by an endpoint that is live or while the UpdateEndpoint or CreateEndpoint operations are being performed on the endpoint. To update an endpoint, you must create a new EndpointConfig. If you delete the EndpointConfig of an endpoint that is active or being created or updated you may lose visibility into the instance type the endpoint is using. The endpoint must be deleted in order to stop incurring charges.

Request Syntax

```
{
  "DeploymentConfig": {
    "AutoRollbackConfiguration": {
      "Alarms": [
        {
          "AlarmName": "string"
        }
      ],
    },
    "BlueGreenUpdatePolicy": {
      "MaximumExecutionTimeoutInSeconds": number,
      "TerminationWaitInSeconds": number,
      "TrafficRoutingConfiguration": {
        "CanarySize": {
          "Type": "string",
          "Value": number
        },
        "LinearStepSize": {
          "Type": "string",
          "Value": number
        },
        "Type": "string",
        "WaitIntervalInSeconds": number
      }
    }
  },
  "EndpointConfigName": "string",
  "EndpointName": "string",
  "ExcludeRetainedVariantProperties": [
    {
      "VariantPropertyType": "string"
    }
  ],
  "RetainAllVariantProperties": boolean,
  "RetainDeploymentConfig": boolean
}
```
Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeploymentConfig (p. 767)**

The deployment configuration for an endpoint, which contains the desired deployment strategy and rollback configurations.

Type: DeploymentConfig (p. 987) object

Required: No

**EndpointConfigName (p. 767)**

The name of the new endpoint configuration.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$

Required: Yes

**EndpointName (p. 767)**

The name of the endpoint whose configuration you want to update.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$

Required: Yes

**ExcludeRetainedVariantProperties (p. 767)**

When you are updating endpoint resources with UpdateEndpoint:RetainAllVariantProperties (p. 768), whose value is set to true, ExcludeRetainedVariantProperties specifies the list of type VariantProperty (p. 1454) to override with the values provided by EndpointConfig. If you don't specify a value for ExcludeAllVariantProperties, no variant properties are overridden.

Type: Array of VariantProperty (p. 1454) objects

Array Members: Minimum number of 0 items. Maximum number of 3 items.

Required: No

**RetainAllVariantProperties (p. 767)**

When updating endpoint resources, enables or disables the retention of variant properties, such as the instance count or the variant weight. To retain the variant properties of an endpoint when updating it, set RetainAllVariantProperties to true. To use the variant properties specified in a new EndpointConfig call when updating an endpoint, set RetainAllVariantProperties to false. The default is false.

Type: Boolean
UpdateEndpoint

Required: No

RetainDeploymentConfig (p. 767)

Specifies whether to reuse the last deployment configuration. The default value is false (the configuration is not reused).

Type: Boolean

Required: No

Response Syntax

```json
{
   "EndpointArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

EndpointArn (p. 769)

The Amazon Resource Name (ARN) of the endpoint.

Type: String


Pattern: `arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:endpoint/.*`

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
UpdateEndpointWeightsAndCapacities
Service: Amazon SageMaker Service

Updates variant weight of one or more variants associated with an existing endpoint, or capacity of one variant associated with an existing endpoint. When it receives the request, SageMaker sets the endpoint status to Updating. After updating the endpoint, it sets the status to InService. To check the status of an endpoint, use the DescribeEndpoint API.

Request Syntax

```json
{
    "DesiredWeightsAndCapacities": [
        {
            "DesiredInstanceCount": number,
            "DesiredWeight": number,
            "VariantName": "string"
        }
    ],
    "EndpointName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

DesiredWeightsAndCapacities (p. 771)

An object that provides new capacity and weight values for a variant.

Type: Array of DesiredWeightAndCapacity objects

Array Members: Minimum number of 1 item.

Required: Yes

EndpointName (p. 771)

The name of an existing SageMaker endpoint.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](?=.*[a-zA-Z0-9])\d*$

Required: Yes

Response Syntax

```json
{
    "EndpointArn": "string"
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**EndpointArn (p. 771)**

The Amazon Resource Name (ARN) of the updated endpoint.

- **Type:** String
- **Length Constraints:** Minimum length of 20. Maximum length of 2048.
- **Pattern:** `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:endpoint/.*`

Errors

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

**ResourceLimitExceeded**

You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateExperiment

Service: Amazon SageMaker Service

Adds, updates, or removes the description of an experiment. Updates the display name of an experiment.

Request Syntax

```json
{
    "Description": "string",
    "DisplayName": "string",
    "ExperimentName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

Description (p. 773)

The description of the experiment.

Type: String

Length Constraints: Maximum length of 3072.

Pattern: .*

Required: No

DisplayName (p. 773)

The name of the experiment as displayed. The name doesn't need to be unique. If DisplayName isn't specified, ExperimentName is displayed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern:^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: No

ExperimentName (p. 773)

The name of the experiment to update.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern:^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: Yes

Response Syntax

```json
{
}
```
"ExperimentArn": "string"
}

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ExperimentArn (p. 773)**

The Amazon Resource Name (ARN) of the experiment.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ConflictException**

There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.

HTTP Status Code: 400

**ResourceNotFoundException**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
**UpdateImage**

Service: Amazon SageMaker Service

Updates the properties of a SageMaker image. To change the image's tags, use the AddTags (p. 17) and DeleteTags (p. 265) APIs.

**Request Syntax**

```json
{
    "DeleteProperties": [ "string" ],
    "Description": "string",
    "DisplayName": "string",
    "ImageName": "string",
    "RoleArn": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DeleteProperties (p. 775)**

A list of properties to delete. Only the Description and DisplayName properties can be deleted.

Type: Array of strings

Array Members: Maximum number of 2 items.


Pattern: ("DisplayName$"|"Description$")

Required: No

**Description (p. 775)**

The new description for the image.

Type: String


Pattern: .*

Required: No

**DisplayName (p. 775)**

The new display name for the image.

Type: String


Pattern: ^\S(.*\S)?$

Required: No
**ImageName (p. 775)**

The name of the image to update.

Type: String


Pattern: $^[a-zA-Z0-9][-.[a-zA-Z0-9]{0,62}$

Required: Yes

**RoleArn (p. 775)**

The new Amazon Resource Name (ARN) for the IAM role that enables Amazon SageMaker to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws([a-z\-\]*)*:iam::*\d{12}:role/\?([a-zA-Z0-9_0-9=,+,.@\-\_/]+*$

Required: No

**Response Syntax**

```
{
  "ImageArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ImageArn (p. 776)**

The Amazon Resource Name (ARN) of the image.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws([-\w\+]*):sagemaker:.+:\d{12}:image/[a-zA-Z0-9\{0-9\}\{\}0-9]+\s+.@\-\_/]+*$

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400
**ResourceNotFound**

   Resource being access is not found.

   HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateModelPackage
Service: Amazon SageMaker Service

Updates a versioned model.

Request Syntax

```json
{
  "AdditionalInferenceSpecificationsToAdd": [
    {
      "Containers": [
        {
          "ContainerHostname": "string",
          "Environment": {
            "string": "string"
          },
          "Framework": "string",
          "FrameworkVersion": "string",
          "Image": "string",
          "ImageDigest": "string",
          "ModelDataUrl": "string",
          "ModelInput": {
            "DataInputConfig": "string"
          },
          "NearestModelName": "string",
          "ProductId": "string"
        }
      ],
      "Description": "string",
      "Name": "string",
      "SupportedContentTypes": [ "string" ],
      "SupportedRealtimeInferenceInstanceTypes": [ "string" ],
      "SupportedResponseMIMETypes": [ "string" ],
      "SupportedTransformInstanceTypes": [ "string" ]
    }
  ],
  "ApprovalDescription": "string",
  "CustomerMetadataProperties": {
    "string": "string"
  },
  "CustomerMetadataPropertiesToRemove": [ "string" ],
  "ModelApprovalStatus": "string",
  "ModelPackageArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AdditionalInferenceSpecificationsToAdd (p. 778)

An array of additional Inference Specification objects to be added to the existing array additional Inference Specification. Total number of additional Inference Specifications can not exceed 15. Each additional Inference Specification specifies artifacts based on this model package that can be used on inference endpoints. Generally used with SageMaker Neo to store the compiled artifacts.

Type: Array of AdditionalInferenceSpecificationDefinition (p. 863) objects
Array Members: Minimum number of 1 item. Maximum number of 15 items.

Required: No

**ApprovalDescription (p. 778)**

A description for the approval status of the model.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: .*

Required: No

**CustomerMetadataProperties (p. 778)**

The metadata properties associated with the model package versions.

Type: String to string map

Map Entries: Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: ^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)${1,128}

Value Length Constraints: Minimum length of 1. Maximum length of 256.

Value Pattern: ^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)${1,256}

Required: No

**CustomerMetadataPropertiesToRemove (p. 778)**

The metadata properties associated with the model package versions to remove.

Type: Array of strings


Pattern: ^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)${1,128}

Required: No

**ModelApprovalStatus (p. 778)**

The approval status of the model.

Type: String

Valid Values: Approved | Rejected | PendingManualApproval

Required: No

**ModelPackageArn (p. 778)**

The Amazon Resource Name (ARN) of the model package.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package/.*
Required: Yes

Response Syntax

```
{
   "ModelPackageArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ModelPackageArn (p. 780)**

The Amazon Resource Name (ARN) of the model.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package/.*`

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateMonitoringSchedule

Service: Amazon SageMaker Service

Updates a previously created schedule.

Request Syntax

```json
{
  "MonitoringScheduleConfig": {
    "MonitoringJobDefinition": {
      "BaselineConfig": {
        "BaseliningJobName": "string",
        "ConstraintsResource": {
          "S3Uri": "string"
        },
        "StatisticsResource": {
          "S3Uri": "string"
        }
      },
      "Environment": {
        "string": "string"
      },
      "MonitoringAppSpecification": {
        "ContainerArguments": [ "string" ],
        "ContainerEntrypoint": [ "string" ],
        "ImageUri": "string",
        "PostAnalyticsProcessorSourceUri": "string",
        "RecordPreprocessorSourceUri": "string"
      },
      "MonitoringInputs": [
        {
          "EndpointInput": {
            "EndpointName": "string",
            "EndTimeOffset": "string",
            "FeaturesAttribute": "string",
            "InferenceAttribute": "string",
            "LocalPath": "string",
            "ProbabilityAttribute": "string",
            "ProbabilityThresholdAttribute": number,
            "S3DataDistributionType": "string",
            "S3InputMode": "string",
            "StartTimeOffset": "string"
          }
        }
      ],
      "MonitoringOutputConfig": {
        "KmsKeyId": "string",
        "MonitoringOutputs": [
          {
            "S3Output": {
              "LocalPath": "string",
              "S3UploadMode": "string",
              "S3Uri": "string"
            }
          }
        ]
      },
      "MonitoringResources": {
        "ClusterConfig": {
          "InstanceCount": number,
          "InstanceType": "string",
          "VolumeKmsKeyId": "string",
          "VolumeSizeInGB": number
        }
      }
    }
  }
}
```
UpdateMonitoringSchedule

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

MonitoringScheduleConfig (p. 781)

The configuration object that specifies the monitoring schedule and defines the monitoring job.

Type: MonitoringScheduleConfig (p. 1215) object

Required: Yes

MonitoringScheduleName (p. 781)

The name of the monitoring schedule. The name must be unique within an AWS Region within an AWS account.

Type: String


Pattern: ^[a-zA-Z0-9](¬*[a-zA-Z0-9]){0,62}$

Required: Yes

Response Syntax

{
  "MonitoringScheduleArn": "string"
}

Response Elements

If the action is successful, the service sends back an HTTP 200 response.
The following data is returned in JSON format by the service.

**MonitoringScheduleArn (p. 782)**

The Amazon Resource Name (ARN) of the monitoring schedule.

Type: String

Length Constraints: Maximum length of 256.

Pattern: . *

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateNotebookInstance
Service: Amazon SageMaker Service

Updates a notebook instance. NotebookInstance updates include upgrading or downgrading the ML compute instance used for your notebook instance to accommodate changes in your workload requirements.

Request Syntax

```json
{
    "AcceleratorTypes": [ "string" ],
    "AdditionalCodeRepositories": [ "string" ],
    "DefaultCodeRepository": "string",
    "DisassociateAcceleratorTypes": boolean,
    "DisassociateAdditionalCodeRepositories": boolean,
    "DisassociateDefaultCodeRepository": boolean,
    "DisassociateLifecycleConfig": boolean,
    "InstanceMetadataServiceConfiguration": {
        "MinimumInstanceMetadataServiceVersion": "string"
    },
    "InstanceType": "string",
    "LifecycleConfigName": "string",
    "NotebookInstanceName": "string",
    "RoleArn": "string",
    "RootAccess": "string",
    "VolumeSizeInGB": number
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

AcceleratorTypes (p. 784)

A list of the Elastic Inference (EI) instance types to associate with this notebook instance. Currently only one EI instance type can be associated with a notebook instance. For more information, see Using Elastic Inference in Amazon SageMaker.

Type: Array of strings

Valid Values: ml.eia1.medium | ml.eia1.large | ml.eia1.xlarge | ml.eia2.medium | ml.eia2.large | ml.eia2.xlarge

Required: No

AdditionalCodeRepositories (p. 784)

An array of up to three Git repositories to associate with the notebook instance. These can be either the names of Git repositories stored as resources in your account, or the URL of Git repositories in AWS CodeCommit or in any other Git repository. These repositories are cloned at the same level as the default repository of your notebook instance. For more information, see Associating Git Repositories with SageMaker Notebook Instances.

Type: Array of strings

Array Members: Maximum number of 3 items.

Pattern: `https://([^/]+)/?(.*)$|^[a-zA-Z0-9](-*[a-zA-Z0-9])*

Required: No

**DefaultCodeRepository (p. 784)**

The Git repository to associate with the notebook instance as its default code repository. This can be either the name of a Git repository stored as a resource in your account, or the URL of a Git repository in AWS CodeCommit or in any other Git repository. When you open a notebook instance, it opens in the directory that contains this repository. For more information, see Associating Git Repositories with SageMaker Notebook Instances.

Type: String


Pattern: `https://([^/]+)/?(.*)$|^[a-zA-Z0-9](-*[a-zA-Z0-9])*

Required: No

**DisassociateAcceleratorTypes (p. 784)**

A list of the Elastic Inference (EI) instance types to remove from this notebook instance. This operation is idempotent. If you specify an accelerator type that is not associated with the notebook instance when you call this method, it does not throw an error.

Type: Boolean

Required: No

**DisassociateAdditionalCodeRepositories (p. 784)**

A list of names or URLs of the default Git repositories to remove from this notebook instance. This operation is idempotent. If you specify a Git repository that is not associated with the notebook instance when you call this method, it does not throw an error.

Type: Boolean

Required: No

**DisassociateDefaultCodeRepository (p. 784)**

The name or URL of the default Git repository to remove from this notebook instance. This operation is idempotent. If you specify a Git repository that is not associated with the notebook instance when you call this method, it does not throw an error.

Type: Boolean

Required: No

**DisassociateLifecycleConfig (p. 784)**

Set to `true` to remove the notebook instance lifecycle configuration currently associated with the notebook instance. This operation is idempotent. If you specify a lifecycle configuration that is not associated with the notebook instance when you call this method, it does not throw an error.

Type: Boolean

Required: No

**InstanceMetadataServiceConfiguration (p. 784)**

Information on the IMDS configuration of the notebook instance
Type: `InstanceMetadataServiceConfiguration (p. 1111)` object

Required: No

**InstanceType (p. 784)**

The Amazon ML compute instance type.

Type: String

Valid Values: `ml.t2.medium | ml.t2.large | ml.t2.xlarge | ml.t2.2xlarge | ml.t3.medium | ml.t3.large | ml.t3.xlarge | ml.t3.2xlarge | ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge | ml.m4.16xlarge | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge | ml.m5.24xlarge | ml.m5d.large | ml.m5d.xlarge | ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.8xlarge | ml.m5d.12xlarge | ml.m5d.16xlarge | ml.m5d.24xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge | ml.c4.8xlarge | ml.c5.xlarge | ml.c5.2xlarge | ml.c5.4xlarge | ml.c5.9xlarge | ml.c5d.xlarge | ml.c5d.2xlarge | ml.c5d.4xlarge | ml.c5d.9xlarge | ml.p2.xlarge | ml.p2.7xlarge | ml.p2.16xlarge | ml.p2.32xlarge | ml.p3.xlarge | ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.p3dn.2xlarge | ml.p4dn.xlarge | ml.p4dn.2xlarge | ml.p4dn.8xlarge | ml.p4dn.16xlarge | ml.p4dn.32xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge | ml.g4dn.16xlarge | ml.g4dn.32xlarge | ml.g5.xlarge | ml.g5.2xlarge | ml.g5.4xlarge | ml.g5.8xlarge | ml.g5.16xlarge | ml.g5.32xlarge | ml.g5.64xlarge | ml.g5d.xlarge | ml.g5d.2xlarge | ml.g5d.4xlarge | ml.g5d.8xlarge | ml.g5d.16xlarge | ml.g5d.32xlarge | ml.g5d.64xlarge | ml.r5.xlarge | ml.r5.2xlarge | ml.r5.4xlarge | ml.r5.8xlarge | ml.r5.16xlarge | ml.r5.32xlarge | ml.r5.64xlarge | ml.r5d.xlarge | ml.r5d.2xlarge | ml.r5d.4xlarge | ml.r5d.8xlarge | ml.r5d.16xlarge | ml.r5d.32xlarge | ml.r5d.64xlarge | ml.xlarge | ml.2xlarge | ml.4xlarge | ml.8xlarge | ml.16xlarge | ml.32xlarge | ml.64xlarge`}

Required: No

**LifecycleConfigName (p. 784)**

The name of a lifecycle configuration to associate with the notebook instance. For information about lifestyle configurations, see Step 2.1: (Optional) Customize a Notebook Instance.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](\-*[a-zA-Z0-9])*`  

Required: No

**NotebookInstanceName (p. 784)**

The name of the notebook instance to update.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](\-*[a-zA-Z0-9])*`  

Required: Yes

**RoleArn (p. 784)**

The Amazon Resource Name (ARN) of the IAM role that SageMaker can assume to access the notebook instance. For more information, see SageMaker Roles.

*Note*  
To be able to pass this role to SageMaker, the caller of this API must have the `iam:PassRole` permission.
Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@-_\//]+$

Required: No

RootAccess (p. 784)

Whether root access is enabled or disabled for users of the notebook instance. The default value is Enabled.

**Note**
If you set this to **Disabled**, users don't have root access on the notebook instance, but lifecycle configuration scripts still run with root permissions.

Type: String

Valid Values: Enabled | Disabled

Required: No

VolumeSizeInGB (p. 784)

The size, in GB, of the ML storage volume to attach to the notebook instance. The default value is 5 GB. ML storage volumes are encrypted, so SageMaker can't determine the amount of available free space on the volume. Because of this, you can increase the volume size when you update a notebook instance, but you can't decrease the volume size. If you want to decrease the size of the ML storage volume in use, create a new notebook instance with the desired size.

Type: Integer


Required: No

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceLimitExceeded

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3
UpdateNotebookInstanceLifecycleConfig

Service: Amazon SageMaker Service

Updates a notebook instance lifecycle configuration created with the CreateNotebookInstanceLifecycleConfig (p. 150) API.

Request Syntax

```json
{
   "NotebookInstanceLifecycleConfigName": "string",
   "OnCreate": [
      {
         "Content": "string"
      }
   ],
   "OnStart": [
      {
         "Content": "string"
      }
   ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

NotebookInstanceLifecycleConfigName (p. 789)

The name of the lifecycle configuration.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9][-\[a-zA-Z0-9]]*$`

Required: Yes

OnCreate (p. 789)

The shell script that runs only once, when you create a notebook instance. The shell script must be a base64-encoded string.

Type: Array of NotebookInstanceLifecycleHook (p. 1225) objects

Array Members: Maximum number of 1 item.

Required: No

OnStart (p. 789)

The shell script that runs every time you start a notebook instance, including when you create the notebook instance. The shell script must be a base64-encoded string.

Type: Array of NotebookInstanceLifecycleHook (p. 1225) objects

Array Members: Maximum number of 1 item.
Required: No

Response Elements
If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors
For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceLimitExceeded
You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdatePipeline
Service: Amazon SageMaker Service

Updates a pipeline.

Request Syntax

```
{
  "ParallelismConfiguration": {
    "MaxParallelExecutionSteps": number
  },
  "PipelineDefinition": "string",
  "PipelineDefinitionS3Location": {
    "Bucket": "string",
    "ObjectKey": "string",
    "VersionId": "string"
  },
  "PipelineDescription": "string",
  "PipelineDisplayName": "string",
  "PipelineName": "string",
  "RoleArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ParallelismConfiguration (p. 791)**

If specified, it applies to all executions of this pipeline by default.

Type: ParallelismConfiguration (p. 1247) object

Required: No

**PipelineDefinition (p. 791)**

The JSON pipeline definition.

Type: String


Pattern: .*(?![ \r\n\t].)*

Required: No

**PipelineDefinitionS3Location (p. 791)**

The location of the pipeline definition stored in Amazon S3. If specified, SageMaker will retrieve the pipeline definition from this location.

Type: PipelineDefinitionS3Location (p. 1261) object

Required: No

**PipelineDescription (p. 791)**

The description of the pipeline.
Type: String
Length Constraints: Minimum length of 0. Maximum length of 3072.
Pattern: .*
Required: No

**PipelineDisplayName (p. 791)**

The display name of the pipeline.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: \^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,255}\{0,255\}
Required: No

**PipelineName (p. 791)**

The name of the pipeline to update.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: \^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,255}\{0,255\}
Required: Yes

**RoleArn (p. 791)**

The Amazon Resource Name (ARN) that the pipeline uses to execute.
Type: String
Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_]+$\+
Required: No

**Response Syntax**

```
{
  "PipelineArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**PipelineArn (p. 792)**

The Amazon Resource Name (ARN) of the updated pipeline.
Type: String
Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]\{12\}:pipeline/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ResourceNotFound

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdatePipelineExecution

Service: Amazon SageMaker Service

Updates a pipeline execution.

Request Syntax

```
{
    "ParallelismConfiguration": {
        "MaxParallelExecutionSteps": number
    },
    "PipelineExecutionArn": "string",
    "PipelineExecutionDescription": "string",
    "PipelineExecutionDisplayName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ParallelismConfiguration (p. 794)**

This configuration, if specified, overrides the parallelism configuration of the parent pipeline for this specific run.

Type: ParallelismConfiguration (p. 1247) object

Required: No

**PipelineExecutionArn (p. 794)**

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline\/[a-z0-9\-]+/.*$

Required: Yes

**PipelineExecutionDescription (p. 794)**

The description of the pipeline execution.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 3072.

Pattern: .*

Required: No

**PipelineExecutionDisplayName (p. 794)**

The display name of the pipeline execution.
Type: String

Length Constraints: Minimum length of 1. Maximum length of 82.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,81\}

Required: No

Response Syntax

```json
{
    "PipelineExecutionArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**PipelineExecutionArn (p. 795)**

The Amazon Resource Name (ARN) of the updated pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:pipeline\/.*/execution\/.*$

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
• AWS SDK for Ruby V3
UpdateProject
Service: Amazon SageMaker Service

Updates a machine learning (ML) project that is created from a template that sets up an ML pipeline from training to deploying an approved model.

Note
You must not update a project that is in use. If you update the ServiceCatalogProvisioningUpdateDetails of a project that is active or being created, or updated, you may lose resources already created by the project.

Request Syntax

```json
{
    "ProjectDescription": "string",
    "ProjectName": "string",
    "ServiceCatalogProvisioningUpdateDetails": {
        "ProvisioningArtifactId": "string",
        "ProvisioningParameters": [
            {
                "Key": "string",
                "Value": "string"
            }
        ]
    },
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**ProjectDescription (p. 797)**

The description for the project.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*
Required: No

**ProjectName (p. 797)**

The name of the project.
Type: String
Pattern: ^(a-zA-Z0-9)(-[a-zA-Z0-9]){0,31}$
UpdateProject

Required: Yes

**ServiceCatalogProvisioningUpdateDetails (p. 797)**

The product ID and provisioning artifact ID to provision a service catalog. The provisioning artifact ID will default to the latest provisioning artifact ID of the product, if you don't provide the provisioning artifact ID. For more information, see What is AWS Service Catalog.

Type: ServiceCatalogProvisioningUpdateDetails (p. 1364) object

Required: No

**Tags (p. 797)**

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Response Syntax**

```json
{
   "ProjectArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**ProjectArn (p. 798)**

The Amazon Resource Name (ARN) of the project.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:project:.*

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateTrainingJob
Service: Amazon SageMaker Service

Update a model training job to request a new Debugger profiling configuration.

Request Syntax

```json
{
   "ProfilerConfig": {
      "DisableProfiler": boolean,
      "ProfilingIntervalInMilliseconds": number,
      "ProfilingParameters": {
         "string": "string"
      },
      "S3OutputPath": "string"
   },
   "ProfilerRuleConfigurations": [
      {
         "InstanceType": "string",
         "LocalPath": "string",
         "RuleConfigurationName": "string",
         "RuleEvaluatorImage": "string",
         "RuleParameters": {
            "string": "string"
         },
         "S3OutputPath": "string",
         "VolumeSizeInGB": number
      }
   ],
   "TrainingJobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

ProfilerConfig (p. 800)

Configuration information for Debugger system monitoring, framework profiling, and storage paths.

Type: ProfilerConfigForUpdate (p. 1303) object

Required: No

ProfilerRuleConfigurations (p. 800)

Configuration information for Debugger rules for profiling system and framework metrics.

Type: Array of ProfilerRuleConfiguration (p. 1305) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

Required: No

TrainingJobName (p. 800)

The name of a training job to update the Debugger profiling configuration.

Type: String

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

Required: Yes

**Response Syntax**

```json
{
  "TrainingJobArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**TrainingJobArn (p. 801)**

The Amazon Resource Name (ARN) of the training job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:training-job/.*

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](https://docs.aws.amazon.com/AmazonSageMaker/latest/APIReference/Errors.html).

**ResourceNotFound**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateTrial
Service: Amazon SageMaker Service

Updates the display name of a trial.

Request Syntax

```json
{
    "DisplayName": "string",
    "TrialName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DisplayName (p. 802)**

The name of the trial as displayed. The name doesn't need to be unique. If DisplayName isn't specified, TrialName is displayed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: No

**TrialName (p. 802)**

The name of the trial to update.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,119}$

Required: Yes

Response Syntax

```json
{
    "TrialArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
**TrialArn (p. 802)**

The Amazon Resource Name (ARN) of the trial.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial/.*`

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ConflictException**

There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.

HTTP Status Code: 400

**ResourceNotFoundException**

Resource being access is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateTrialComponent
Service: Amazon SageMaker Service

Updates one or more properties of a trial component.

Request Syntax

```
{
    "DisplayName": "string",
    "EndTime": number,
    "InputArtifacts": {
        "string": {
            "MediaType": "string",
            "Value": "string"
        }
    },
    "InputArtifactsToRemove": [ "string" ],
    "OutputArtifacts": {
        "string": {
            "MediaType": "string",
            "Value": "string"
        }
    },
    "OutputArtifactsToRemove": [ "string" ],
    "Parameters": {
        "string": {
            "NumberValue": number,
            "StringValue": "string"
        }
    },
    "ParametersToRemove": [ "string" ],
    "StartTime": number,
    "Status": {
        "Message": "string",
        "PrimaryStatus": "string"
    },
    "TrialComponentName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

DisplayName (p. 804)

The name of the component as displayed. The name doesn't need to be unique. If DisplayName isn't specified, TrialComponentName is displayed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]{0,119}

Required: No

EndTime (p. 804)

When the component ended.
Type: Timestamp
Required: No

**InputArtifacts (p. 804)**

Replaces all of the component's input artifacts with the specified artifacts.

Type: String to `TrialComponentArtifact (p. 1428)` object map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 64.

Key Pattern: . *

Required: No

**InputArtifactsToRemove (p. 804)**

The input artifacts to remove from the component.

Type: Array of strings

Length Constraints: Maximum length of 256.

Pattern: . *

Required: No

**OutputArtifacts (p. 804)**

Replaces all of the component's output artifacts with the specified artifacts.

Type: String to `TrialComponentArtifact (p. 1428)` object map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 64.

Key Pattern: . *

Required: No

**OutputArtifactsToRemove (p. 804)**

The output artifacts to remove from the component.

Type: Array of strings

Length Constraints: Maximum length of 256.

Pattern: . *

Required: No

**Parameters (p. 804)**

Replaces all of the component's hyperparameters with the specified hyperparameters.

Type: String to `TrialComponentParameterValue (p. 1431)` object map

Map Entries: Maximum number of 150 items.

Key Length Constraints: Maximum length of 256.
Key Pattern: .*
Required: No

**ParametersToRemove (p. 804)**
The hyperparameters to remove from the component.
Type: Array of strings
Length Constraints: Maximum length of 256.
Pattern: .*
Required: No

**StartTime (p. 804)**
When the component started.
Type: Timestamp
Required: No

**Status (p. 804)**
The new status of the component.
Type: TrialComponentStatus (p. 1436) object
Required: No

**TrialComponentName (p. 804)**
The name of the component to update.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$
Required: Yes

**Response Syntax**

```json
{
   "TrialComponentArn": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**TrialComponentArn (p. 806)**
The Amazon Resource Name (ARN) of the trial component.
Type: String
Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial-component/.*

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

ConflictException

There was a conflict when you attempted to modify a SageMaker entity such as an Experiment or Artifact.

HTTP Status Code: 400

ResourceNotFoundException

Resource being access is not found.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateUserProfile

Service: Amazon SageMaker Service

Updates a user profile.

Request Syntax

```json
{  
  "DomainId": "string",
  "UserProfileName": "string",
  "UserSettings": {  
    "ExecutionRole": "string",
    "JupyterServerAppSettings": { 
      "DefaultResourceSpec": { 
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      },
      "LifecycleConfigArns": [ "string" ]
    },
    "KernelGatewayAppSettings": { 
      "CustomImages": [ 
        { 
          "AppImageConfigName": "string",
          "ImageName": "string",
          "ImageVersionNumber": number
        }
      ],
      "DefaultResourceSpec": { 
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      },
      "LifecycleConfigArns": [ "string" ]
    },
    "RSessionAppSettings": { 
      "CustomImages": [ 
        { 
          "AppImageConfigName": "string",
          "ImageName": "string",
          "ImageVersionNumber": number
        }
      ],
      "DefaultResourceSpec": { 
        "InstanceType": "string",
        "LifecycleConfigArn": "string",
        "SageMakerImageArn": "string",
        "SageMakerImageVersionArn": "string"
      }
    },
    "RStudioServerProAppSettings": { 
      "AccessStatus": "string",
      "UserGroup": "string"
    },
    "SecurityGroups": [ "string" ],
    "SharingSettings": { 
      "NotebookOutputOption": "string",
      "S3KmsKeyId": "string",
      "S3OutputPath": "string"
    },
    "TensorBoardAppSettings": {
  }
}
```
"DefaultResourceSpec": {
    "InstanceType": "string",
    "LifecycleConfigArn": "string",
    "SageMakerImageArn": "string",
    "SageMakerImageVersionArn": "string"
  }
}
}

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**DomainId (p. 808)**

The domain ID.

Type: String

Length Constraints: Maximum length of 63.

Required: Yes

**UserProfileName (p. 808)**

The user profile name.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

Required: Yes

**UserSettings (p. 808)**

A collection of settings.

Type: UserSettings (p. 1452) object

Required: No

Response Syntax

```json
{
  "UserProfileArn": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.
**UserProfileArn (p. 809)**

The user profile Amazon Resource Name (ARN).

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:user-profile/.*`

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)](#).

**ResourceInUse**

Resource being accessed is in use.

HTTP Status Code: 400

**ResourceLimitExceeded**

You have exceeded a SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

**ResourceNotFound**

Resource being accessed is not found.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateWorkforce

Service: Amazon SageMaker Service

Use this operation to update your workforce. You can use this operation to require that workers use specific IP addresses to work on tasks and to update your OpenID Connect (OIDC) Identity Provider (IdP) workforce configuration.

Use SourceIpConfig to restrict worker access to tasks to a specific range of IP addresses. You specify allowed IP addresses by creating a list of up to ten CIDRs. By default, a workforce isn't restricted to specific IP addresses. If you specify a range of IP addresses, workers who attempt to access tasks using any IP address outside the specified range are denied and get a Not Found error message on the worker portal.

Use OidcConfig to update the configuration of a workforce created using your own OIDC IdP.

**Important**
You can only update your OIDC IdP configuration when there are no work teams associated with your workforce. You can delete work teams using the DeleteWorkteam (p. 275) operation.

After restricting access to a range of IP addresses or updating your OIDC IdP configuration with this operation, you can view details about your update workforce using the DescribeWorkforce (p. 476) operation.

**Important**
This operation only applies to private workforces.

**Request Syntax**

```json
{
    "OidcConfig": {
        "AuthorizationEndpoint": "string",
        "ClientId": "string",
        "ClientSecret": "string",
        "Issuer": "string",
        "JwksUri": "string",
        "LogoutEndpoint": "string",
        "TokenEndpoint": "string",
        "UserInfoEndpoint": "string"
    },
    "SourceIpConfig": {
        "Cidrs": [ "string" ]
    },
    "WorkforceName": "string"
}
```

**Request Parameters**

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

**OidcConfig (p. 811)**

Use this parameter to update your OIDC Identity Provider (IdP) configuration for a workforce made using your own IdP.

Type: OidcConfig (p. 1233) object

Required: No
**SourceIpConfig (p. 811)**

A list of one to ten worker IP address ranges (CIDRs) that can be used to access tasks assigned to this workforce.

- **Maximum:** Ten CIDR values
- **Type:** SourceIpConfig (p. 1369) object
- **Required:** No

**WorkforceName (p. 811)**

The name of the private workforce that you want to update. You can find your workforce name by using the ListWorkforces (p. 671) operation.

- **Type:** String
- **Length Constraints:** Minimum length of 1. Maximum length of 63.
- **Pattern:** ^[a-zA-Z0-9-]{0,62}$
- **Required:** Yes

**Response Syntax**

```json
{
    "Workforce": {
        "CognitoConfig": {
            "ClientId": "string",
            "UserPool": "string"
        },
        "CreateDate": number,
        "LastUpdatedDate": number,
        "OidcConfig": {
            "AuthorizationEndpoint": "string",
            "ClientId": "string",
            "Issuer": "string",
            "JwksUri": "string",
            "LogoutEndpoint": "string",
            "TokenEndpoint": "string",
            "UserInfoEndpoint": "string"
        },
        "SourceIpConfig": {
            "Cidrs": [ "string" ]
        },
        "SubDomain": "string",
        "WorkforceArn": "string",
        "WorkforceName": "string"
    }
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**Workforce (p. 812)**

A single private workforce. You can create one private workforce in each AWS Region. By default, any workforce-related API operation used in a specific region will apply to the workforce created in that region. To learn how to create a private workforce, see Create a Private Workforce.
Type: **Workforce (p. 1457) object**

**Errors**

For information about the errors that are common to all actions, see [Common Errors (p. 1472)].

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
UpdateWorkteam
Service: Amazon SageMaker Service

Updates an existing work team with new member definitions or description.

Request Syntax

```json
{
    "Description": "string",
    "MemberDefinitions": [
        {
            "CognitoMemberDefinition": {
                "ClientId": "string",
                "UserGroup": "string",
                "UserPool": "string"
            },
            "OidcMemberDefinition": {
                "Groups": [ "string" ]
            }
        }
    ],
    "NotificationConfiguration": {
        "NotificationTopicArn": "string"
    },
    "WorkteamName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters (p. 1470).

The request accepts the following data in JSON format.

Description (p. 814)

An updated description for the work team.

Type: String


Pattern: .+

Required: No

MemberDefinitions (p. 814)

A list of MemberDefinition objects that contains objects that identify the workers that make up the work team.

Workforces can be created using Amazon Cognito or your own OIDC Identity Provider (IdP). For private workforces created using Amazon Cognito use CognitoMemberDefinition. For workforces created using your own OIDC identity provider (IdP) use OidcMemberDefinition. You should not provide input for both of these parameters in a single request.

For workforces created using Amazon Cognito, private work teams correspond to Amazon Cognito user groups within the user pool used to create a workforce. All of the CognitoMemberDefinition objects that make up the member definition must have the same ClientId and UserPool values.
To add an Amazon Cognito user group to an existing worker pool, see *Adding groups to a User Pool.* For more information about user pools, see *Amazon Cognito User Pools.*

For workforces created using your own OIDC IdP, specify the user groups that you want to include in your private work team by listing those groups in `Groups` of `OidcMemberDefinition`. Be aware that user groups that are already in the work team must also be listed in `Groups` when you make this request to remain on the work team. If you do not include these user groups, they will no longer be associated with the work team you update.

**Type:** Array of MemberDefinition (p. 1142) objects

**Array Members:** Minimum number of 1 item. Maximum number of 10 items.

**Required:** No

**NotificationConfiguration (p. 814)**

Configures SNS topic notifications for available or expiring work items

**Type:** NotificationConfiguration (p. 1229) object

**Required:** No

**WorkteamName (p. 814)**

The name of the work team to update.

**Type:** String

**Length Constraints:** Minimum length of 1. Maximum length of 63.

**Pattern:** `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$`

**Required:** Yes

---

**Response Syntax**

```json
{
  "Workteam": {
    "CreateDate": number,
    "Description": "string",
    "LastUpdatedDate": number,
    "MemberDefinitions": [
      {
        "CognitoMemberDefinition": {
          "ClientId": "string",
          "UserGroup": "string",
          "UserPool": "string"
        },
        "OidcMemberDefinition": {
          "Groups": [ "string" ]
        }
      }
    ],
    "NotificationConfiguration": {
      "NotificationTopicArn": "string"
    },
    "ProductListingIds": [ "string" ],
    "SubDomain": "string",
    "WorkforceArn": "string",
    "WorkteamArn": "string",
    "WorkteamName": "string"
  }
}
```
Response Elements

If the action is successful, the service sends back an HTTP 200 response. The following data is returned in JSON format by the service.

**Workteam (p. 815)**

A `Workteam` object that describes the updated work team.

Type: `Workteam (p. 1459)` object

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**ResourceLimitExceeded**

You have exceeded an SageMaker resource limit. For example, you might have too many training jobs created.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

Amazon SageMaker Runtime

The following actions are supported by Amazon SageMaker Runtime:

- **InvokeEndpoint (p. 817)**
- **InvokeEndpointAsync (p. 823)**
InvokeEndpoint

Service: Amazon SageMaker Runtime

After you deploy a model into production using Amazon SageMaker hosting services, your client applications use this API to get inferences from the model hosted at the specified endpoint.

For an overview of Amazon SageMaker, see How It Works.

Amazon SageMaker strips all POST headers except those supported by the API. Amazon SageMaker might add additional headers. You should not rely on the behavior of headers outside those enumerated in the request syntax.

Calls to InvokeEndpoint are authenticated by using AWS Signature Version 4. For information, see Authenticating Requests (AWS Signature Version 4) in the Amazon S3 API Reference.

A customer's model containers must respond to requests within 60 seconds. The model itself can have a maximum processing time of 60 seconds before responding to invocations. If your model is going to take 50-60 seconds of processing time, the SDK socket timeout should be set to be 70 seconds.

Note
Endpoints are scoped to an individual account, and are not public. The URL does not contain the account ID, but Amazon SageMaker determines the account ID from the authentication token that is supplied by the caller.

Request Syntax

POST /endpoints/EndpointName/invocations HTTP/1.1
Content-Type: Content-Type
Accept: Accept
X-Amzn-SageMaker-Custom-Attributes: CustomAttributes
X-Amzn-SageMaker-Target-Model: TargetModel
X-Amzn-SageMaker-Target-Variant: TargetVariant
X-Amzn-SageMaker-Target-Container-Hostname: TargetContainerHostname
X-Amzn-SageMaker-Inference-Id: InferenceId

Body

URI Request Parameters

The request uses the following URI parameters.

Accept (p. 817)
The desired MIME type of the inference in the response.
Length Constraints: Maximum length of 1024.
Pattern: \p{ASCII}*

ContentType (p. 817)
The MIME type of the input data in the request body.
Length Constraints: Maximum length of 1024.
Pattern: \p{ASCII}*

CustomAttributes (p. 817)
Provides additional information about a request for an inference submitted to a model hosted at an Amazon SageMaker endpoint. The information is an opaque value that is forwarded verbatim. You
could use this value, for example, to provide an ID that you can use to track a request or to provide other metadata that a service endpoint was programmed to process. The value must consist of no more than 1024 visible US-ASCII characters as specified in Section 3.3.6. Field Value Components of the Hypertext Transfer Protocol (HTTP/1.1).

The code in your model is responsible for setting or updating any custom attributes in the response. If your code does not set this value in the response, an empty value is returned. For example, if a custom attribute represents the trace ID, your model can prepend the custom attribute with `Trace ID:` in your post-processing function.

This feature is currently supported in the AWS SDKs but not in the Amazon SageMaker Python SDK.

Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

**EndpointName (p. 817)**

The name of the endpoint that you specified when you created the endpoint using the `CreateEndpoint` API.

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]*)*

Required: Yes

**InferenceId (p. 817)**

If you provide a value, it is added to the captured data when you enable data capture on the endpoint. For information about data capture, see Capture Data.

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: \A\S[\p{Print}]\z

**TargetContainerHostname (p. 817)**

If the endpoint hosts multiple containers and is configured to use direct invocation, this parameter specifies the host name of the container to invoke.

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]*)*

**TargetModel (p. 817)**

The model to request for inference when invoking a multi-model endpoint.


Pattern: \A\S[\p{Print}]\z

**TargetVariant (p. 817)**

Specify the production variant to send the inference request to when invoking an endpoint that is running two or more variants. Note that this parameter overrides the default behavior for the endpoint, which is to distribute the invocation traffic based on the variant weights.

For information about how to use variant targeting to perform a/b testing, see Test models in production

Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*  

Request Body

The request accepts the following binary data.

Body (p. 817)

Provides input data, in the format specified in the ContentType request header. Amazon SageMaker passes all of the data in the body to the model.

For information about the format of the request body, see Common Data Formats-Inference.

Length Constraints: Maximum length of 6291456.

Required: Yes

Response Syntax

HTTP/1.1 200
Content-Type: ContentType
x-Amzn-Invoked-Production-Variant: InvokedProductionVariant
X-Amzn-SageMaker-Custom-Attributes: CustomAttributes

Body

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The response returns the following HTTP headers.

ContentType (p. 819)

The MIME type of the inference returned in the response body.

Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

CustomAttributes (p. 819)

Provides additional information in the response about the inference returned by a model hosted at an Amazon SageMaker endpoint. The information is an opaque value that is forwarded verbatim. You could use this value, for example, to return an ID received in the CustomAttributes header of a request or other metadata that a service endpoint was programmed to produce. The value must consist of no more than 1024 visible US-ASCII characters as specified in Section 3.3.6. Field Value Components of the Hypertext Transfer Protocol (HTTP/1.1). If the customer wants the custom attribute returned, the model must set the custom attribute to be included on the way back.

The code in your model is responsible for setting or updating any custom attributes in the response. If your code does not set this value in the response, an empty value is returned. For example, if a custom attribute represents the trace ID, your model can prepend the custom attribute with Trace ID: in your post-processing function.

This feature is currently supported in the AWS SDKs but not in the Amazon SageMaker Python SDK.
Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

**InvokedProductionVariant (p. 819)**

Identifies the production variant that was invoked.

Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

The response returns the following as the HTTP body.

**Body (p. 819)**

Includes the inference provided by the model.

For information about the format of the response body, see Common Data Formats-Inference.

Length Constraints: Maximum length of 6291456.

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**InternalDependencyException**

Your request caused an exception with an internal dependency. Contact customer support.

HTTP Status Code: 530

**InternalFailure**

An internal failure occurred.

HTTP Status Code: 500

**ModelError**

Model (owned by the customer in the container) returned 4xx or 5xx error code.

HTTP Status Code: 424

**ModelNotReadyException**

Either a serverless endpoint variant's resources are still being provisioned, or a multi-model endpoint is still downloading or loading the target model. Wait and try your request again.

HTTP Status Code: 429

**ServiceUnavailable**

The service is unavailable. Try your call again.

HTTP Status Code: 503

**ValidationError**

Inspect your request and try again.

HTTP Status Code: 400
Examples

Pass a trace ID in the CustomAttribute of a request and return it in the CustomAttribute of the response.

In this example a trace ID is passed to the service endpoint in the CustomAttributes header of the request and then retrieved and returned in the CustomAttributes header of the response.

Sample Request

```python
import boto3
client = boto3.client('sagemaker-runtime')
custom_attributes = "c000b4f9-df62-4c85-a0bf-7c525f9104a4"  # An example of a trace ID.
endpoint_name = "..."                                       # Your endpoint name.
content_type = "..."                                        # The MIME type of the input
data in the request body.
accept = "..."                                              # The desired MIME type of the inference in the response.
payload = "..."                                             # Payload for inference.
response = client.invoke_endpoint(
    EndpointName=endpoint_name,
   CustomAttributes=custom_attributes,  # If model receives and updates
    ContentType=content_type,          # by adding "Trace: " in
    Accept=accept,                    # front of custom_attributes in the request,
    Body=payload                      # becomes
)
print(response['CustomAttributes'])                         # custom_attributes in response
```

Sample Response

```
Trace ID: c000b4f9-df62-4c85-a0bf-7c525f9104a4
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
InvokeEndpointAsync
Service: Amazon SageMaker Runtime

After you deploy a model into production using Amazon SageMaker hosting services, your client applications use this API to get inferences from the model hosted at the specified endpoint in an asynchronous manner.

Inference requests sent to this API are enqueued for asynchronous processing. The processing of the inference request may or may not complete before you receive a response from this API. The response from this API will not contain the result of the inference request but contain information about where you can locate it.

Amazon SageMaker strips all POST headers except those supported by the API. Amazon SageMaker might add additional headers. You should not rely on the behavior of headers outside those enumerated in the request syntax.

Calls to InvokeEndpointAsync are authenticated by using AWS Signature Version 4. For information, see Authenticating Requests (AWS Signature Version 4) in the Amazon S3 API Reference.

Request Syntax

<table>
<thead>
<tr>
<th>POST /endpoints/EndpointName/async-invocations HTTP/1.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-Amzn-SageMaker-Content-Type: ContentType</td>
</tr>
<tr>
<td>X-Amzn-SageMaker-Accept: Accept</td>
</tr>
<tr>
<td>X-Amzn-SageMaker-Custom-Attributes: CustomAttributes</td>
</tr>
<tr>
<td>X-Amzn-SageMaker-Inference-Id: InferenceId</td>
</tr>
<tr>
<td>X-Amzn-SageMaker-InputLocation: InputLocation</td>
</tr>
<tr>
<td>X-Amzn-SageMaker-RequestTTLSeconds: RequestTTLSeconds</td>
</tr>
</tbody>
</table>

URI Request Parameters

The request uses the following URI parameters.

Accept (p. 823)

The desired MIME type of the inference in the response.

Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

ContentType (p. 823)

The MIME type of the input data in the request body.

Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

CustomAttributes (p. 823)

Provides additional information about a request for an inference submitted to a model hosted at an Amazon SageMaker endpoint. The information is an opaque value that is forwarded verbatim. You could use this value, for example, to provide an ID that you can use to track a request or to provide other metadata that a service endpoint was programmed to process. The value must consist of no more than 1024 visible US-ASCII characters as specified in Section 3.3.6. Field Value Components of the Hypertext Transfer Protocol (HTTP/1.1).

The code in your model is responsible for setting or updating any custom attributes in the response. If your code does not set this value in the response, an empty value is returned. For example, if a
custom attribute represents the trace ID, your model can prepend the custom attribute with `Trace ID:` in your post-processing function.

This feature is currently supported in the AWS SDKs but not in the Amazon SageMaker Python SDK.

Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

**EndpointName (p. 823)**

The name of the endpoint that you specified when you created the endpoint using the `CreateEndpoint` API.

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](\-[a-zA-Z0-9])*\$

Required: Yes

**InferenceId (p. 823)**

The identifier for the inference request. Amazon SageMaker will generate an identifier for you if none is specified.

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: \A\S[\p{Print}]\z

**InputLocation (p. 823)**

The Amazon S3 URI where the inference request payload is stored.


Pattern: ^(https|s3)://([^/]+)/?\*(.*)$

Required: Yes

**RequestTTLSeconds (p. 823)**

Maximum age in seconds a request can be in the queue before it is marked as expired.

Valid Range: Minimum value of 60. Maximum value of 21600.

**Request Body**

The request does not have a request body.

**Response Syntax**

```
HTTP/1.1 202
X-Amzn-SageMaker-OutputLocation: OutputLocation
Content-type: application/json

{
  "InferenceId": "string"
}
```

**Response Elements**

If the action is successful, the service sends back an HTTP 202 response.
The response returns the following HTTP headers.

**OutputLocation (p. 824)**

The Amazon S3 URI where the inference response payload is stored.

Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

The following data is returned in JSON format by the service.

**InferenceId (p. 824)**

Identifier for an inference request. This will be the same as the `InferenceId` specified in the input. Amazon SageMaker will generate an identifier for you if you do not specify one.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: \p{ASCII}*

**Errors**

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**InternalFailure**

An internal failure occurred.

HTTP Status Code: 500

**ServiceUnavailable**

The service is unavailable. Try your call again.

HTTP Status Code: 503

**ValidationError**

Inspect your request and try again.

HTTP Status Code: 400

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
• AWS SDK for Ruby V3

Amazon SageMaker Edge Manager

The following actions are supported by Amazon SageMaker Edge Manager:

• GetDeviceRegistration (p. 827)
• SendHeartbeat (p. 829)
GetDeviceRegistration
Service: Amazon Sagemaker Edge Manager
Use to check if a device is registered with SageMaker Edge Manager.

Request Syntax

POST /GetDeviceRegistration HTTP/1.1
Content-type: application/json

{
  "DeviceFleetName": "string",
  "DeviceName": "string"
}

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

DeviceFleetName (p. 827)
  The name of the fleet that the device belongs to.
  Type: String
  Pattern: ^[a-zA-Z0-9](-*_*[a-zA-Z0-9])*$
  Required: Yes

DeviceName (p. 827)
  The unique name of the device you want to get the registration status from.
  Type: String
  Pattern: ^[a-zA-Z0-9](-*_*[a-zA-Z0-9])*$
  Required: Yes

Response Syntax

HTTP/1.1 200
Content-type: application/json

{
  "CacheTTL": "string",
  "DeviceRegistration": "string"
}
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CacheTTL (p. 827)

The amount of time, in seconds, that the registration status is stored on the device's cache before it is refreshed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1000.

DeviceRegistration (p. 827)

Describes if the device is currently registered with SageMaker Edge Manager.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1000.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

InternalServiceException

An internal failure occurred. Try your request again. If the problem persists, contact AWS customer support.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
SendHeartbeat
Service: Amazon Sagemaker Edge Manager

Use to get the current status of devices registered on SageMaker Edge Manager.

Request Syntax

```
POST /SendHeartbeat HTTP/1.1
Content-type: application/json
{
  "AgentMetrics": [
    {
      "Dimension": "string",
      "MetricName": "string",
      "Timestamp": number,
      "Value": number
    }
  ],
  "AgentVersion": "string",
  "DeviceFleetName": "string",
  "DeviceName": "string",
  "Models": [
    {
      "LatestInference": number,
      "LatestSampleTime": number,
      "ModelMetrics": [
        {
          "Dimension": "string",
          "MetricName": "string",
          "Timestamp": number,
          "Value": number
        }
      ],
      "ModelName": "string",
      "ModelVersion": "string"
    }
  ]
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

**AgentMetrics (p. 829)**

For internal use. Returns a list of SageMaker Edge Manager agent operating metrics.

Type: Array of EdgeMetric (p. 1462) objects

Required: No

**AgentVersion (p. 829)**

Returns the version of the agent.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 64.
Pattern: [a-zA-Z0-9\ \_\.]+
Required: Yes

DeviceFleetName (p. 829)
The name of the fleet that the device belongs to.
Type: String
Pattern: ^[a-zA-Z0-9](-*[_*][a-zA-Z0-9])*$
Required: Yes

DeviceName (p. 829)
The unique name of the device.
Type: String
Pattern: ^[a-zA-Z0-9](-*[_*][a-zA-Z0-9])*$
Required: Yes

Models (p. 829)
Returns a list of models deployed on the the device.
Type: Array of Model (p. 1463) objects
Required: No

Response Syntax

HTTP/1.1 200

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

InternalServiceException

An internal failure occurred. Try your request again. If the problem persists, contact AWS customer support.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS Command Line Interface
• AWS SDK for .NET
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for JavaScript
• AWS SDK for PHP V3
• AWS SDK for Python
• AWS SDK for Ruby V3

Amazon SageMaker Feature Store Runtime

The following actions are supported by Amazon SageMaker Feature Store Runtime:

• BatchGetRecord (p. 832)
• DeleteRecord (p. 835)
• GetRecord (p. 837)
• PutRecord (p. 840)
BatchGetRecord

Service: Amazon SageMaker Feature Store Runtime

Retrieves a batch of Records from a FeatureGroup.

Request Syntax

POST /BatchGetRecord HTTP/1.1
Content-type: application/json

{
   "Identifiers": [
      {
         "FeatureGroupName": "string",
         "FeatureNames": [ "string" ],
         "RecordIdentifiersValueAsString": [ "string" ]
      }
   ]
}

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

Identifiers (p. 832)

A list of FeatureGroup names, with their corresponding RecordIdentifier value, and Feature name that have been requested to be retrieved in batch.

Type: Array of BatchGetRecordIdentifier (p. 1467) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Required: Yes

Response Syntax

HTTP/1.1 200
Content-type: application/json

{
   "Errors": [
      {
         "ErrorCode": "string",
         "ErrorMessage": "string",
         "FeatureGroupName": "string",
         "RecordIdentifierValueAsString": "string"
      }
   ],
   "Records": [
      {
         "FeatureGroupName": "string",
         "Record": [
            "string"
          ]
      }
   ]
}
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors (p. 832)

A list of errors that have occurred when retrieving a batch of Records.

Type: Array of BatchGetRecordError (p. 1465) objects

Array Members: Minimum number of 0 items.

Records (p. 832)

A list of Records you requested to be retrieved in batch.

Type: Array of BatchGetRecordResultDetail (p. 1468) objects

Array Members: Minimum number of 0 items.

UnprocessedIdentifiers (p. 832)

A unprocessed list of FeatureGroup names, with their corresponding RecordIdentifier value, and Feature name.

Type: Array of BatchGetRecordIdentifier (p. 1467) objects

Array Members: Minimum number of 0 items.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

AccessForbidden

You do not have permission to perform an action.

HTTP Status Code: 403

InternalFailure

An internal failure occurred. Try your request again. If the problem persists, contact AWS customer support.
HTTP Status Code: 500

**ServiceUnavailable**

The service is currently unavailable.

HTTP Status Code: 503

**ValidationError**

There was an error validating your request.

HTTP Status Code: 400

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
**DeleteRecord**
Service: Amazon SageMaker Feature Store Runtime

Deletes a Record from a FeatureGroup. When the DeleteRecord API is called a new record will be added to the OfflineStore and the Record will be removed from the OnlineStore. This record will have a value of True in the is_deleted column.

**Request Syntax**

```
DELETE /FeatureGroup/FeatureGroupName?
EventTime=EventTime&RecordIdentifierValueAsString=RecordIdentifierValueAsString HTTP/1.1
```

**URI Request Parameters**

The request uses the following URI parameters.

**EventTime** (p. 835)

Timestamp indicating when the deletion event occurred. EventTime can be used to query data at a certain point in time.

Length Constraints: Maximum length of 358400.

Pattern: .*

Required: Yes

**FeatureGroupName** (p. 835)

The name of the feature group to delete the record from.

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*`

Required: Yes

**RecordIdentifierValueAsString** (p. 835)

The value for the RecordIdentifier that uniquely identifies the record, in string format.

Length Constraints: Maximum length of 358400.

Pattern: .*

Required: Yes

**Request Body**

The request does not have a request body.

**Response Syntax**

```
HTTP/1.1 200
```

**Response Elements**

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.
Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

AccessForbidden

You do not have permission to perform an action.

HTTP Status Code: 403

InternalFailure

An internal failure occurred. Try your request again. If the problem persists, contact AWS customer support.

HTTP Status Code: 500

ServiceUnavailable

The service is currently unavailable.

HTTP Status Code: 503

ValidationError

There was an error validating your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
GetRecord

Service: Amazon SageMaker Feature Store Runtime

Use for OnlineStore serving from a FeatureStore. Only the latest records stored in the OnlineStore can be retrieved. If no Record with RecordIdentifierValue is found, then an empty result is returned.

Request Syntax

GET /FeatureGroup/FeatureGroupName?
FeatureName=FeatureNames&RecordIdentifierValueAsString=RecordIdentifierValueAsString
HTTP/1.1

URI Request Parameters

The request uses the following URI parameters.

FeatureGroupName (p. 837)

The name of the feature group from which you want to retrieve a record.

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])* 

Required: Yes

FeatureNames (p. 837)

List of names of Features to be retrieved. If not specified, the latest value for all the Features are returned.

Array Members: Minimum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9][-_]*[a-zA-Z0-9]]*

RecordIdentifierValueAsString (p. 837)

The value that corresponds to RecordIdentifier type and uniquely identifies the record in the FeatureGroup.

Length Constraints: Maximum length of 358400.

Pattern: .*

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

HTTP/1.1 200
Content-type: application/json
Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

**Record (p. 837)**

The record you requested. A list of `FeatureValues`.

Type: Array of `FeatureValue (p. 1469)` objects

Array Members: Minimum number of 1 item.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 1472).

**AccessForbidden**

You do not have permission to perform an action.

HTTP Status Code: 403

**InternalFailure**

An internal failure occurred. Try your request again. If the problem persists, contact AWS customer support.

HTTP Status Code: 500

**ResourceNotFound**

A resource that is required to perform an action was not found.

HTTP Status Code: 404

**ServiceUnavailable**

The service is currently unavailable.

HTTP Status Code: 503

**ValidationError**

There was an error validating your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
PutRecord
Service: Amazon SageMaker Feature Store Runtime

Used for data ingestion into the FeatureStore. The PutRecord API writes to both the OnlineStore and OfflineStore. If the record is the latest record for the recordIdentifier, the record is written to both the OnlineStore and OfflineStore. If the record is a historic record, it is written only to the OfflineStore.

Request Syntax

```
PUT /FeatureGroup/<FeatureGroupName> HTTP/1.1
Content-type: application/json

{
    "Record": [
        {
            "FeatureName": "string",
            "ValueAsString": "string"
        }
    ]
}
```

URI Request Parameters

The request uses the following URI parameters.

**FeatureGroupName (p. 840)**

The name of the feature group that you want to insert the record into.

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*  

Required: Yes

Request Body

The request accepts the following data in JSON format.

**Record (p. 840)**

List of FeatureValues to be inserted. This will be a full over-write. If you only want to update few of the feature values, do the following:

- Use GetRecord to retrieve the latest record.
- Update the record returned from GetRecord.
- Use PutRecord to update feature values.

Type: Array of FeatureValue (p. 1469) objects

Array Members: Minimum number of 1 item.

Required: Yes
Response Syntax

HTTP/1.1 200

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see CommonErrors (p. 1472).

AccessForbidden
You do not have permission to perform an action.

HTTP Status Code: 403

InternalFailure
An internal failure occurred. Try your request again. If the problem persists, contact AWS customer support.

HTTP Status Code: 500

ServiceUnavailable
The service is currently unavailable.

HTTP Status Code: 503

ValidationError
There was an error validating your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface
- AWS SDK for .NET
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for JavaScript
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3
Data Types

The following data types are supported by Amazon SageMaker Service:

- ActionSource (p. 860)
- ActionSummary (p. 861)
- AdditionalInferenceSpecificationDefinition (p. 863)
- AgentVersion (p. 865)
- Alarm (p. 866)
- AlgorithmSpecification (p. 867)
- AlgorithmStatusDetails (p. 870)
- AlgorithmStatusItem (p. 871)
- AlgorithmSummary (p. 872)
- AlgorithmValidationProfile (p. 874)
- AlgorithmValidationSpecification (p. 875)
- AnnotationConsolidationConfig (p. 876)
- AppDetails (p. 887)
- AppImageConfigDetails (p. 889)
- AppSpecification (p. 891)
- ArtifactSource (p. 892)
- ArtifactSourceType (p. 893)
- ArtifactSummary (p. 894)
- AssociationSummary (p. 896)
- AsyncInferenceClientConfig (p. 898)
- AsyncInferenceConfig (p. 899)
- AsyncInferenceNotificationConfig (p. 900)
- AsyncInferenceOutputConfig (p. 901)
- AthenaDatasetDefinition (p. 902)
- AutoMLCandidate (p. 904)
- AutoMLCandidateGenerationConfig (p. 906)
- AutoMLCandidateStep (p. 907)
- AutoMLChannel (p. 908)
- AutoMLContainerDefinition (p. 910)
- AutoMLDataSource (p. 912)
- AutoMLDataSplitConfig (p. 913)
- AutoMLJobArtifacts (p. 914)
- AutoMLJobCompletionCriteria (p. 915)
- AutoMLJobConfig (p. 916)
- AutoMLJobObjective (p. 917)
- AutoMLJobSummary (p. 919)
- AutoMLOutputDataConfig (p. 921)
- AutoMLPartialFailureReason (p. 922)
- AutoMLS3DataSource (p. 923)
- AutoMLSecurityConfig (p. 924)
• AutoRollbackConfig (p. 925)
• BatchDescribeModelPackageError (p. 926)
• BatchDescribeModelPackageSummary (p. 927)
• Bias (p. 929)
• BlueGreenUpdatePolicy (p. 930)
• CacheHitResult (p. 931)
• CallbackStepMetadata (p. 932)
• CandidateArtifactLocations (p. 933)
• CandidateProperties (p. 934)
• CapacitySize (p. 935)
• CaptureContentTypeHeader (p. 936)
• CaptureOption (p. 937)
• CategoricalParameter (p. 938)
• CategoricalParameterRange (p. 939)
• CategoricalParameterRangeSpecification (p. 940)
• Channel (p. 941)
• ChannelSpecification (p. 943)
• CheckpointConfig (p. 945)
• ClarifyCheckStepMetadata (p. 946)
• CodeRepositorySummary (p. 948)
• CognitoConfig (p. 950)
• CognitoMemberDefinition (p. 951)
• CollectionConfiguration (p. 952)
• CompilationJobSummary (p. 953)
• ConditionStepMetadata (p. 955)
• ContainerDefinition (p. 956)
• ContextSource (p. 959)
• ContextSummary (p. 960)
• ContinuousParameterRange (p. 962)
• ContinuousParameterRangeSpecification (p. 964)
• CustomImage (p. 965)
• DataCaptureConfig (p. 966)
• DataCaptureConfigSummary (p. 968)
• DataCatalogConfig (p. 970)
• DataProcessing (p. 971)
• DataQualityAppSpecification (p. 973)
• DataQualityBaselineConfig (p. 975)
• DataQualityJobInput (p. 976)
• DatasetDefinition (p. 977)
• DataSource (p. 979)
• DebugHookConfig (p. 980)
• DebugRuleConfiguration (p. 982)
• DebugRuleEvaluationStatus (p. 984)
• DeployedImage (p. 986)
• DeploymentConfig (p. 987)
• DesiredWeightAndCapacity (p. 988)
- Device (p. 989)
- DeviceFleetSummary (p. 990)
- DeviceStats (p. 991)
- DeviceSummary (p. 992)
- DomainDetails (p. 994)
- DomainSettings (p. 996)
- DomainSettingsForUpdate (p. 997)
- DriftCheckBaselines (p. 998)
- DriftCheckBias (p. 999)
- DriftCheckExplainability (p. 1000)
- DriftCheckModelDataQuality (p. 1001)
- DriftCheckModelQuality (p. 1002)
- Edge (p. 1003)
- EdgeModel (p. 1004)
- EdgeModelStat (p. 1005)
- EdgeModelSummary (p. 1007)
- EdgeOutputConfig (p. 1008)
- EdgePackagingJobSummary (p. 1010)
- EdgePresetDeploymentOutput (p. 1012)
- EMRStepMetadata (p. 1013)
- Endpoint (p. 1014)
- EndpointConfigSummary (p. 1016)
- EndpointInput (p. 1017)
- EndpointInputConfiguration (p. 1019)
- EndpointOutputConfiguration (p. 1021)
- EndpointSummary (p. 1023)
- EnvironmentParameter (p. 1025)
- EnvironmentParameterRanges (p. 1026)
- Experiment (p. 1027)
- ExperimentConfig (p. 1029)
- ExperimentSource (p. 1031)
- ExperimentSummary (p. 1032)
- Explainability (p. 1034)
- FailStepMetadata (p. 1035)
- FeatureDefinition (p. 1036)
- FeatureGroup (p. 1037)
- FeatureGroupSummary (p. 1040)
- FileSource (p. 1042)
- FileSystemConfig (p. 1043)
- FileSystemDataSource (p. 1044)
- Filter (p. 1046)
- FinalAutoMLJobObjectiveMetric (p. 1049)
- FinalHyperParameterTuningJobObjectiveMetric (p. 1050)
- FlowDefinitionOutputConfig (p. 1051)
- FlowDefinitionSummary (p. 1052)
- GitConfig (p. 1054)
• GitConfigForUpdate (p. 1055)
• HumanLoopActivationConditionsConfig (p. 1056)
• HumanLoopActivationConfig (p. 1057)
• HumanLoopConfig (p. 1058)
• HumanLoopRequestSource (p. 1063)
• HumanTaskConfig (p. 1064)
• HumanTaskUiSummary (p. 1077)
• HyperParameterAlgorithmSpecification (p. 1078)
• HyperParameterSpecification (p. 1080)
• HyperParameterTrainingJobDefinition (p. 1082)
• HyperParameterTrainingJobSummary (p. 1086)
• HyperParameterTuningJobConfig (p. 1089)
• HyperParameterTuningJobObjective (p. 1091)
• HyperParameterTuningJobSummary (p. 1092)
• HyperParameterTuningJobWarmStartConfig (p. 1094)
• Image (p. 1096)
• ImageConfig (p. 1098)
• ImageVersion (p. 1099)
• InferenceExecutionConfig (p. 1101)
• InferenceRecommendation (p. 1102)
• InferenceRecommendationsJob (p. 1103)
• InferenceSpecification (p. 1105)
• InputConfig (p. 1107)
• InstanceMetadataServiceConfiguration (p. 1111)
• IntegerParameterRange (p. 1112)
• IntegerParameterRangeSpecification (p. 1114)
• JupyterServerAppSettings (p. 1115)
• KernelGatewayAppSettings (p. 1116)
• KernelGatewayImageConfig (p. 1117)
• KernelSpec (p. 1118)
• LabelCounters (p. 1119)
• LabelCountersForWorkteam (p. 1121)
• LabelingJobAlgorithmsConfig (p. 1122)
• LabelingJobDataAttributes (p. 1124)
• LabelingJobDataSource (p. 1125)
• LabelingJobForWorkteamSummary (p. 1126)
• LabelingJobInputConfig (p. 1128)
• LabelingJobOutput (p. 1129)
• LabelingJobOutputConfig (p. 1130)
• LabelingJobResourceConfig (p. 1132)
• LabelingJobS3DataSource (p. 1133)
• LabelingJobSnsDataSource (p. 1134)
• LabelingJobStoppingConditions (p. 1135)
• LabelingJobSummary (p. 1136)
• LambdaStepMetadata (p. 1139)
• LineageGroupSummary (p. 1140)
• MemberDefinition (p. 1142)
• MetadataProperties (p. 1143)
• MetricData (p. 1145)
• MetricDatum (p. 1146)
• MetricDefinition (p. 1147)
• MetricsSource (p. 1148)
• ModelArtifacts (p. 1149)
• ModelBiasAppSpecification (p. 1150)
• ModelBiasBaselineConfig (p. 1151)
• ModelBiasJobInput (p. 1152)
• ModelClientConfig (p. 1153)
• ModelConfiguration (p. 1154)
• ModelDataQuality (p. 1155)
• ModelDeployConfig (p. 1156)
• ModelDeployResult (p. 1157)
• ModelDigests (p. 1158)
• ModelExplainabilityAppSpecification (p. 1159)
• ModelExplainabilityBaselineConfig (p. 1160)
• ModelExplainabilityJobInput (p. 1161)
• ModelInput (p. 1162)
• ModelLatencyThreshold (p. 1163)
• ModelMetadataFilter (p. 1164)
• ModelMetadataSearchExpression (p. 1165)
• ModelMetadataSummary (p. 1166)
• ModelMetrics (p. 1167)
• ModelPackage (p. 1168)
• ModelPackageContainerDefinition (p. 1173)
• ModelPackageGroup (p. 1176)
• ModelPackageGroupSummary (p. 1178)
• ModelPackageStatusDetails (p. 1180)
• ModelPackageStatusItem (p. 1181)
• ModelPackageSummary (p. 1182)
• ModelPackageValidationProfile (p. 1184)
• ModelPackageValidationSpecification (p. 1185)
• ModelQuality (p. 1186)
• ModelQualityAppSpecification (p. 1187)
• ModelQualityBaselineConfig (p. 1189)
• ModelQualityJobInput (p. 1190)
• ModelStepMetadata (p. 1191)
• ModelSummary (p. 1192)
• MonitoringAppSpecification (p. 1193)
• MonitoringBaselineConfig (p. 1195)
• MonitoringClusterConfig (p. 1196)
• MonitoringConstraintsResource (p. 1198)
• MonitoringExecutionSummary (p. 1199)
• MonitoringGroundTruthS3Input (p. 1201)
• MonitoringInput (p. 1202)
• MonitoringJobDefinition (p. 1203)
• MonitoringJobDefinitionSummary (p. 1205)
• MonitoringNetworkConfig (p. 1207)
• MonitoringOutput (p. 1208)
• MonitoringOutputConfig (p. 1209)
• MonitoringResources (p. 1210)
• MonitoringS3Output (p. 1211)
• MonitoringSchedule (p. 1212)
• MonitoringScheduleConfig (p. 1215)
• MonitoringScheduleSummary (p. 1216)
• MonitoringStatisticsResource (p. 1218)
• MonitoringStoppingCondition (p. 1219)
• MultiModelConfig (p. 1220)
• NeoVpcConfig (p. 1221)
• NestedFilters (p. 1222)
• NetworkConfig (p. 1223)
• NotebookInstanceLifecycleConfigSummary (p. 1224)
• NotebookInstanceLifecycleHook (p. 1225)
• NotebookInstanceSummary (p. 1226)
• NotificationConfiguration (p. 1229)
• ObjectiveStatusCounters (p. 1230)
• OfflineStoreConfig (p. 1231)
• OfflineStoreStatus (p. 1232)
• OidcConfig (p. 1233)
• OidcConfigForResponse (p. 1235)
• OidcMemberDefinition (p. 1237)
• OnlineStoreConfig (p. 1238)
• OnlineStoreSecurityConfig (p. 1239)
• OutputConfig (p. 1240)
• OutputDataConfig (p. 1244)
• OutputParameter (p. 1246)
• ParallelismConfiguration (p. 1247)
• Parameter (p. 1248)
• ParameterRange (p. 1249)
• ParameterRanges (p. 1250)
• Parent (p. 1251)
• ParentHyperParameterTuningJob (p. 1252)
• PendingDeploymentSummary (p. 1253)
• PendingProductionVariantSummary (p. 1254)
• Phase (p. 1257)
• Pipeline (p. 1258)
• PipelineDefinitionS3Location (p. 1261)
• PipelineExecution (p. 1262)
• PipelineExecutionStep (p. 1265)
• PipelineExecutionStepMetadata (p. 1267)
• PipelineExecutionSummary (p. 1270)
• PipelineExperimentConfig (p. 1272)
• PipelineSummary (p. 1273)
• ProcessingClusterConfig (p. 1275)
• ProcessingFeatureStoreOutput (p. 1277)
• ProcessingInput (p. 1278)
• ProcessingJob (p. 1279)
• ProcessingJobStepMetadata (p. 1283)
• ProcessingJobSummary (p. 1284)
• ProcessingOutput (p. 1286)
• ProcessingOutputConfig (p. 1287)
• ProcessingResources (p. 1288)
• ProcessingS3Input (p. 1289)
• ProcessingS3Output (p. 1291)
• ProcessingStoppingCondition (p. 1292)
• ProductionVariant (p. 1293)
• ProductionVariantCoreDumpConfig (p. 1295)
• ProductionVariantServerlessConfig (p. 1297)
• ProductionVariantStatus (p. 1298)
• ProductionVariantSummary (p. 1299)
• ProfilerConfig (p. 1301)
• ProfilerConfigForUpdate (p. 1303)
• ProfilerRuleConfiguration (p. 1305)
• ProfilerRuleEvaluationStatus (p. 1307)
• Project (p. 1309)
• ProjectSummary (p. 1312)
• PropertyNameQuery (p. 1314)
• PropertyNameSuggestion (p. 1315)
• ProvisioningParameter (p. 1316)
• PublicWorkforceTaskPrice (p. 1317)
• QualityCheckStepMetadata (p. 1320)
• QueryFilters (p. 1322)
• RecommendationJobCompiledOutputConfig (p. 1324)
• RecommendationJobInputConfig (p. 1325)
• RecommendationJobOutputConfig (p. 1327)
• RecommendationJobResourceLimit (p. 1328)
• RecommendationJobStoppingConditions (p. 1329)
• RecommendationMetrics (p. 1330)
• RedshiftDatasetDefinition (p. 1331)
• RegisterModelStepMetadata (p. 1333)
• RenderableTask (p. 1334)
• RenderingError (p. 1335)
• RepositoryAuthConfig (p. 1336)
• ResolvedAttributes (p. 1337)
• ResourceConfig (p. 1338)
• ResourceLimits (p. 1340)
• ResourceSpec (p. 1341)
• RetentionPolicy (p. 1343)
• RetryStrategy (p. 1344)
• RSessionAppSettings (p. 1345)
• RStudioServerProAppSettings (p. 1346)
• RStudioServerProDomainSettings (p. 1347)
• RStudioServerProDomainSettingsForUpdate (p. 1348)
• S3DataSource (p. 1349)
• S3StorageConfig (p. 1352)
• ScheduleConfig (p. 1354)
• SearchExpression (p. 1355)
• SearchRecord (p. 1357)
• SecondaryStatusTransition (p. 1359)
• ServiceCatalogProvisionedProductDetails (p. 1361)
• ServiceCatalogProvisioningDetails (p. 1362)
• ServiceCatalogProvisioningUpdateDetails (p. 1364)
• SharingSettings (p. 1365)
• ShuffleConfig (p. 1366)
• SourceAlgorithm (p. 1367)
• SourceAlgorithmSpecification (p. 1368)
• SourceIpConfig (p. 1369)
• StoppingCondition (p. 1370)
• StudioLifecycleConfigDetails (p. 1372)
• SubscribedWorkteam (p. 1374)
• SuggestionQuery (p. 1376)
• Tag (p. 1377)
• TargetPlatform (p. 1378)
• TensorBoardAppSettings (p. 1380)
• TensorBoardOutputConfig (p. 1381)
• TrafficPattern (p. 1382)
• TrafficRoutingConfig (p. 1383)
• TrainingJob (p. 1385)
• TrainingJobDefinition (p. 1393)
• TrainingJobStatusCounters (p. 1395)
• TrainingJobStepMetadata (p. 1397)
• TrainingJobSummary (p. 1398)
• TrainingSpecification (p. 1400)
• TransformDataSource (p. 1402)
• TransformInput (p. 1403)
• TransformJob (p. 1405)
• TransformJobDefinition (p. 1410)
• TransformJobStepMetadata (p. 1412)
• TransformJobSummary (p. 1413)
• TransformOutput (p. 1415)
• TransformResources (p. 1417)
• TransformS3DataSource (p. 1419)
The following data types are supported by Amazon SageMaker Runtime:

The following data types are supported by Amazon Sagemaker Edge Manager:

The following data types are supported by Amazon SageMaker Feature Store Runtime:

The following data types are supported by Amazon SageMaker Service:

Amazon SageMaker Amazon Sagemaker API Reference
Amazon SageMaker Service

- Trial (p. 1421)
- TrialComponent (p. 1424)
- TrialComponentArtifact (p. 1428)
- TrialComponentMetricSummary (p. 1429)
- TrialComponentParameterValue (p. 1431)
- TrialComponentSimpleSummary (p. 1432)
- TrialComponentSource (p. 1434)
- TrialComponentSourceDetail (p. 1435)
- TrialComponentStatus (p. 1436)
- TrialComponentSummary (p. 1437)
- TrialSource (p. 1439)
- TrialSummary (p. 1440)
- TuningJobCompletionCriteria (p. 1442)
- TuningJobStepMetaData (p. 1443)
- UiConfig (p. 1444)
- UiTemplate (p. 1446)
- UiTemplateInfo (p. 1447)
- USD (p. 1448)
- UserContext (p. 1449)
- UserProfileDetails (p. 1450)
- UserSettings (p. 1452)
- VariantProperty (p. 1454)
- Vertex (p. 1455)
- VpcConfig (p. 1456)
- Workforce (p. 1457)
- Workteam (p. 1459)

The following data types are supported by Amazon SageMaker Runtime:

The following data types are supported by Amazon Sagemaker Edge Manager:

- EdgeMetric (p. 1462)
- Model (p. 1463)

The following data types are supported by Amazon SageMaker Feature Store Runtime:

- BatchGetRecordError (p. 1465)
- BatchGetRecordIdentifier (p. 1467)
- BatchGetRecordResultDetail (p. 1468)
- FeatureValue (p. 1469)

Amazon SageMaker Service

The following data types are supported by Amazon SageMaker Service:

- ActionSource (p. 860)
- ActionSummary (p. 861)
• AdditionalInferenceSpecificationDefinition (p. 863)
• AgentVersion (p. 865)
• Alarm (p. 866)
• AlgorithmSpecification (p. 867)
• AlgorithmStatusDetails (p. 870)
• AlgorithmStatusItem (p. 871)
• AlgorithmSummary (p. 872)
• AlgorithmValidationProfile (p. 874)
• AlgorithmValidationSpecification (p. 875)
• AnnotationConsolidationConfig (p. 876)
• AppDetails (p. 887)
• AppImageConfigDetails (p. 889)
• AppSpecification (p. 891)
• ArtifactSource (p. 892)
• ArtifactSourceType (p. 893)
• ArtifactSummary (p. 894)
• AssociationSummary (p. 896)
• AsyncInferenceClientConfig (p. 898)
• AsyncInferenceConfig (p. 899)
• AsyncInferenceNotificationConfig (p. 900)
• AsyncInferenceOutputConfig (p. 901)
• AthenaDatasetDefinition (p. 902)
• AutoMLCandidate (p. 904)
• AutoMLCandidateGenerationConfig (p. 906)
• AutoMLCandidateStep (p. 907)
• AutoMLChannel (p. 908)
• AutoMLContainerDefinition (p. 910)
• AutoMLDataSource (p. 912)
• AutoMLDataSplitConfig (p. 913)
• AutoMLJobArtifacts (p. 914)
• AutoMLJobCompletionCriteria (p. 915)
• AutoMLJobConfig (p. 916)
• AutoMLJobObjective (p. 917)
• AutoMLJobSummary (p. 919)
• AutoMLOutputDataConfig (p. 921)
• AutoMLPartialFailureReason (p. 922)
• AutoMLS3DataSource (p. 923)
• AutoMLSecurityConfig (p. 924)
• AutoRollbackConfig (p. 925)
• BatchDescribeModelPackageError (p. 926)
• BatchDescribeModelPackageSummary (p. 927)
• Bias (p. 929)
• BlueGreenUpdatePolicy (p. 930)
• CacheHitResult (p. 931)
• CallbackStepMetadata (p. 932)
• CandidateArtifactLocations (p. 933)
- CandidateProperties (p. 934)
- CapacitySize (p. 935)
- CaptureContentTypeHeader (p. 936)
- CaptureOption (p. 937)
- CategoricalParameter (p. 938)
- CategoricalParameterRange (p. 939)
- CategoricalParameterRangeSpecification (p. 940)
- Channel (p. 941)
- ChannelSpecification (p. 943)
- CheckpointConfig (p. 945)
- ClarifyCheckStepMetadata (p. 946)
- CodeRepositorySummary (p. 948)
- CognitoConfig (p. 950)
- CognitoMemberDefinition (p. 951)
- CollectionConfiguration (p. 952)
- CompilationJobSummary (p. 953)
- ConditionStepMetadata (p. 955)
- ContainerDefinition (p. 956)
- ContextSource (p. 959)
- ContextSummary (p. 960)
- ContinuousParameterRange (p. 962)
- ContinuousParameterRangeSpecification (p. 964)
- CustomImage (p. 965)
- DataCaptureConfig (p. 966)
- DataCaptureConfigSummary (p. 968)
- DataCatalogConfig (p. 970)
- DataProcessing (p. 971)
- DataQualityAppSpecification (p. 973)
- DataQualityBaselineConfig (p. 975)
- DataQualityJobInput (p. 976)
- DatasetDefinition (p. 977)
- DataSource (p. 979)
- DebugHookConfig (p. 980)
- DebugRuleConfiguration (p. 982)
- DebugRuleEvaluationStatus (p. 984)
- DeployedImage (p. 986)
- DeploymentConfig (p. 987)
- DesiredWeightAndCapacity (p. 988)
- Device (p. 989)
- DeviceFleetSummary (p. 990)
- DeviceStats (p. 991)
- DeviceSummary (p. 992)
- DomainDetails (p. 994)
- DomainSettings (p. 996)
- DomainSettingsForUpdate (p. 997)
- DriftCheckBaselines (p. 998)
• DriftCheckBias (p. 999)
• DriftCheckExplainability (p. 1000)
• DriftCheckModelDataQuality (p. 1001)
• DriftCheckModelQuality (p. 1002)
• Edge (p. 1003)
• EdgeModel (p. 1004)
• EdgeModelStat (p. 1005)
• EdgeModelSummary (p. 1007)
• EdgeOutputConfig (p. 1008)
• EdgePackagingJobSummary (p. 1010)
• EdgePresetDeploymentOutput (p. 1012)
• EMRStepMetadata (p. 1013)
• Endpoint (p. 1014)
• EndpointConfigSummary (p. 1016)
• EndpointInput (p. 1017)
• EndpointInputConfiguration (p. 1019)
• EndpointOutputConfiguration (p. 1021)
• EndpointSummary (p. 1023)
• EnvironmentParameter (p. 1025)
• EnvironmentParameterRanges (p. 1026)
• Experiment (p. 1027)
• ExperimentConfig (p. 1029)
• ExperimentSource (p. 1031)
• ExperimentSummary (p. 1032)
• Explainability (p. 1034)
• FailStepMetadata (p. 1035)
• FeatureDefinition (p. 1036)
• FeatureGroup (p. 1037)
• FeatureGroupSummary (p. 1040)
• FileSource (p. 1042)
• FileSystemConfig (p. 1043)
• FileSystemDataSource (p. 1044)
• Filter (p. 1046)
• FinalAutoMLJobObjectiveMetric (p. 1049)
• FinalHyperParameterTuningJobObjectiveMetric (p. 1050)
• FlowDefinitionOutputConfig (p. 1051)
• FlowDefinitionSummary (p. 1052)
• GitConfig (p. 1054)
• GitConfigForUpdate (p. 1055)
• HumanLoopActivationConditionsConfig (p. 1056)
• HumanLoopActivationConfig (p. 1057)
• HumanLoopConfig (p. 1058)
• HumanLoopRequestSource (p. 1063)
• HumanTaskConfig (p. 1064)
• HumanTaskUiSummary (p. 1077)
• HyperParameterAlgorithmSpecification (p. 1078)
• HyperParameterSpecification (p. 1080)
• HyperParameterTrainingJobDefinition (p. 1082)
• HyperParameterTrainingJobSummary (p. 1086)
• HyperParameterTuningJobConfig (p. 1089)
• HyperParameterTuningJobObjective (p. 1091)
• HyperParameterTuningJobSummary (p. 1092)
• HyperParameterTuningJobWarmStartConfig (p. 1094)
• Image (p. 1096)
• ImageConfig (p. 1098)
• ImageVersion (p. 1099)
• InferenceExecutionConfig (p. 1101)
• InferenceRecommendation (p. 1102)
• InferenceRecommendationsJob (p. 1103)
• InferenceSpecification (p. 1105)
• InputConfig (p. 1107)
• InstanceMetadataServiceConfiguration (p. 1111)
• IntegerParameterRange (p. 1112)
• IntegerParameterRangeSpecification (p. 1114)
• JupyterServerAppSettings (p. 1115)
• KernelGatewayAppSettings (p. 1116)
• KernelGatewayImageConfig (p. 1117)
• KernelSpec (p. 1118)
• LabelCounters (p. 1119)
• LabelCountersForWorkteam (p. 1121)
• LabelingJobAlgorithmsConfig (p. 1122)
• LabelingJobDataAttributes (p. 1124)
• LabelingJobDataSource (p. 1125)
• LabelingJobForWorkteamSummary (p. 1126)
• LabelingJobInputConfig (p. 1128)
• LabelingJobOutput (p. 1129)
• LabelingJobOutputConfig (p. 1130)
• LabelingJobResourceConfig (p. 1132)
• LabelingJobS3DataSource (p. 1133)
• LabelingJobSnsDataSource (p. 1134)
• LabelingJobStoppingConditions (p. 1135)
• LabelingJobSummary (p. 1136)
• LambdaStepMetadata (p. 1139)
• LineageGroupSummary (p. 1140)
• MemberDefinition (p. 1142)
• MetadataProperties (p. 1143)
• MetricData (p. 1145)
• MetricDatum (p. 1146)
• MetricDefinition (p. 1147)
• MetricsSource (p. 1148)
• ModelArtifacts (p. 1149)
• ModelBiasAppSpecification (p. 1150)
• ModelBiasBaselineConfig (p. 1151)
• ModelBiasJobInput (p. 1152)
• ModelClientConfig (p. 1153)
• ModelConfiguration (p. 1154)
• ModelDataQuality (p. 1155)
• ModelDeployConfig (p. 1156)
• ModelDeployResult (p. 1157)
• ModelDigests (p. 1158)
• ModelExplainabilityAppSpecification (p. 1159)
• ModelExplainabilityBaselineConfig (p. 1160)
• ModelExplainabilityJobInput (p. 1161)
• ModelInput (p. 1162)
• ModelLatencyThreshold (p. 1163)
• ModelMetadataFilter (p. 1164)
• ModelMetadataSearchExpression (p. 1165)
• ModelMetadataSummary (p. 1166)
• ModelMetrics (p. 1167)
• ModelPackage (p. 1168)
• ModelPackageContainerDefinition (p. 1173)
• ModelPackageGroup (p. 1176)
• ModelPackageGroupSummary (p. 1178)
• ModelPackageStatusDetails (p. 1180)
• ModelPackageStatusItem (p. 1181)
• ModelPackageSummary (p. 1182)
• ModelPackageValidationProfile (p. 1184)
• ModelPackageValidationSpecification (p. 1185)
• ModelQuality (p. 1186)
• ModelQualityAppSpecification (p. 1187)
• ModelQualityBaselineConfig (p. 1189)
• ModelQualityJobInput (p. 1190)
• ModelStepMetadata (p. 1191)
• ModelSummary (p. 1192)
• MonitoringAppSpecification (p. 1193)
• MonitoringBaselineConfig (p. 1195)
• MonitoringClusterConfig (p. 1196)
• MonitoringConstraintsResource (p. 1198)
• MonitoringExecutionSummary (p. 1199)
• MonitoringGroundTruthS3Input (p. 1201)
• MonitoringInput (p. 1202)
• MonitoringJobDefinition (p. 1203)
• MonitoringJobDefinitionSummary (p. 1205)
• MonitoringNetworkConfig (p. 1207)
• MonitoringOutput (p. 1208)
• MonitoringOutputConfig (p. 1209)
• MonitoringResources (p. 1210)
• MonitoringS3Output (p. 1211)
- MonitoringSchedule (p. 1212)
- MonitoringScheduleConfig (p. 1215)
- MonitoringScheduleSummary (p. 1216)
- MonitoringStatisticsResource (p. 1218)
- MonitoringStoppingCondition (p. 1219)
- MultiModelConfig (p. 1220)
- NeoVpcConfig (p. 1221)
- NestedFilters (p. 1222)
- NetworkConfig (p. 1223)
- NotebookInstanceLifecycleConfigSummary (p. 1224)
- NotebookInstanceLifecycleHook (p. 1225)
- NotebookInstanceSummary (p. 1226)
- NotificationConfiguration (p. 1229)
- ObjectiveStatusCounters (p. 1230)
- OfflineStoreConfig (p. 1231)
- OfflineStoreStatus (p. 1232)
- OidcConfig (p. 1233)
- OidcConfigForResponse (p. 1235)
- OidcMemberDefinition (p. 1237)
- OnlineStoreConfig (p. 1238)
- OnlineStoreSecurityConfig (p. 1239)
- OutputConfig (p. 1240)
- OutputDataConfig (p. 1244)
- OutputParameter (p. 1246)
- ParallelismConfiguration (p. 1247)
- Parameter (p. 1248)
- ParameterRange (p. 1249)
- ParameterRanges (p. 1250)
- Parent (p. 1251)
- ParentHyperParameterTuningJob (p. 1252)
- PendingDeploymentSummary (p. 1253)
- PendingProductionVariantSummary (p. 1254)
- Phase (p. 1257)
- Pipeline (p. 1258)
- PipelineDefinitionS3Location (p. 1261)
- PipelineExecution (p. 1262)
- PipelineExecutionStep (p. 1265)
- PipelineExecutionStepMetadata (p. 1267)
- PipelineExecutionSummary (p. 1270)
- PipelineExperimentConfig (p. 1272)
- PipelineSummary (p. 1273)
- ProcessingClusterConfig (p. 1275)
- ProcessingFeatureStoreOutput (p. 1277)
- ProcessingInput (p. 1278)
- ProcessingJob (p. 1279)
- ProcessingJobStepMetadata (p. 1283)
• ProcessingJobSummary (p. 1284)
• ProcessingOutput (p. 1286)
• ProcessingOutputConfig (p. 1287)
• ProcessingResources (p. 1288)
• ProcessingS3Input (p. 1289)
• ProcessingS3Output (p. 1291)
• ProcessingStoppingCondition (p. 1292)
• ProductionVariant (p. 1293)
• ProductionVariantCoreDumpConfig (p. 1295)
• ProductionVariantServerlessConfig (p. 1297)
• ProductionVariantStatus (p. 1298)
• ProductionVariantSummary (p. 1299)
• ProfilerConfig (p. 1301)
• ProfilerConfigForUpdate (p. 1303)
• ProfilerRuleConfiguration (p. 1305)
• ProfilerRuleEvaluationStatus (p. 1307)
• Project (p. 1309)
• ProjectSummary (p. 1312)
• PropertyNameQuery (p. 1314)
• PropertyNameSuggestion (p. 1315)
• ProvisioningParameter (p. 1316)
• PublicWorkforceTaskPrice (p. 1317)
• QualityCheckStepMetadata (p. 1320)
• QueryFilters (p. 1322)
• RecommendationJobCompiledOutputConfig (p. 1324)
• RecommendationJobInputConfig (p. 1325)
• RecommendationJobOutputConfig (p. 1327)
• RecommendationJobResourceLimit (p. 1328)
• RecommendationJobStoppingConditions (p. 1329)
• RecommendationMetrics (p. 1330)
• RedshiftDatasetDefinition (p. 1331)
• RegisterModelStepMetadata (p. 1333)
• RenderableTask (p. 1334)
• RenderingError (p. 1335)
• RepositoryAuthConfig (p. 1336)
• ResolvedAttributes (p. 1337)
• ResourceConfig (p. 1338)
• ResourceLimits (p. 1340)
• ResourceSpec (p. 1341)
• RetentionPolicy (p. 1343)
• RetryStrategy (p. 1344)
• RSessionAppSettings (p. 1345)
• RStudioServerProAppSettings (p. 1346)
• RStudioServerProDomainSettings (p. 1347)
• RStudioServerProDomainSettingsForUpdate (p. 1348)
• S3DataSource (p. 1349)
• S3StorageConfig (p. 1352)
• ScheduleConfig (p. 1354)
• SearchExpression (p. 1355)
• SearchRecord (p. 1357)
• SecondaryStatusTransition (p. 1359)
• ServiceCatalogProvisionedProductDetails (p. 1361)
• ServiceCatalogProvisioningDetails (p. 1362)
• ServiceCatalogProvisioningUpdateDetails (p. 1364)
• SharingSettings (p. 1365)
• ShuffleConfig (p. 1366)
• SourceAlgorithm (p. 1367)
• SourceAlgorithmSpecification (p. 1368)
• SourceIpConfig (p. 1369)
• StoppingCondition (p. 1370)
• StudioLifecycleConfigDetails (p. 1372)
• SubscribedWorkteam (p. 1374)
• SuggestionQuery (p. 1376)
• Tag (p. 1377)
• TargetPlatform (p. 1378)
• TensorBoardAppSettings (p. 1380)
• TensorBoardOutputConfig (p. 1381)
• TrafficPattern (p. 1382)
• TrafficRoutingConfig (p. 1383)
• TrainingJob (p. 1385)
• TrainingJobDefinition (p. 1393)
• TrainingJobStatusCounters (p. 1395)
• TrainingJobStepMetadata (p. 1397)
• TrainingJobSummary (p. 1398)
• TrainingSpecification (p. 1400)
• TransformDataSource (p. 1402)
• TransformInput (p. 1403)
• TransformJob (p. 1405)
• TransformJobDefinition (p. 1410)
• TransformJobStepMetadata (p. 1412)
• TransformJobSummary (p. 1413)
• TransformOutput (p. 1415)
• TransformResources (p. 1417)
• TransformS3DataSource (p. 1419)
• Trial (p. 1421)
• TrialComponent (p. 1424)
• TrialComponentArtifact (p. 1428)
• TrialComponentMetricSummary (p. 1429)
• TrialComponentParameterValue (p. 1431)
• TrialComponentSimpleSummary (p. 1432)
• TrialComponentSource (p. 1434)
• TrialComponentSourceDetail (p. 1435)
- TrialComponentStatus (p. 1436)
- TrialComponentSummary (p. 1437)
- TrialSource (p. 1439)
- TrialSummary (p. 1440)
- TuningJobCompletionCriteria (p. 1442)
- TuningJobStepMetaData (p. 1443)
- UiConfig (p. 1444)
- UiTemplate (p. 1446)
- UiTemplateInfo (p. 1447)
- USD (p. 1448)
- UserContext (p. 1449)
- UserProfileDetails (p. 1450)
- UserSettings (p. 1452)
- VariantProperty (p. 1454)
- Vertex (p. 1455)
- VpcConfig (p. 1456)
- Workforce (p. 1457)
- Workteam (p. 1459)
ActionSource
Service: Amazon SageMaker Service

A structure describing the source of an action.

Contents

SourceId

The ID of the source.
Type: String
Length Constraints: Maximum length of 256.
Required: No

SourceType

The type of the source.
Type: String
Length Constraints: Maximum length of 256.
Required: No

SourceUri

The URI of the source.
Type: String
Length Constraints: Maximum length of 2048.
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ActionSummary
Service: Amazon SageMaker Service

Lists the properties of an action. An action represents an action or activity. Some examples are a workflow step and a model deployment. Generally, an action involves at least one input artifact or output artifact.

Contents

ActionArn
The Amazon Resource Name (ARN) of the action.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:action/.*
Required: No

ActionName
The name of the action.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9\-]{0,119}$
Required: No

ActionType
The type of the action.
Type: String
Length Constraints: Maximum length of 64.
Required: No

CreationTime
When the action was created.
Type: Timestamp
Required: No

LastModifiedTime
When the action was last modified.
Type: Timestamp
Required: No

Source
The source of the action.
Type: ActionSource (p. 860) object
**Required:** No

**Status**

The status of the action.

Type: String

Valid Values: Unknown | InProgress | Completed | Failed | Stopping | Stopped

**Required:** No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AdditionalInferenceSpecificationDefinition

Service: Amazon SageMaker Service

A structure of additional Inference Specification. Additional Inference Specification specifies details about inference jobs that can be run with models based on this model package.

Contents

Containers

The Amazon ECR registry path of the Docker image that contains the inference code.

Type: Array of ModelPackageContainerDefinition (p. 1173) objects

Array Members: Minimum number of 1 item. Maximum number of 15 items.

Required: Yes

Description

A description of the additional Inference specification

Type: String

Length Constraints: Maximum length of 1024.

Pattern: \[p\{L\}p\{M\}p\{Z\}p\{S\}p\{N\}p\{P\}]*

Required: No

Name

A unique name to identify the additional inference specification. The name must be unique within the list of your additional inference specifications for a particular model package.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

SupportedContentTypes

The supported MIME types for the input data.

Type: Array of strings

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

SupportedRealtimeInferenceInstanceTypes

A list of the instance types that are used to generate inferences in real-time.

Type: Array of strings

Valid Values: ml.t2.medium | ml.t2.large | ml.t2.xlarge | ml.t2.2xlarge | ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge |
AdditionalInferenceSpecificationDefinition

The supported MIME types for the output data.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Pattern: ^[-\w]+\/.+$

SupportedTransformInstanceTypes

A list of the instance types on which a transformation job can be run or on which an endpoint can be deployed.

Type: Array of strings

Valid Values:

| ml.m4.16xlarge | ml.m5.large | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge |
| ml.m5.12xlarge | ml.m5.24xlarge | ml.m5d.large | ml.m5d.xlarge |
| ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.12xlarge | ml.m5d.24xlarge |
| ml.c4.large | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge | ml.c4.8xlarge |
| ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge | ml.p3.2xlarge |
| ml.p3.8xlarge | ml.p3.16xlarge | ml.c5.large | ml.c5.xlarge |
| ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge | ml.c5d.large |
| ml.c5d.xlarge | ml.c5d.2xlarge | ml.c5d.4xlarge | ml.c5d.9xlarge |
| ml.c5d.18xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge |
| ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.r5.large |
| ml.r5.xlarge | ml.r5.2xlarge | ml.r5.4xlarge | ml.r5.12xlarge |
| ml.r5.24xlarge | ml.r5d.large | ml.r5d.xlarge | ml.r5d.2xlarge |
| ml.r5d.4xlarge | ml.r5d.12xlarge | ml.r5d.24xlarge | ml.inf1.xlarge |
| ml.inf1.2xlarge | ml.inf1.6xlarge | ml.inf1.24xlarge | ml.r5d.16xlarge |
AgentVersion
Service: Amazon SageMaker Service

Edge Manager agent version.

Contents

AgentCount

The number of Edge Manager agents.

Type: Long

Required: Yes

Version

Version of the agent.

Type: String


Pattern: \[a-zA-Z0-9\ \_\ \.]+\]

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Alarm
Service: Amazon SageMaker Service
An Amazon CloudWatch alarm configured to monitor metrics on an endpoint.

Contents

AlarmName
The name of a CloudWatch alarm in your account.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 255.
Pattern: ^(?![\s*$]).+
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AlgorithmSpecification

Service: Amazon SageMaker Service

Specifies the training algorithm to use in a CreateTrainingJob (p. 173) request.

For more information about algorithms provided by SageMaker, see Algorithms. For information about using your own algorithms, see Using Your Own Algorithms with Amazon SageMaker.

Contents

AlgorithmName

The name of the algorithm resource to use for the training job. This must be an algorithm resource that you created or subscribe to on AWS Marketplace.

Note
You must specify either the algorithm name to the AlgorithmName parameter or the image URI of the algorithm container to the TrainingImage parameter.

Note that the AlgorithmName parameter is mutually exclusive with the TrainingImage parameter. If you specify a value for the AlgorithmName parameter, you can't specify a value for TrainingImage, and vice versa.

If you specify values for both parameters, the training job might break; if you don't specify any value for both parameters, the training job might raise a null error.

Type: String


Pattern: (arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:a-z\-\/[0-9]{12}([a-zA-Z0-9]\[a-zA-Z0-9-]+)?(\-\d)\[0,62]\(\d\)?

Required: No

EnableSageMakerMetricsTimeSeries

To generate and save time-series metrics during training, set to true. The default is false and time-series metrics aren't generated except in the following cases:

• You use one of the SageMaker built-in algorithms
• You use one of the following Prebuilt SageMaker Docker Images:
  • Tensorflow (version >= 1.15)
  • MXNet (version >= 1.6)
  • PyTorch (version >= 1.3)
• You specify at least one MetricDefinition (p. 1147)

Type: Boolean

Required: No

MetricDefinitions

A list of metric definition objects. Each object specifies the metric name and regular expressions used to parse algorithm logs. SageMaker publishes each metric to Amazon CloudWatch.

Type: Array of MetricDefinition (p. 1147) objects

Array Members: Minimum number of 0 items. Maximum number of 40 items.

Required: No
### TrainingImage

The registry path of the Docker image that contains the training algorithm. For information about docker registry paths for SageMaker built-in algorithms, see [Docker Registry Paths and Example Code](https://docs.aws.amazon.com/sagemaker/latest/dg/how-to-debug-algorithms.html) in the Amazon SageMaker developer guide. SageMaker supports both `registry/repository[:tag]` and `registry/repository[@digest]` image path formats. For more information about using your custom training container, see [Using Your Own Algorithms with Amazon SageMaker](https://docs.aws.amazon.com/sagemaker/latest/dg/how-to-debug-algorithms.html).

**Note**

You must specify either the algorithm name to the `AlgorithmName` parameter or the image URI of the algorithm container to the `TrainingImage` parameter.

For more information, see the note in the `AlgorithmName` parameter description.

**Type:** String

**Length Constraints:** Maximum length of 255.

**Pattern:** .*

**Required:** No

### TrainingInputMode

The training input mode that the algorithm supports. For more information about input modes, see [Algorithms](https://docs.aws.amazon.com/sagemaker/latest/dg/how-to-debug-algorithms.html).

#### Pipe mode

If an algorithm supports Pipe mode, Amazon SageMaker streams data directly from Amazon S3 to the container.

#### File mode

If an algorithm supports File mode, SageMaker downloads the training data from S3 to the provisioned ML storage volume, and mounts the directory to the Docker volume for the training container.

You must provision the ML storage volume with sufficient capacity to accommodate the data downloaded from S3. In addition to the training data, the ML storage volume also stores the output model. The algorithm container uses the ML storage volume to also store intermediate information, if any.

For distributed algorithms, training data is distributed uniformly. Your training duration is predictable if the input data objects sizes are approximately the same. SageMaker does not split the files any further for model training. If the object sizes are skewed, training won't be optimal as the data distribution is also skewed when one host in a training cluster is overloaded, thus becoming a bottleneck in training.

#### FastFile mode

If an algorithm supports FastFile mode, SageMaker streams data directly from S3 to the container with no code changes, and provides file system access to the data. Users can author their training script to interact with these files as if they were stored on disk.

FastFile mode works best when the data is read sequentially. Augmented manifest files aren't supported. The startup time is lower when there are fewer files in the S3 bucket provided.

**Type:** String

**Valid Values:** Pipe | File | FastFile
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AlgorithmStatusDetails
Service: Amazon SageMaker Service

Specifies the validation and image scan statuses of the algorithm.

Contents

**ImageScanStatuses**

The status of the scan of the algorithm's Docker image container.

Type: Array of AlgorithmStatusItem (p. 871) objects

Required: No

**ValidationStatuses**

The status of algorithm validation.

Type: Array of AlgorithmStatusItem (p. 871) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AlgorithmStatusItem

Service: Amazon SageMaker Service

Represents the overall status of an algorithm.

Contents

FailureReason

if the overall status is Failed, the reason for the failure.

Type: String

Required: No

Name

The name of the algorithm for which the overall status is being reported.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$

Required: Yes

Status

The current status.

Type: String

Valid Values: NotStarted | InProgress | Completed | Failed

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AlgorithmSummary

Service: Amazon SageMaker Service

Provides summary information about an algorithm.

Contents

AlgorithmArn

The Amazon Resource Name (ARN) of the algorithm.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:algorithm/.*`

Required: Yes

AlgorithmDescription

A brief description of the algorithm.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `[^\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*`

Required: No

AlgorithmName

The name of the algorithm that is described by the summary.

Type: String


Pattern: `^[a-zA-20-9]{0,62}$`

Required: Yes

AlgorithmStatus

The overall status of the algorithm.

Type: String

Valid Values: Pending | InProgress | Completed | Failed | Deleting

Required: Yes

CreationTime

A timestamp that shows when the algorithm was created.

Type: Timestamp

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
AlgorithmValidationProfile
Service: Amazon SageMaker Service

Defines a training job and a batch transform job that SageMaker runs to validate your algorithm. The data provided in the validation profile is made available to your buyers on AWS Marketplace.

Contents

ProfileName

The name of the profile for the algorithm. The name must have 1 to 63 characters. Valid characters are a-z, A-Z, 0-9, and - (hyphen).

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

TrainingJobDefinition

The TrainingJobDefinition object that describes the training job that SageMaker runs to validate your algorithm.

Type: TrainingJobDefinition (p. 1393) object

Required: Yes

TransformJobDefinition

The TransformJobDefinition object that describes the transform job that SageMaker runs to validate your algorithm.

Type: TransformJobDefinition (p. 1410) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AlgorithmValidationSpecification

Service: Amazon SageMaker Service

Specifies configurations for one or more training jobs that SageMaker runs to test the algorithm.

Contents

ValidationProfiles

An array of AlgorithmValidationProfile objects, each of which specifies a training job and batch transform job that SageMaker runs to validate your algorithm.

Type: Array of AlgorithmValidationProfile (p. 874) objects

Array Members: Fixed number of 1 item.

Required: Yes

ValidationRole

The IAM roles that SageMaker uses to run the training jobs.

Type: String


Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_\//]+$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AnnotationConsolidationConfig

Service: Amazon SageMaker Service

Configures how labels are consolidated across human workers and processes output data.

Contents

AnnotationConsolidationLambdaArn

The Amazon Resource Name (ARN) of a Lambda function implements the logic for annotation consolidation and to process output data.

This parameter is required for all labeling jobs. For built-in task types, use one of the following Amazon SageMaker Ground Truth Lambda function ARNs for AnnotationConsolidationLambdaArn. For custom labeling workflows, see Post-annotation Lambda.

**Bounding box** - Finds the most similar boxes from different workers based on the Jaccard index of the boxes.
- arn:aws:lambda:us-east-1:432418664414:function:ACS-BoundingBox
- arn:aws:lambda:eu-west-1:568282634449:function:ACS-BoundingBox
- arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-BoundingBox
- arn:aws:lambda:ap-south-1:565803892007:function:ACS-BoundingBox
- arn:aws:lambda:eu-west-1:20300106190:function:ACS-BoundingBox
- arn:aws:lambda:ap-south-1:20300106190:function:ACS-BoundingBox
- arn:aws:lambda:eu-west-2:487402164563:function:ACS-BoundingBox
- arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-BoundingBox
- arn:aws:lambda:ca-central-1:918755190332:function:ACS-BoundingBox

**Image classification** - Uses a variant of the Expectation Maximization approach to estimate the true class of an image based on annotations from individual workers.
- arn:aws:lambda:us-east-1:432418664414:function:ACS-ImageMultiClass
- arn:aws:lambda:eu-west-1:568282634449:function:ACS-ImageMultiClass
- arn:aws:lambda:eu-west-1:20300106190:function:ACS-ImageMultiClass

**Multi-label image classification** - Uses a variant of the Expectation Maximization approach to estimate the true classes of an image based on annotations from individual workers.
• arn:aws:lambda:eu-west-1:568282634449:function:ACS-ImageMultiClassMultiLabel
• arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-ImageMultiClassMultiLabel
• arn:aws:lambda:ap-south-1:565803892007:function:ACS-ImageMultiClassMultiLabel
• arn:aws:lambda:eu-central-1:203001061592:function:ACS-ImageMultiClassMultiLabel
• arn:aws:lambda:eu-west-2:487402164563:function:ACS-ImageMultiClassMultiLabel
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-ImageMultiClassMultiLabel
• arn:aws:lambda:ca-central-1:918755190332:function:ACS-ImageMultiClassMultiLabel

**Semantic segmentation** - Treats each pixel in an image as a multi-class classification and treats pixel annotations from workers as "votes" for the correct label.
• arn:aws:lambda:us-east-1:432418664414:function:ACS-SemanticSegmentation
• arn:aws:lambda:eu-west-1:568282634449:function:ACS-SemanticSegmentation
• arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-SemanticSegmentation
• arn:aws:lambda:ap-south-1:565803892007:function:ACS-SemanticSegmentation
• arn:aws:lambda:eu-central-1:203001061592:function:ACS-SemanticSegmentation
• arn:aws:lambda:eu-west-2:487402164563:function:ACS-SemanticSegmentation
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-SemanticSegmentation

**Text classification** - Uses a variant of the Expectation Maximization approach to estimate the true class of text based on annotations from individual workers.
• arn:aws:lambda:us-east-1:432418664414:function:ACS-TextMultiClass
• arn:aws:lambda:us-east-2:266458841044:function:ACS-TextMultiClass
• arn:aws:lambda:us-west-2:081040173940:function:ACS-TextMultiClass
• arn:aws:lambda:eu-west-1:568282634449:function:ACS-TextMultiClass
• arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-TextMultiClass
• arn:aws:lambda:ap-southeast-2:454466003867:function:ACS-TextMultiClass
Multi-label text classification - Uses a variant of the Expectation Maximization approach to estimate the true classes of text based on annotations from individual workers.


Named entity recognition - Groups similar selections and calculates aggregate boundaries, resolving to most-assigned label.


Multi-label text classification - Uses a variant of the Expectation Maximization approach to estimate the true classes of text based on annotations from individual workers.

- arn:aws:lambda:eu-west-1:56828264449:function:ACS-TextMultiClassMultiLabel
- arn:aws:lambda:eu-west-1:56828264449:function:ACS-TextMultiClassMultiLabel

Named entity recognition - Groups similar selections and calculates aggregate boundaries, resolving to most-assigned label.

Video Classification - Use this task type when you need workers to classify videos using predefined labels that you specify. Workers are shown videos and are asked to choose one label for each video.

Video Frame Object Detection - Use this task type to have workers identify and locate objects in a sequence of video frames (images extracted from a video) using bounding boxes. For example, you can use this task to ask workers to identify and localize various objects in a series of video frames, such as cars, bikes, and pedestrians.

Video Frame Object Tracking - Use this task type to have workers track the movement of objects in a sequence of video frames (images extracted from a video) using bounding boxes. For example, you can use this task to ask workers to track the movement of objects, such as cars, bikes, and pedestrians.
3D Point Cloud Object Detection - Use this task type when you want workers to classify objects in a 3D point cloud by drawing 3D cuboids around objects. For example, you can use this task type to ask workers to identify different types of objects in a point cloud, such as cars, bikes, and pedestrians.

- arn:aws:lambda:eu-west-1:568282634449:function:ACS-3DPointCloudObjectDetection

3D Point Cloud Object Tracking - Use this task type when you want workers to draw 3D cuboids around objects that appear in a sequence of 3D point cloud frames. For example, you can use this task type to ask workers to track the movement of vehicles across multiple point cloud frames.

- arn:aws:lambda:us-east-1:432418664414:function:ACS-3DPointCloudObjectTracking
- arn:aws:lambda:eu-west-1:568282634449:function:ACS-3DPointCloudObjectTracking
- arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-3DPointCloudObjectTracking
- arn:aws:lambda:ap-south-1:565803892007:function:ACS-3DPointCloudObjectTracking
AnnotationConsolidationConfig

- 
  - arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-3DPointCloudObjectTracking
  - arn:aws:lambda:ap-south-1:565803892007:function:ACS-3DPointCloudObjectTracking

3D Point Cloud Semantic Segmentation - Use this task type when you want workers to create a point-level semantic segmentation masks by painting objects in a 3D point cloud using different colors where each color is assigned to one of the classes you specify.

- 
  - arn:aws:lambda:eu-west-1:568282634449:function:ACS-3DPointCloudSemanticSegmentation

Use the following ARNs for Label Verification and Adjustment Jobs

Use label verification and adjustment jobs to review and adjust labels. To learn more, see Verify and Adjust Labels.

Semantic Segmentation Adjustment - Treats each pixel in an image as a multi-class classification and treats pixel adjusted annotations from workers as "votes" for the correct label.
• `arn:aws:lambda:us-east-1:432418664414:function:ACS-AdjustmentSemanticSegmentation`
• `arn:aws:lambda:eu-west-1:568282634449:function:ACS-AdjustmentSemanticSegmentation`
• `arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-AdjustmentSemanticSegmentation`
• `arn:aws:lambda:ap-south-1:565803892007:function:ACS-AdjustmentSemanticSegmentation`
• `arn:aws:lambda:eu-central-1:203001061592:function:ACS-AdjustmentSemanticSegmentation`
• `arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-AdjustmentSemanticSegmentation`

**Semantic Segmentation Verification** - Uses a variant of the Expectation Maximization approach to estimate the true class of verification judgment for semantic segmentation labels based on annotations from individual workers.

• `arn:aws:lambda:us-east-1:432418664414:function:ACS-VerificationSemanticSegmentation`
• `arn:aws:lambda:eu-west-1:568282634449:function:ACS-VerificationSemanticSegmentation`
• `arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-VerificationSemanticSegmentation`
• `arn:aws:lambda:ap-south-1:565803892007:function:ACS-VerificationSemanticSegmentation`
• `arn:aws:lambda:eu-central-1:203001061592:function:ACS-VerificationSemanticSegmentation`
• `arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-VerificationSemanticSegmentation`

**Bounding Box Adjustment** - Finds the most similar boxes from different workers based on the Jaccard index of the adjusted annotations.

• `arn:aws:lambda:us-east-1:432418664414:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:us-east-2:266458841044:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:us-west-2:081040173940:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:eu-west-1:568282634449:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:ap-southeast-2:454466003867:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:ap-south-1:565803892007:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:eu-central-1:203001061592:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:ap-northeast-2:845288260483:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:eu-west-2:487402164563:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-AdjustmentBoundingBox`
• `arn:aws:lambda:ca-central-1:918755190332:function:ACS-AdjustmentBoundingBox`

**Bounding Box Verification** - Uses a variant of the Expectation Maximization approach to estimate the true class of verification judgment for bounding box labels based on annotations from individual workers.

• `arn:aws:lambda:us-east-1:432418664414:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:us-east-2:266458841044:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:us-west-2:081040173940:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:eu-west-1:568282634449:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:ap-southeast-2:454466003867:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:ap-south-1:565803892007:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:eu-central-1:203001061592:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:ap-northeast-2:845288260483:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:eu-west-2:487402164563:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-VerificationBoundingBox`
• `arn:aws:lambda:ca-central-1:918755190332:function:ACS-VerificationBoundingBox`

**Video Frame Object Detection Adjustment** - Use this task type when you want workers to adjust bounding boxes that workers have added to video frames to classify and localize objects in a sequence of video frames.
• arn:aws:lambda:us-east-1:432418664414:function:ACS-AdjustmentVideoObjectDetection
• arn:aws:lambda:eu-west-1:568282634449:function:ACS-AdjustmentVideoObjectDetection
• arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-AdjustmentVideoObjectDetection
• arn:aws:lambda:ap-south-1:565803892007:function:ACS-AdjustmentVideoObjectDetection
• arn:aws:lambda:eu-central-1:203001061592:function:ACS-AdjustmentVideoObjectDetection
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-AdjustmentVideoObjectDetection

**Video Frame Object Tracking Adjustment** - Use this task type when you want workers to adjust bounding boxes that workers have added to video frames to track object movement across a sequence of video frames.

• arn:aws:lambda:us-east-1:432418664414:function:ACS-AdjustmentVideoObjectTracking
• arn:aws:lambda:eu-west-1:568282634449:function:ACS-AdjustmentVideoObjectTracking
• arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-AdjustmentVideoObjectTracking
• arn:aws:lambda:ap-south-1:565803892007:function:ACS-AdjustmentVideoObjectTracking
• arn:aws:lambda:eu-central-1:203001061592:function:ACS-AdjustmentVideoObjectTracking
• arn:aws:lambda:eu-west-2:487402164563:function:ACS-AdjustmentVideoObjectTracking
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-AdjustmentVideoObjectTracking
• arn:aws:lambda:ca-central-1:918755190332:function:ACS-AdjustmentVideoObjectTracking

3D Point Cloud Object Detection Adjustment - Use this task type when you want workers to adjust 3D cuboids around objects in a 3D point cloud.
• arn:aws:lambda:us-east-1:432418664414:function:ACS-Adjustment3DPointCloudObjectDetection
• arn:aws:lambda:eu-west-1:568282634449:function:ACS-Adjustment3DPointCloudObjectDetection
• arn:aws:lambda:eu-central-1:203001061592:function:ACS-Adjustment3DPointCloudObjectDetection
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-Adjustment3DPointCloudObjectDetection
• arn:aws:lambda:ap-south-1:565803892007:function:ACS-Adjustment3DPointCloudObjectDetection
• arn:aws:lambda:eu-central-1:203001061592:function:ACS-Adjustment3DPointCloudObjectDetection
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-Adjustment3DPointCloudObjectDetection
• arn:aws:lambda:ca-central-1:918755190332:function:ACS-Adjustment3DPointCloudObjectDetection

3D Point Cloud Object Tracking Adjustment - Use this task type when you want workers to adjust 3D cuboids around objects that appear in a sequence of 3D point cloud frames.
• arn:aws:lambda:us-east-1:432418664414:function:ACS-Adjustment3DPointCloudObjectTracking
• arn:aws:lambda:us-east-2:266458841044:function:ACS-Adjustment3DPointCloudObjectTracking
• arn:aws:lambda:us-west-2:081040173940:function:ACS-Adjustment3DPointCloudObjectTracking
• arn:aws:lambda:eu-west-1:487402164563:function:ACS-Adjustment3DPointCloudObjectTracking
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-Adjustment3DPointCloudObjectTracking
• arn:aws:lambda:ca-central-1:918755190332:function:ACS-Adjustment3DPointCloudObjectTracking
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-Adjustment3DPointCloudObjectTracking
• arn:aws:lambda:ca-central-1:918755190332:function:ACS-Adjustment3DPointCloudObjectTracking

3D Point Cloud Semantic Segmentation Adjustment - Use this task type when you want workers to adjust a point-level semantic segmentation masks using a paint tool.
• arn:aws:lambda:us-east-1:432418664414:function:ACS-3DPointCloudSemanticSegmentation
• arn:aws:lambda:us-east-1:432418664414:function:ACS-Adjustment3DPointCloudSemanticSegmentation
• arn:aws:lambda:eu-west-1:568282634449:function:ACS-Adjustment3DPointCloudSemanticSegmentation
• arn:aws:lambda:ap-northeast-1:477331159723:function:ACS-Adjustment3DPointCloudSemanticSegmentation
• arn:aws:lambda:ap-south-1:565803892007:function:ACS-Adjustment3DPointCloudSemanticSegmentation
• arn:aws:lambda:eu-central-1:203001061592:function:ACS-Adjustment3DPointCloudSemanticSegmentation
• arn:aws:lambda:eu-west-2:487402164563:function:ACS-Adjustment3DPointCloudSemanticSegmentation
• arn:aws:lambda:ap-southeast-1:377565633583:function:ACS-Adjustment3DPointCloudSemanticSegmentation
• arn:aws:lambda:ca-central-1:918755190332:function:ACS-Adjustment3DPointCloudSemanticSegmentation

Type: String

Length Constraints: Maximum length of 2048.

Pattern: arn:[a-z-]*:lambda:[a-z0-9-]*:[0-9]{12}:function:.*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
AppDetails
Service: Amazon SageMaker Service
Details about an Amazon SageMaker app.

Contents

AppName
The name of the app.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]{0,62})
Required: No

AppType
The type of app.
Type: String
Valid Values: JupyterServer | KernelGateway | TensorBoard | RStudioServerPro | RSessionGateway
Required: No

CreationTime
The creation time.
Type: Timestamp
Required: No

DomainId
The domain ID.
Type: String
Length Constraints: Maximum length of 63.
Required: No

Status
The status.
Type: String
Valid Values: Deleted | Deleting | Failed | InService | Pending
Required: No

UserProfileName
The user profile name.
Type: String
Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9])*\{0,62\}

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AppImageConfigDetails

Service: Amazon SageMaker Service

The configuration for running a SageMaker image as a KernelGateway app.

Contents

AppImageConfigArn

The Amazon Resource Name (ARN) of the AppImageConfig.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:app-image-config/.*

Required: No

AppImageConfigName

The name of the AppImageConfig. Must be unique to your account.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-](-*[a-zA-Z0-9\-])\{0,62

Required: No

CreationTime

When the AppImageConfig was created.

Type: Timestamp

Required: No

KernelGatewayImageConfig

The configuration for the file system and kernels in the SageMaker image.

Type: KernelGatewayImageConfig (p. 1117) object

Required: No

LastModifiedTime

When the AppImageConfig was last modified.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AppSpecification
Service: Amazon SageMaker Service
Configuration to run a processing job in a specified container image.

Contents

ContainerArguments
The arguments for a container used to run a processing job.
Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 100 items.
Length Constraints: Maximum length of 256.
Pattern: . *
Required: No

ContainerEntrypoint
The entrypoint for a container used to run a processing job.
Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 100 items.
Length Constraints: Maximum length of 256.
Pattern: . *
Required: No

ImageUri
The container image to be run by the processing job.
Type: String
Length Constraints: Maximum length of 255.
Pattern: . *
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ArtifactSource
Service: Amazon SageMaker Service
A structure describing the source of an artifact.

Contents

SourceType
A list of source types.
Type: Array of ArtifactSourceType (p. 893) objects
Required: No

SourceUri
The URI of the source.
Type: String
Length Constraints: Maximum length of 2048.
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ArtifactSourceType

Service: Amazon SageMaker Service

The ID and ID type of an artifact source.

Contents

SourceIdType

The type of ID.

Type: String

Valid Values: MD5Hash | S3ETag | S3Version | Custom

Required: Yes

Value

The ID.

Type: String

Length Constraints: Maximum length of 256.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ArtifactSummary
Service: Amazon SageMaker Service

Lists a summary of the properties of an artifact. An artifact represents a URI addressable object or data. Some examples are a dataset and a model.

Contents

ArtifactArn

The Amazon Resource Name (ARN) of the artifact.

Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:artifact/.*
Required: No

ArtifactName

The name of the artifact.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9\-]*(\-[a-zA-Z0-9\-])\{0,119\}
Required: No

ArtifactType

The type of the artifact.

Type: String
Length Constraints: Maximum length of 256.
Required: No

CreationTime

When the artifact was created.

Type: Timestamp
Required: No

LastModifiedTime

When the artifact was last modified.

Type: Timestamp
Required: No

Source

The source of the artifact.

Type: ArtifactSource (p. 892) object
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**AssociationSummary**

Service: Amazon SageMaker Service

Lists a summary of the properties of an association. An association is an entity that links other lineage or experiment entities. An example would be an association between a training job and a model.

**Contents**

**AssociationType**

The type of the association.

Type: String

Valid Values: ContributedTo | AssociatedWith | DerivedFrom | Produced

Required: No

**CreatedBy**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

Required: No

**CreationTime**

When the association was created.

Type: Timestamp

Required: No

**DestinationArn**

The Amazon Resource Name (ARN) of the destination.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*`

Required: No

**DestinationName**

The name of the destination.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9-]*([-*][a-zA-Z0-9]{0,119})`

Required: No

**DestinationType**

The destination type.
Type: String
Length Constraints: Maximum length of 256.
Required: No

**SourceArn**

The ARN of the source.

Type: String
Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

Required: No

**SourceName**

The name of the source.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]{0,119}$

Required: No

**SourceType**

The source type.

Type: String
Length Constraints: Maximum length of 256.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AsyncInferenceClientConfig

Service: Amazon SageMaker Service

Configures the behavior of the client used by SageMaker to interact with the model container during asynchronous inference.

Contents

MaxConcurrentInvocationsPerInstance

The maximum number of concurrent requests sent by the SageMaker client to the model container. If no value is provided, SageMaker chooses an optimal value.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AsyncInferenceConfig
Service: Amazon SageMaker Service

Specifies configuration for how an endpoint performs asynchronous inference.

Contents

ClientConfig

Configures the behavior of the client used by SageMaker to interact with the model container during asynchronous inference.

Type: AsyncInferenceClientConfig (p. 898) object

Required: No

OutputConfig

Specifies the configuration for asynchronous inference invocation outputs.

Type: AsyncInferenceOutputConfig (p. 901) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AsyncInferenceNotificationConfig

Service: Amazon SageMaker Service

Specifies the configuration for notifications of inference results for asynchronous inference.

Contents

ErrorTopic

Amazon SNS topic to post a notification to when inference fails. If no topic is provided, no notification is sent on failure.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sns:[a-z0-9\-]*:[0-9]{12}:[a-zA-Z0-9_.-]+

Required: No

SuccessTopic

Amazon SNS topic to post a notification to when inference completes successfully. If no topic is provided, no notification is sent on success.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sns:[a-z0-9\-]*:[0-9]{12}:[a-zA-Z0-9_.-]+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AsyncInferenceOutputConfig

Service: Amazon SageMaker Service

Specifies the configuration for asynchronous inference invocation outputs.

Contents

KmsKeyId

The AWS Key Management Service (AWS KMS) key that SageMaker uses to encrypt the asynchronous inference output in Amazon S3.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

NotificationConfig

Specifies the configuration for notifications of inference results for asynchronous inference.

Type: AsyncInferenceNotificationConfig (p. 900) object

Required: No

S3OutputPath

The Amazon S3 location to upload inference responses to.

Type: String

Length Constraints: Maximum length of 512.

Pattern: ^(https|s3)://([^/])/(.*)$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AthenaDatasetDefinition

Service: Amazon SageMaker Service

Configuration for Athena Dataset Definition input.

Contents

Catalog

The name of the data catalog used in Athena query execution.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: 

Required: Yes

Database

The name of the database used in the Athena query execution.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: . *

Required: Yes

KmsKeyId

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt data generated from an Athena query execution.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: . *

Required: No

OutputCompression

The compression used for Athena query results.

Type: String

Valid Values: GZIP | SNAPPY | ZLIB

Required: No

OutputFormat

The data storage format for Athena query results.

Type: String

Valid Values: PARQUET | ORC | AVRO | JSON | TEXTFILE

Required: Yes
**OutputS3Uri**

The location in Amazon S3 where Athena query results are stored.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `(https|s3)://([^/]+)/(\.)*$`

Required: Yes

**QueryString**

The SQL query statements, to be executed.

Type: String


Pattern: `\s\S+`

Required: Yes

**WorkGroup**

The name of the workgroup in which the Athena query is being started.

Type: String


Pattern: `[a-zA-Z0-9\-\_\-]+`

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLCandidate
Service: Amazon SageMaker Service

Information about a candidate produced by an AutoML training job, including its status, steps, and other properties.

Contents

CandidateName

The name of the candidate.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Required: Yes

CandidateProperties

The properties of an AutoML candidate job.

Type: CandidateProperties (p. 934) object

Required: No

CandidateStatus

The candidate's status.

Type: String

Valid Values: Completed | InProgress | Failed | Stopped | Stopping

Required: Yes

CandidateSteps

Information about the candidate's steps.

Type: Array of AutoMLCandidateStep (p. 907) objects

Required: Yes

CreationTime

The creation time.

Type: Timestamp

Required: Yes

EndTime

The end time.

Type: Timestamp

Required: No

FailureReason

The failure reason.

Type: String
AutoMLCandidate

Length Constraints: Maximum length of 1024.

Required: No

**FinalAutoMLJobObjectiveMetric**

The best candidate result from an AutoML training job.

Type: FinalAutoMLJobObjectiveMetric (p. 1049) object

Required: No

**InferenceContainers**

Information about the inference container definitions.

Type: Array of AutoMLContainerDefinition (p. 910) objects

Array Members: Maximum number of 5 items.

Required: No

**LastModifiedTime**

The last modified time.

Type: Timestamp

Required: Yes

**ObjectiveStatus**

The objective's status.

Type: String

Valid Values: Succeeded | Pending | Failed

Required: Yes

---

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLCandidateGenerationConfig

Service: Amazon SageMaker Service

Stores the config information for how a candidate is generated (optional).

Contents

FeatureSpecificationS3Uri

A URL to the Amazon S3 data source containing selected features from the input data source to run an Autopilot job (optional). This file should be in json format as shown below:

{ "FeatureAttributeNames": ["col1", "col2", ...] }.

The key name FeatureAttributeNames is fixed. The values listed in ["col1", "col2", ...] is case sensitive and should be a list of strings containing unique values that are a subset of the column names in the input data. The list of columns provided must not include the target column.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)/?(.*)$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLCandidateStep
Service: Amazon SageMaker Service

Information about the steps for a candidate and what step it is working on.

Contents

CandidateStepArn
The ARN for the candidate's step.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:.*/.*
Required: Yes

CandidateStepName
The name for the candidate's step.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 64.
Required: Yes

CandidateStepType
Whether the candidate is at the transform, training, or processing step.
Type: String
Valid Values: AWS::SageMaker::TrainingJob | AWS::SageMaker::TransformJob | AWS::SageMaker::ProcessingJob
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLChannel
Service: Amazon SageMaker Service

A channel is a named input source that training algorithms can consume. The validation dataset size is limited to less than 2 GB. The training dataset size must be less than 100 GB. For more information, see `Channel (p. 941)`.

**Note**
A validation dataset must contain the same headers as the training dataset.

**Contents**

**ChannelType**
The channel type (optional) is an `enum` string. The default value is `training`. Channels for training and validation must share the same `ContentType` and `TargetAttributeName`. For information on specifying training and validation channel types, see `How to specify training and validation datasets`.

Type: String

Valid Values: `training` | `validation`

Required: No

**CompressionType**
You can use `Gzip` or `None`. The default value is `None`.

Type: String

Valid Values: `None` | `Gzip`

Required: No

**ContentType**
The content type of the data from the input source. You can use `text/csv;header=present` or `x-application/vnd.amazon+parquet`. The default value is `text/csv;header=present`.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `.*`

Required: No

**DataSource**
The data source for an AutoML channel.

Type: `AutoMLDataSource (p. 912)` object

Required: Yes

**TargetAttributeName**
The name of the target variable in supervised learning, usually represented by 'y'.

Type: String
Length Constraints: Minimum length of 1.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLContainerDefinition
Service: Amazon SageMaker Service

A list of container definitions that describe the different containers that make up an AutoML candidate. For more information, see ContainerDefinition (p. 956).

Contents

Environment

The environment variables to set in the container. For more information, see ContainerDefinition (p. 956).

Type: String to string map

Map Entries: Maximum number of 16 items.

Key Length Constraints: Maximum length of 1024.

Key Pattern: [a-zA-Z_][a-zA-Z0-9_]*

Value Length Constraints: Maximum length of 1024.

Value Pattern: \Ss*\s*

Required: No

Image

The Amazon Elastic Container Registry (Amazon ECR) path of the container. For more information, see ContainerDefinition (p. 956).

Type: String

Length Constraints: Maximum length of 255.

Pattern: \S+

Required: Yes

ModelDataUrl

The location of the model artifacts. For more information, see ContainerDefinition (p. 956).

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^https:\/s3\://([^/]+)\/[.\s]*$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLDataSource

Service: Amazon SageMaker Service

The data source for the Autopilot job.

Contents

S3DataSource

The Amazon S3 location of the input data.

Type: AutoMLS3DataSource (p. 923) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLDataSplitConfig

Service: Amazon SageMaker Service

This structure specifies how to split the data into train and validation datasets. The validation and training datasets must contain the same headers. The validation dataset must be less than 2 GB in size.

Contents

ValidationFraction

The validation fraction (optional) is a float that specifies the portion of the training dataset to be used for validation. The default value is 0.2, and values must be greater than 0 and less than 1. We recommend setting this value to be less than 0.5.

Type: Float

Valid Range: Minimum value of 0. Maximum value of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLJobArtifacts
Service: Amazon SageMaker Service

The artifacts that are generated during an AutoML job.

Contents

CandidateDefinitionNotebookLocation
The URL of the notebook location.
Type: String
Length Constraints: Minimum length of 1.
Required: No

DataExplorationNotebookLocation
The URL of the notebook location.
Type: String
Length Constraints: Minimum length of 1.
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLJobCompletionCriteria
Service: Amazon SageMaker Service

How long a job is allowed to run, or how many candidates a job is allowed to generate.

Contents

MaxAutoMLJobRuntimeInSeconds
The maximum runtime, in seconds, an AutoML job has to complete.

If an AutoML job exceeds the maximum runtime, the job is stopped automatically and its processing is ended gracefully. The AutoML job identifies the best model whose training was completed and marks it as the best-performing model. Any unfinished steps of the job, such as automatic one-click Autopilot model deployment, will not be completed.

Type: Integer
Valid Range: Minimum value of 1.
Required: No

MaxCandidates
The maximum number of times a training job is allowed to run.

Type: Integer
Valid Range: Minimum value of 1.
Required: No

MaxRuntimePerTrainingJobInSeconds
The maximum time, in seconds, that each training job is allowed to run as part of a hyperparameter tuning job. For more information, see the StoppingCondition (p. 1370) used by the CreateHyperParameterTuningJob (p. 90) action.

Type: Integer
Valid Range: Minimum value of 1.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLJobConfig

Service: Amazon SageMaker Service

A collection of settings used for an AutoML job.

Contents

CandidateGenerationConfig

The configuration for generating a candidate for an AutoML job (optional).

Type: AutoMLCandidateGenerationConfig (p. 906) object

Required: No

CompletionCriteria

How long an AutoML job is allowed to run, or how many candidates a job is allowed to generate.

Type: AutoMLJobCompletionCriteria (p. 915) object

Required: No

DataSplitConfig

The configuration for splitting the input training dataset.

Type: AutoMLDataSplitConfig

Type: AutoMLDataSplitConfig (p. 913) object

Required: No

SecurityConfig

The security configuration for traffic encryption or Amazon VPC settings.

Type: AutoMLSecurityConfig (p. 924) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLJobObjective

Service: Amazon SageMaker Service

Specifies a metric to minimize or maximize as the objective of a job.

Contents

MetricName

The name of the objective metric used to measure the predictive quality of a machine learning system. This metric is optimized during training to provide the best estimate for model parameter values from data.

Here are the options:

- **MSE**: The mean squared error (MSE) is the average of the squared differences between the predicted and actual values. It is used for regression. MSE values are always positive: the better a model is at predicting the actual values, the smaller the MSE value is. When the data contains outliers, they tend to dominate the MSE, which might cause subpar prediction performance.

- **Accuracy**: The ratio of the number of correctly classified items to the total number of (correctly and incorrectly) classified items. It is used for binary and multiclass classification. It measures how close the predicted class values are to the actual values. Accuracy values vary between zero and one: one indicates perfect accuracy and zero indicates perfect inaccuracy.

- **F1**: The F1 score is the harmonic mean of the precision and recall. It is used for binary classification into classes traditionally referred to as positive and negative. Predictions are said to be true when they match their actual (correct) class and false when they do not. Precision is the ratio of the true positive predictions to all positive predictions (including the false positives) in a data set and measures the quality of the prediction when it predicts the positive class. Recall (or sensitivity) is the ratio of the true positive predictions to all actual positive instances and measures how completely a model predicts the actual class members in a data set. The standard F1 score weighs precision and recall equally. But which metric is paramount typically depends on specific aspects of a problem. F1 scores vary between zero and one: one indicates the best possible performance and zero the worst.

- **AUC**: The area under the curve (AUC) metric is used to compare and evaluate binary classification by algorithms such as logistic regression that return probabilities. A threshold is needed to map the probabilities into classifications. The relevant curve is the receiver operating characteristic curve that plots the true positive rate (TPR) of predictions (or recall) against the false positive rate (FPR) as a function of the threshold value, above which a prediction is considered positive. Increasing the threshold results in fewer false positives but more false negatives. AUC is the area under this receiver operating characteristic curve and so provides an aggregated measure of the model performance across all possible classification thresholds. The AUC score can also be interpreted as the probability that a randomly selected positive data point is more likely to be predicted positive than a randomly selected negative example. AUC scores vary between zero and one: a score of one indicates perfect accuracy and a score of one half indicates that the prediction is not better than a random classifier. Values under one half predict less accurately than a random predictor. But such consistently bad predictors can simply be inverted to obtain better than random predictors.

- **F1macro**: The F1macro score applies F1 scoring to multiclass classification. In this context, you have multiple classes to predict. You just calculate the precision and recall for each class as you did for the positive class in binary classification. Then, use these values to calculate the F1 score for each class and average them to obtain the F1macro score. F1macro scores vary between zero and one: one indicates the best possible performance and zero the worst.

If you do not specify a metric explicitly, the default behavior is to automatically use:

- **MSE**: for regression.
- **F1**: for binary classification.
• **Accuracy**: for multiclass classification.
  
  Type: String
  
  Valid Values: Accuracy | MSE | F1 | F1macro | AUC
  
  Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLJobSummary

Service: Amazon SageMaker Service
Provides a summary about an AutoML job.

Contents

AutoMLJobArn
The ARN of the AutoML job.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*
Required: Yes

AutoMLJobName
The name of the AutoML job you are requesting.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}
Required: Yes

AutoMLJobSecondaryStatus
The secondary status of the AutoML job.
Type: String
Required: Yes

AutoMLJobStatus
The status of the AutoML job.
Type: String
Valid Values: Completed | InProgress | Failed | Stopped | Stopping
Required: Yes

CreationTime
When the AutoML job was created.
Type: Timestamp
Required: Yes
EndTime
The end time of an AutoML job.
Type: Timestamp
Required: No
FailureReason
The failure reason of an AutoML job.
Type: String
Length Constraints: Maximum length of 1024.
Required: No
LastModifedTime
When the AutoML job was last modified.
Type: Timestamp
Required: Yes
PartialFailureReasons
The list of reasons for partial failures within an AutoML job.
Type: Array of AutoMLPartialFailureReason (p. 922) objects
Array Members: Minimum number of 1 item. Maximum number of 5 items.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLOutputDataConfig
Service: Amazon SageMaker Service
The output data configuration.

Contents

KmsKeyId

The AWS KMS encryption key ID.
Type: String
Length Constraints: Maximum length of 2048.
Pattern: .*
Required: No

S3OutputPath

The Amazon S3 output path. Must be 128 characters or less.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: ^(https|s3)://([a-zA-Z0-9]+)((/[^/]+)?)?(.*)$
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLPartialFailureReason
Service: Amazon SageMaker Service
The reason for a partial failure of an AutoML job.

Contents

PartialFailureMessage
The message containing the reason for a partial failure of an AutoML job.

Type: String
Length Constraints: Maximum length of 1024.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLS3DataSource
Service: Amazon SageMaker Service

The Amazon S3 data source.

Contents

S3DataType

The data type.

A ManifestFile should have the format shown below:

```javascript
[ {"prefix": "s3://DOC-EXAMPLE-BUCKET/DOC-EXAMPLE-FOLDER/DOC-EXAMPLE-PREFIX/"},

"DOC-EXAMPLE-RELATIVE-PATH/DOC-EXAMPLE-FOLDER/DATA-1",

"DOC-EXAMPLE-RELATIVE-PATH/DOC-EXAMPLE-FOLDER/DATA-2",

... "DOC-EXAMPLE-RELATIVE-PATH/DOC-EXAMPLE-FOLDER/DATA-N" ]
```

An S3Prefix should have the following format:

`s3://DOC-EXAMPLE-BUCKET/DOC-EXAMPLE-FOLDER-OR-FILE`

Type: String

Valid Values: ManifestFile | S3Prefix

Required: Yes

S3Uri

The URL to the Amazon S3 data source.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(https|s3)://([^/]+)/?([^/]+?)$`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoMLSecurityConfig
Service: Amazon SageMaker Service

Security options.

Contents

EnableInterContainerTrafficEncryption
Whether to use traffic encryption between the container layers.
Type: Boolean
Required: No

VolumeKmsKeyId
The key used to encrypt stored data.
Type: String
Length Constraints: Maximum length of 2048.
Pattern: . *
Required: No

VpcConfig
The VPC configuration.
Type: VpcConfig (p. 1456) object
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
AutoRollbackConfig

Service: Amazon SageMaker Service

Automatic rollback configuration for handling endpoint deployment failures and recovery.

Contents

Alarms

List of CloudWatch alarms in your account that are configured to monitor metrics on an endpoint. If any alarms are tripped during a deployment, SageMaker rolls back the deployment.

Type: Array of Alarm (p. 866) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
BatchDescribeModelPackageError

Service: Amazon SageMaker Service

The error code and error description associated with the resource.

Contents

ErrorCode

Type: String
Required: Yes

ErrorResponse

Type: String
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
BatchDescribeModelPackageSummary
Service: Amazon SageMaker Service

Provides summary information about the model package.

Contents

**CreationTime**

The creation time of the model package summary.

Type: Timestamp

Required: Yes

**InferenceSpecification**

Defines how to perform inference generation after a training job is run.

Type: InferenceSpecification (p. 1105) object

Required: Yes

**ModelApprovalStatus**

The approval status of the model.

Type: String

Valid Values: Approved | Rejected | PendingManualApproval

Required: No

**ModelPackageArn**

The Amazon Resource Name (ARN) of the model package.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:model-package/.*

Required: Yes

**ModelPackageDescription**

The description of the model package.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*

Required: No

**ModelPackageGroupName**

The group name for the model package.

Type: String

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$
Required: Yes

**ModelPackageStatus**

The status of the mortgage package.
Type: String

Valid Values: Pending | InProgress | Completed | Failed | Deleting
Required: Yes

**ModelPackageVersion**

The version number of a versioned model.
Type: Integer

Valid Range: Minimum value of 1.
Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Bias
Service: Amazon SageMaker Service
Contains bias metrics for a model.

Contents

PostTrainingReport
The post-training bias report for a model.
Type: MetricsSource (p. 1148) object
Required: No

PreTrainingReport
The pre-training bias report for a model.
Type: MetricsSource (p. 1148) object
Required: No

Report
The bias report for a model
Type: MetricsSource (p. 1148) object
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**BlueGreenUpdatePolicy**  
Service: Amazon SageMaker Service

Update policy for a blue/green deployment. If this update policy is specified, SageMaker creates a new fleet during the deployment while maintaining the old fleet. SageMaker flips traffic to the new fleet according to the specified traffic routing configuration. Only one update policy should be used in the deployment configuration. If no update policy is specified, SageMaker uses a blue/green deployment strategy with all at once traffic shifting by default.

**Contents**

MaximumExecutionTimeoutInSeconds

Maximum execution timeout for the deployment. Note that the timeout value should be larger than the total waiting time specified in TerminationWaitInSeconds and WaitIntervalInSeconds.

Type: Integer


Required: No

TerminationWaitInSeconds

Additional waiting time in seconds after the completion of an endpoint deployment before terminating the old endpoint fleet. Default is 0.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 3600.

Required: No

TrafficRoutingConfiguration

Defines the traffic routing strategy to shift traffic from the old fleet to the new fleet during an endpoint deployment.

Type: TrafficRoutingConfig (p. 1383) object

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CacheHitResult
Service: Amazon SageMaker Service
Details on the cache hit of a pipeline execution step.

Contents

SourcePipelineExecutionArn
The Amazon Resource Name (ARN) of the pipeline execution.
Type: String
Length Constraints: Maximum length of 256.
Pattern: ^arn:aws[\w\-]*:sagemaker:[\w\-]*:[0-9]{12}:pipeline\/*.*/\execution\/.*$
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CallbackStepMetadata

Service: Amazon SageMaker Service

Metadata about a callback step.

Contents

CallbackToken

The pipeline generated token from the Amazon SQS queue.

Type: String

Length Constraints: Fixed length of 10.

Pattern: ^[a-zA-Z0-9]+$

Required: No

OutputParameters

A list of the output parameters of the callback step.

Type: Array of OutputParameter (p. 1246) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

SqsQueueUrl

The URL of the Amazon Simple Queue Service (Amazon SQS) queue used by the callback step.

Type: String

Length Constraints: Maximum length of 256.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CandidateArtifactLocations

Service: Amazon SageMaker Service

The location of artifacts for an AutoML candidate job.

**Contents**

**Explainability**

The Amazon S3 prefix to the explainability artifacts generated for the AutoML candidate.

Type: String

Length Constraints: Minimum length of 1.

Required: Yes

**ModelInsights**

The Amazon S3 prefix to the model insight artifacts generated for the AutoML candidate.

Type: String

Length Constraints: Minimum length of 1.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CandidateProperties
Service: Amazon SageMaker Service
The properties of an AutoML candidate job.

Contents

CandidateArtifactLocations
The Amazon S3 prefix to the artifacts generated for an AutoML candidate.
Type: CandidateArtifactLocations (p. 933) object
Required: No

CandidateMetrics
Information about the candidate metrics for an AutoML job.
Type: Array of MetricDatum (p. 1146) objects
Array Members: Minimum number of 0 items. Maximum number of 40 items.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CapacitySize
Service: Amazon SageMaker Service
Specifies the endpoint capacity to activate for production.

Contents

Type
Specifies the endpoint capacity type.
- INSTANCE_COUNT: The endpoint activates based on the number of instances.
- CAPACITY_PERCENT: The endpoint activates based on the specified percentage of capacity.

Type: String
Valid Values: INSTANCE_COUNT | CAPACITY_PERCENT
Required: Yes

Value
Defines the capacity size, either as a number of instances or a capacity percentage.

Type: Integer
Valid Range: Minimum value of 1.
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CaptureContentTypeHeader
Service: Amazon SageMaker Service

Configuration specifying how to treat different headers. If no headers are specified SageMaker will by default base64 encode when capturing the data.

Contents

CsvContentTypes
The list of all content type headers that SageMaker will treat as CSV and capture accordingly.
Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 10 items.
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*\/[a-zA-Z0-9](-*[a-zA-Z0-9.])* Required: No

JsonContentTypes
The list of all content type headers that SageMaker will treat as JSON and capture accordingly.
Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 10 items.
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*\/[a-zA-Z0-9](-*[a-zA-Z0-9.])* Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CaptureOption
Service: Amazon SageMaker Service
Specifies data Model Monitor will capture.

Contents

CaptureMode
Specify the boundary of data to capture.

Type: String

Valid Values: Input | Output

Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CategoricalParameter
Service: Amazon SageMaker Service

Environment parameters you want to benchmark your load test against.

Contents

Name

The Name of the environment variable.
Type: String
Length Constraints: Maximum length of 64.
Required: Yes

Value

The list of values you can pass.
Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 3 items.
Length Constraints: Maximum length of 128.
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CategoricalParameterRange
Service: Amazon SageMaker Service
A list of categorical hyperparameters to tune.

Contents

Name
The name of the categorical hyperparameter to tune.
Type: String
Length Constraints: Maximum length of 256.
Pattern: . *
Required: Yes

Values
A list of the categories for the hyperparameter.
Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 20 items.
Length Constraints: Maximum length of 256.
Pattern: . *
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CategoricalParameterRangeSpecification
Service: Amazon SageMaker Service

Defines the possible values for a categorical hyperparameter.

Contents

Values

The allowed categories for the hyperparameter.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Length Constraints: Maximum length of 256.

Pattern: .*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Channel
Service: Amazon SageMaker Service

A channel is a named input source that training algorithms can consume.

Contents

**ChannelName**

The name of the channel.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: [A-Za-z0-9\-_]+  

Required: Yes

**CompressionType**

If training data is compressed, the compression type. The default value is None. CompressionType is used only in Pipe input mode. In File mode, leave this field unset or set it to None.

Type: String

Valid Values: None | Gzip

Required: No

**ContentType**

The MIME type of the data.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

**DataSource**

The location of the channel data.

Type: [DataSource](p. 979) object

Required: Yes

**InputMode**

(Optional) The input mode to use for the data channel in a training job. If you don't set a value for InputMode, SageMaker uses the value set for TrainingInputMode. Use this parameter to override the TrainingInputMode setting in a [AlgorithmSpecification](p. 867) request when you have a channel that needs a different input mode from the training job's general setting. To download the data from Amazon Simple Storage Service (Amazon S3) to the provisioned ML storage volume, and mount the directory to a Docker volume, use File input mode. To stream data directly from Amazon S3 to the container, choose Pipe input mode.

To use a model for incremental training, choose File input model.

Type: String
Valid Values: Pipe | File | FastFile

Required: No

**RecordWrapperType**

Specify RecordIO as the value when input data is in raw format but the training algorithm requires the RecordIO format. In this case, SageMaker wraps each individual S3 object in a RecordIO record. If the input data is already in RecordIO format, you don't need to set this attribute. For more information, see Create a Dataset Using RecordIO.

In File mode, leave this field unset or set it to None.

Type: String

Valid Values: None | RecordIO

Required: No

**ShuffleConfig**

A configuration for a shuffle option for input data in a channel. If you use S3Prefix for S3DataType, this shuffles the results of the S3 key prefix matches. If you use ManifestFile, the order of the S3 object references in the ManifestFile is shuffled. If you use AugmentedManifestFile, the order of the JSON lines in the AugmentedManifestFile is shuffled. The shuffling order is determined using the Seed value.

For Pipe input mode, shuffling is done at the start of every epoch. With large datasets this ensures that the order of the training data is different for each epoch, it helps reduce bias and possible overfitting. In a multi-node training job when ShuffleConfig is combined with S3DataDistributionType of ShardedByS3Key, the data is shuffled across nodes so that the content sent to a particular node on the first epoch might be sent to a different node on the second epoch.

Type: ShuffleConfig (p. 1366) object

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ChannelSpecification
Service: Amazon SageMaker Service
Defines a named input source, called a channel, to be used by an algorithm.

Contents

Description

A brief description of the channel.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*
Required: No

IsRequired

Indicates whether the channel is required by the algorithm.
Type: Boolean
Required: No

Name

The name of the channel.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 64.
Pattern: [A-Za-z0-9._-]+
Required: Yes

SupportedCompressionTypes

The allowed compression types, if data compression is used.
Type: Array of strings
Valid Values: None | Gzip
Required: No

SupportedContentTypes

The supported MIME types for the data.
Type: Array of strings
Length Constraints: Maximum length of 256.
Pattern: .*
Required: Yes

SupportedInputModes

The allowed input mode, either FILE or PIPE.
In FILE mode, Amazon SageMaker copies the data from the input source onto the local Amazon Elastic Block Store (Amazon EBS) volumes before starting your training algorithm. This is the most commonly used input mode.

In PIPE mode, Amazon SageMaker streams input data from the source directly to your algorithm without using the EBS volume.

Type: Array of strings

Array Members: Minimum number of 1 item.

Valid Values: Pipe | File | FastFile

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CheckpointConfig
Service: Amazon SageMaker Service

Contains information about the output location for managed spot training checkpoint data.

Contents

LocalPath
(Optional) The local directory where checkpoints are written. The default directory is /opt/ml/checkpoints/.

Type: String
Length Constraints: Maximum length of 4096.
Pattern: .*
Required: No

S3Uri
Identifies the S3 path where you want SageMaker to store checkpoints. For example, s3://bucket-name/key-name-prefix.

Type: String
Length Constraints: Maximum length of 1024.
Pattern: ^(https|s3)://([^/]+)/([^/]+)/?([^/]+)$
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**ClarifyCheckStepMetadata**

*Service: Amazon SageMaker Service*

The container for the metadata for the ClarifyCheck step. For more information, see the topic on ClarifyCheck step in the *Amazon SageMaker Developer Guide*.

**Contents**

**BaselineUsedForDriftCheckConstraints**

The Amazon S3 URI of baseline constraints file to be used for the drift check.

- **Type:** String
- **Length Constraints:** Maximum length of 1024.
- **Required:** No

**CalculatedBaselineConstraints**

The Amazon S3 URI of the newly calculated baseline constraints file.

- **Type:** String
- **Length Constraints:** Maximum length of 1024.
- **Required:** No

**CheckJobArn**

The Amazon Resource Name (ARN) of the check processing job that was run by this step's execution.

- **Type:** String
- **Length Constraints:** Maximum length of 256.
- **Required:** No

**CheckType**

The type of the Clarify Check step

- **Type:** String
- **Length Constraints:** Maximum length of 256.
- **Required:** No

**ModelPackageName**

The model package group name.

- **Type:** String
- **Length Constraints:** Maximum length of 256.
- **Required:** No

**RegisterNewBaseline**

This flag indicates if a newly calculated baseline can be accessed through step properties `BaselineUsedForDriftCheckConstraints` and `BaselineUsedForDriftCheckStatistics`. If it is set to `False`, the previous baseline of the configured check type must also be available. These can be accessed through the `BaselineUsedForDriftCheckConstraints` property.
Type: Boolean
Required: No

**SkipCheck**

This flag indicates if the drift check against the previous baseline will be skipped or not. If it is set to `False`, the previous baseline of the configured check type must be available.

Type: Boolean
Required: No

**ViolationReport**

The Amazon S3 URI of the violation report if violations are detected.

Type: String
Length Constraints: Maximum length of 1024.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CodeRepositorySummary

Service: Amazon SageMaker Service

Specifies summary information about a Git repository.

Contents

**CodeRepositoryArn**

The Amazon Resource Name (ARN) of the Git repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:code-repository/.*`

Required: Yes

**CodeRepositoryName**

The name of the Git repository.

Type: String


Pattern: `^[a-zA-Z0-9\-]*(\-*[a-zA-Z0-9\-])\{0,62}\$`

Required: Yes

**CreationTime**

The date and time that the Git repository was created.

Type: Timestamp

Required: Yes

**GitConfig**

Configuration details for the Git repository, including the URL where it is located and the ARN of the AWS Secrets Manager secret that contains the credentials used to access the repository.

Type: `GitConfig (p. 1054)` object

Required: No

**LastModifiedTime**

The date and time that the Git repository was last modified.

Type: Timestamp

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
CognitoConfig

Service: Amazon SageMaker Service

Use this parameter to configure your Amazon Cognito workforce. A single Cognito workforce is created using and corresponds to a single Amazon Cognito user pool.

Contents

ClientId

The client ID for your Amazon Cognito user pool.

Type: String


Pattern: \[\w\-]+\[\w\-]+\[0-9a-zA-Z]+\]

Required: Yes

UserPool

A user pool is a user directory in Amazon Cognito. With a user pool, your users can sign in to your web or mobile app through Amazon Cognito. Your users can also sign in through social identity providers like Google, Facebook, Amazon, or Apple, and through SAML identity providers.

Type: String


Pattern: \[\w\-]+\[0-9a-zA-Z]+\]

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CognitoMemberDefinition
Service: Amazon SageMaker Service

Identifies a Amazon Cognito user group. A user group can be used in on or more work teams.

Contents

ClientId
An identifier for an application client. You must create the app client ID using Amazon Cognito.
Type: String
Pattern: [- ~]+
Required: Yes

UserGroup
An identifier for a user group.
Type: String
Pattern: [\p{L}\p{M}\p{S}\p{N}\p{P}]+
Required: Yes

UserPool
An identifier for a user pool. The user pool must be in the same region as the service that you are calling.
Type: String
Pattern: [\w-]+_[0-9a-zA-Z]+
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CollectionConfiguration
Service: Amazon SageMaker Service

Configuration information for the Debugger output tensor collections.

Contents

CollectionName

The name of the tensor collection. The name must be unique relative to other rule configuration names.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: .*

Required: No

CollectionParameters

Parameter values for the tensor collection. The allowed parameters are "name", "include_regex", "reduction_config", "save_config", "tensor_names", and "save_histogram".

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 20 items.

Key Length Constraints: Minimum length of 1. Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 256.

Value Pattern: .*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CompilationJobSummary
Service: Amazon SageMaker Service
A summary of a model compilation job.

Contents

CompilationEndTime
The time when the model compilation job completed.
Type: Timestamp
Required: No

CompilationJobArn
The Amazon Resource Name (ARN) of the model compilation job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[\a-z\-]*:sagemaker:[\a-z0-9\-]*:[0-9]{12}:compilation-job/.*
Required: Yes

CompilationJobName
The name of the model compilation job that you want a summary for.
Type: String
Pattern:^[a-zA-Z0-9\-\d\*\[\]\{\}\(\)\।\?\,\.\:\;\_\+\-\#\$\%\^\&\*\(\)\{\}\[\]\|\-\=\+\?\.\;\[\]\*\#\$\%\^\&\*\(\)\{\}\[\]\|\-\=\+\?\]
Required: Yes

CompilationJobStatus
The status of the model compilation job.
Type: String
Valid Values: INPROGRESS | COMPLETED | FAILED | STARTING | STOPPING | STOPPED
Required: Yes

CompilationStartTime
The time when the model compilation job started.
Type: Timestamp
Required: No

CompilationTargetDevice
The type of device that the model will run on after the compilation job has completed.
Type: String
Valid Values: lambda | ml_m4 | ml_m5 | ml_c4 | ml_c5 | ml_p2 | ml_p3 | ml_g4dn | ml_inf1 | ml_eia2 | jetson_tx1 | jetson_tx2 | jetsonNano | jetson_xavier
CompilationJobSummary

<table>
<thead>
<tr>
<th>rasp3b</th>
<th>imx8qm</th>
<th>deplens</th>
<th>rk3399</th>
<th>rk3288</th>
<th>aisage</th>
<th>sbe_c</th>
<th>qcs605</th>
</tr>
</thead>
<tbody>
<tr>
<td>qcs603</td>
<td>sitara_am57x</td>
<td>amba_cv2</td>
<td>amba_cv22</td>
<td>amba_cv25</td>
<td>x86_win32</td>
<td>x86_win64</td>
<td>coreml</td>
</tr>
</tbody>
</table>

Required: No

CompilationTargetPlatformAccelerator

The type of accelerator that the model will run on after the compilation job has completed.

Type: String

Valid Values: INTEL_GRAPHICS | MALI | NVIDIA | NNA

Required: No

CompilationTargetPlatformArch

The type of architecture that the model will run on after the compilation job has completed.

Type: String

Valid Values: X86_64 | X86 | ARM64 | ARM_EABI | ARM_EABIHF

Required: No

CompilationTargetPlatformOs

The type of OS that the model will run on after the compilation job has completed.

Type: String

Valid Values: ANDROID | LINUX

Required: No

CreationTime

The time when the model compilation job was created.

Type: Timestamp

Required: Yes

LastModifiedTime

The time when the model compilation job was last modified.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ConditionStepMetadata
Service: Amazon SageMaker Service
Metadata for a Condition step.

Contents

Outcome
The outcome of the Condition step evaluation.
Type: String
Valid Values: True | False
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**ContainerDefinition**

Service: Amazon SageMaker Service

Describes the container, as part of model definition.

**Contents**

**ContainerHostname**

This parameter is ignored for models that contain only a PrimaryContainer.

When a ContainerDefinition is part of an inference pipeline, the value of the parameter uniquely identifies the container for the purposes of logging and metrics. For information, see Use Logs and Metrics to Monitor an Inference Pipeline. If you don't specify a value for this parameter for a ContainerDefinition that is part of an inference pipeline, a unique name is automatically assigned based on the position of the ContainerDefinition in the pipeline. If you specify a value for the ContainerHostName for any ContainerDefinition that is part of an inference pipeline, you must specify a value for the ContainerHostName parameter of every ContainerDefinition in that pipeline.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9]{0,62}$

Required: No

**Environment**

The environment variables to set in the Docker container. Each key and value in the Environment string to string map can have length of up to 1024. We support up to 16 entries in the map.

Type: String to string map

Map Entries: Maximum number of 16 items.

Key Length Constraints: Maximum length of 1024.

Key Pattern: [a-zA-Z_.][a-zA-Z0-9_.]*

Value Length Constraints: Maximum length of 1024.

Value Pattern: [\S\s]*

Required: No

**Image**

The path where inference code is stored. This can be either in Amazon EC2 Container Registry or in a Docker registry that is accessible from the same VPC that you configure for your endpoint. If you are using your own custom algorithm instead of an algorithm provided by SageMaker, the inference code must meet SageMaker requirements. SageMaker supports both registry/repository[:tag] and registry/repository[@digest] image path formats. For more information, see Using Your Own Algorithms with Amazon SageMaker

Type: String

Length Constraints: Maximum length of 255.

Pattern: [\S]+
ContainerDefinition

Required: No

**ImageConfig**

Specifies whether the model container is in Amazon ECR or a private Docker registry accessible from your Amazon Virtual Private Cloud (VPC). For information about storing containers in a private Docker registry, see Use a Private Docker Registry for Real-Time Inference Containers

Type: ImageConfig (p. 1098) object

Required: No

**InferenceSpecificationName**

The inference specification name in the model package version.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}\$

Required: No

**Mode**

Whether the container hosts a single model or multiple models.

Type: String

Valid Values: SingleModel | MultiModel

Required: No

**ModelDataUrl**

The S3 path where the model artifacts, which result from model training, are stored. This path must point to a single gzip compressed tar archive (.tar.gz suffix). The S3 path is required for SageMaker built-in algorithms, but not if you use your own algorithms. For more information on built-in algorithms, see Common Parameters.

**Note**

The model artifacts must be in an S3 bucket that is in the same region as the model or endpoint you are creating.

If you provide a value for this parameter, SageMaker uses AWS Security Token Service to download model artifacts from the S3 path you provide. AWS STS is activated in your IAM user account by default. If you previously deactivated AWS STS for a region, you need to reactivate AWS STS for that region. For more information, see Activating and Deactivating AWS STS in an AWS Region in the AWS Identity and Access Management User Guide.

**Important**

If you use a built-in algorithm to create a model, SageMaker requires that you provide a S3 path to the model artifacts in ModelDataUrl.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3):/(/[^/]+/)?(.*)$

Required: No

**ModelPackageName**

The name or Amazon Resource Name (ARN) of the model package to use to create the model.
Type: String


Pattern: (arn:aws[a-z\-]*:sagemaker:a-z0-9\*:0-9\{12\}:a-z\*/\/)?(a-zA-Z0-9\(a-zA-Z0-9\{0,62\}(?!-)(/0-9\{1,5\})?#

Required: No

MultiModelConfig

Specifies additional configuration for multi-model endpoints.

Type: MultiModelConfig (p. 1220) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ContextSource
Service: Amazon SageMaker Service

A structure describing the source of a context.

Contents

SourceId
The ID of the source.
Type: String
Length Constraints: Maximum length of 256.
Required: No

SourceType
The type of the source.
Type: String
Length Constraints: Maximum length of 256.
Required: No

SourceUri
The URI of the source.
Type: String
Length Constraints: Maximum length of 2048.
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ContextSummary

Service: Amazon SageMaker Service

Lists a summary of the properties of a context. A context provides a logical grouping of other entities.

Contents

ContextArn

The Amazon Resource Name (ARN) of the context.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-zA-Z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:context/.*

Required: No

ContextName

The name of the context.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9\-](-*[a-zA-Z0-9])\{0,119\}

Required: No

ContextType

The type of the context.

Type: String

Length Constraints: Maximum length of 256.

Required: No

CreationTime

When the context was created.

Type: Timestamp

Required: No

LastModifiedTime

When the context was last modified.

Type: Timestamp

Required: No

Source

The source of the context.

Type: ContextSource (p. 959) object

Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ContinuousParameterRange
Service: Amazon SageMaker Service

A list of continuous hyperparameters to tune.

Contents

MaxValue

The maximum value for the hyperparameter. The tuning job uses floating-point values between MinValue value and this value for tuning.

Type: String
Length Constraints: Maximum length of 256.
Pattern: .*
Required: Yes

MinValue

The minimum value for the hyperparameter. The tuning job uses floating-point values between this value and MaxValue for tuning.

Type: String
Length Constraints: Maximum length of 256.
Pattern: .*
Required: Yes

Name

The name of the continuous hyperparameter to tune.

Type: String
Length Constraints: Maximum length of 256.
Pattern: .*
Required: Yes

ScalingType

The scale that hyperparameter tuning uses to search the hyperparameter range. For information about choosing a hyperparameter scale, see Hyperparameter Scaling. One of the following values:

Auto
SageMaker hyperparameter tuning chooses the best scale for the hyperparameter.
Linear
Hyperparameter tuning searches the values in the hyperparameter range by using a linear scale.
Logarithmic
Hyperparameter tuning searches the values in the hyperparameter range by using a logarithmic scale.

Logarithmic scaling works only for ranges that have only values greater than 0.
ReverseLogarithmic

Hyperparameter tuning searches the values in the hyperparameter range by using a reverse logarithmic scale.

Reverse logarithmic scaling works only for ranges that are entirely within the range 0<=x<1.0.

Type: String

Valid Values: Auto | Linear | Logarithmic | ReverseLogarithmic

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ContinuousParameterRangeSpecification
Service: Amazon SageMaker Service
Defines the possible values for a continuous hyperparameter.

Contents

MaxValue
The maximum floating-point value allowed.
Type: String
Length Constraints: Maximum length of 256.
Pattern: . *
Required: Yes

MinValue
The minimum floating-point value allowed.
Type: String
Length Constraints: Maximum length of 256.
Pattern: . *
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
CustomImage

Service: Amazon SageMaker Service

A custom SageMaker image. For more information, see Bring your own SageMaker image.

Contents

AppImageConfigName

The name of the AppImageConfig.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

ImageName

The name of the CustomImage. Must be unique to your account.

Type: String


Pattern: ^[a-zA-Z0-9][-.]?[a-zA-Z0-9]{0,62}$

Required: Yes

ImageVersionNumber

The version number of the CustomImage.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DataCaptureConfig
Service: Amazon SageMaker Service

Configuration to control how SageMaker captures inference data.

Contents

CaptureContentTypeHeader
Configuration specifying how to treat different headers. If no headers are specified SageMaker will by default base64 encode when capturing the data.

Type: CaptureContentTypeHeader (p. 936) object

Required: No

CaptureOptions
Specifies data Model Monitor will capture. You can configure whether to collect only input, only output, or both

Type: Array of CaptureOption (p. 937) objects

Array Members: Minimum number of 1 item. Maximum number of 2 items.

Required: Yes

DestinationS3Uri
The Amazon S3 location used to capture the data.

Type: String

Length Constraints: Maximum length of 512.

Pattern: ^(https|s3)://([^/]+)?([^/]+)$

Required: Yes

EnableCapture
Whether data capture should be enabled or disabled (defaults to enabled).

Type: Boolean

Required: No

InitialSamplingPercentage
The percentage of requests SageMaker will capture. A lower value is recommended for Endpoints with high traffic.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: Yes

KmsKeyId
The Amazon Resource Name (ARN) of a AWS Key Management Service key that SageMaker uses to encrypt data on the storage volume attached to the ML compute instance that hosts the endpoint.

The KmsKeyId can be any of the following formats:
• Key ID: 1234abcd-12ab-34cd-56ef-1234567890ab
• Key ARN: arn:aws:kms:us-west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab
• Alias name: alias/ExampleAlias
• Alias name ARN: arn:aws:kms:us-west-2:111122223333:alias/ExampleAlias

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
DataCaptureConfigSummary
Service: Amazon SageMaker Service
The currently active data capture configuration used by your Endpoint.

Contents

CaptureStatus
Whether data capture is currently functional.
Type: String
Valid Values: Started | Stopped
Required: Yes

CurrentSamplingPercentage
The percentage of requests being captured by your Endpoint.
Type: Integer
Valid Range: Minimum value of 0. Maximum value of 100.
Required: Yes

DestinationS3Uri
The Amazon S3 location being used to capture the data.
Type: String
Length Constraints: Maximum length of 512.
Pattern: ^(https|s3):/([^/]+/[^/]+)$
Required: Yes

EnableCapture
Whether data capture is enabled or disabled.
Type: Boolean
Required: Yes

KmsKeyId
The KMS key being used to encrypt the data in Amazon S3.
Type: String
Length Constraints: Maximum length of 2048.
Pattern: .*
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
DataCatalogConfig
Service: Amazon SageMaker Service

The meta data of the Glue table which serves as data catalog for the OfflineStore.

Contents

Catalog
The name of the Glue table catalog.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 255.
Pattern: [ -퟿-�𐀀-􏿿	]*
Required: Yes

Database
The name of the Glue table database.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 255.
Pattern: [ -퟿-�𐀀-􏿿	]*
Required: Yes

TableName
The name of the Glue table.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 255.
Pattern: [ -퟿-�𐀀-􏿿	]*
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Data Processing

Service: Amazon SageMaker Service

The data structure used to specify the data to be used for inference in a batch transform job and to associate the data that is relevant to the prediction results in the output. The input filter provided allows you to exclude input data that is not needed for inference in a batch transform job. The output filter provided allows you to include input data relevant to interpreting the predictions in the output from the job. For more information, see Associate Prediction Results with their Corresponding Input Records.

Contents

Input Filter

A JSONPath expression used to select a portion of the input data to pass to the algorithm. Use the InputFilter parameter to exclude fields, such as an ID column, from the input. If you want SageMaker to pass the entire input dataset to the algorithm, accept the default value $.

Examples: "$", "$[1:]", "$.features"

Type: String

Length Constraints: Minimum length of 0. Maximum length of 63.

Required: No

Join Source

Specifies the source of the data to join with the transformed data. The valid values are None and Input. The default value is None, which specifies not to join the input with the transformed data. If you want the batch transform job to join the original input data with the transformed data, set JoinSource to Input. You can specify OutputFilter as an additional filter to select a portion of the joined dataset and store it in the output file.

For JSON or JSONLines objects, such as a JSON array, SageMaker adds the transformed data to the input JSON object in an attribute called SageMakerOutput. The joined result for JSON must be a key-value pair object. If the input is not a key-value pair object, SageMaker creates a new JSON file. In the new JSON file, and the input data is stored under the SageMakerInput key and the results are stored in SageMakerOutput.

For CSV data, SageMaker takes each row as a JSON array and joins the transformed data with the input by appending each transformed row to the end of the input. The joined data has the original input data followed by the transformed data and the output is a CSV file.

For information on how joining in applied, see Workflow for Associating Inferences with Input Records.

Type: String

Valid Values: Input | None

Required: No

Output Filter

A JSONPath expression used to select a portion of the joined dataset to save in the output file for a batch transform job. If you want SageMaker to store the entire input dataset in the output file, leave the default value, $. If you specify indexes that aren't within the dimension size of the joined dataset, you get an error.

Examples: "$", "$[0,5:]", "$['id','SageMakerOutput']"
Type: String
Length Constraints: Minimum length of 0. Maximum length of 63.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DataQualityAppSpecification

Service: Amazon SageMaker Service

Information about the container that a data quality monitoring job runs.

Contents

ContainerArguments

The arguments to send to the container that the monitoring job runs.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 50 items.

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

ContainerEntrypoint

The entrypoint for a container used to run a monitoring job.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

Environment

Sets the environment variables in the container that the monitoring job runs.

Type: String to string map

Map Entries: Maximum number of 50 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: [a-zA-Z_] [a-zA-Z0-9_]*

Value Length Constraints: Maximum length of 256.

Value Pattern: [\S\s]*

Required: No

ImageUri

The container image that the data quality monitoring job runs.

Type: String

Length Constraints: Maximum length of 255.

Pattern: .*
**PostAnalyticsProcessorSourceUri**

An Amazon S3 URI to a script that is called after analysis has been performed. Applicable only for the built-in (first party) containers.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(https|s3)://([/]+)/?(.*)$`

Required: No

**RecordPreprocessorSourceUri**

An Amazon S3 URI to a script that is called per row prior to running analysis. It can base64 decode the payload and convert it into a flattened json so that the built-in container can use the converted data. Applicable only for the built-in (first party) containers.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(https|s3)://([/]+)/?(.*)$`

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DataQualityBaselineConfig
Service: Amazon SageMaker Service

Configuration for monitoring constraints and monitoring statistics. These baseline resources are compared against the results of the current job from the series of jobs scheduled to collect data periodically.

Contents

BaseliningJobName

The name of the job that performs baselining for the data quality monitoring job.

Type: String


Pattern: ^[a-zA-Z0-9][-*[a-zA-Z0-9]]{0,62}

Required: No

ConstraintsResource

The constraints resource for a monitoring job.

Type: MonitoringConstraintsResource (p. 1198) object

Required: No

StatisticsResource

The statistics resource for a monitoring job.

Type: MonitoringStatisticsResource (p. 1218) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DataQualityJobInput
Service: Amazon SageMaker Service

The input for the data quality monitoring job. Currently endpoints are supported for input.

Contents

EndpointInput

  Input object for the endpoint

  Type: EndpointInput (p. 1017) object

  Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**DatasetDefinition**

Service: Amazon SageMaker Service

Configuration for Dataset Definition inputs. The Dataset Definition input must specify exactly one of either AthenaDatasetDefinition or RedshiftDatasetDefinition types.

**Contents**

**AthenaDatasetDefinition**

Configuration for Athena Dataset Definition input.

Type: AthenaDatasetDefinition (p. 902) object

Required: No

**DataDistributionType**

Whether the generated dataset is FullyReplicated or ShardedByS3Key (default).

Type: String

Valid Values: FullyReplicated | ShardedByS3Key

Required: No

**InputMode**

Whether to use File or Pipe input mode. In File (default) mode, Amazon SageMaker copies the data from the input source onto the local Amazon Elastic Block Store (Amazon EBS) volumes before starting your training algorithm. This is the most commonly used input mode. In Pipe mode, Amazon SageMaker streams input data from the source directly to your algorithm without using the EBS volume.

Type: String

Valid Values: Pipe | File

Required: No

**LocalPath**

The local path where you want Amazon SageMaker to download the Dataset Definition inputs to run a processing job. LocalPath is an absolute path to the input data. This is a required parameter when AppManaged is False (default).

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

**RedshiftDatasetDefinition**

Configuration for Redshift Dataset Definition input.

Type: RedshiftDatasetDefinition (p. 1331) object

Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DataSource
Service: Amazon SageMaker Service

Describes the location of the channel data.

Contents

FileSystemDataSource

The file system that is associated with a channel.

Type: FileSystemDataSource (p. 1044) object

Required: No

S3DataSource

The S3 location of the data source that is associated with a channel.

Type: S3DataSource (p. 1349) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DebugHookConfig

Service: Amazon SageMaker Service

Configuration information for the Debugger hook parameters, metric and tensor collections, and storage paths. To learn more about how to configure the `DebugHookConfig` parameter, see Use the SageMaker and Debugger Configuration API Operations to Create, Update, and Debug Your Training Job.

Contents

CollectionConfigurations

Configuration information for Debugger tensor collections. To learn more about how to configure the `CollectionConfiguration` parameter, see Use the SageMaker and Debugger Configuration API Operations to Create, Update, and Debug Your Training Job.

Type: Array of `CollectionConfiguration` (p. 952) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

Required: No

HookParameters

Configuration information for the Debugger hook parameters.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 20 items.

Key Length Constraints: Minimum length of 1. Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 256.

Value Pattern: .*

Required: No

LocalPath

Path to local storage location for metrics and tensors. Defaults to `/opt/ml/output/tensors/`.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: .*

Required: No

S3OutputPath

Path to Amazon S3 storage location for metrics and tensors.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3):/(/[^/]+)/?(.*)$

Required: Yes
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DebugRuleConfiguration

Service: Amazon SageMaker Service

Configuration information for SageMaker Debugger rules for debugging. To learn more about how to configure the DebugRuleConfiguration parameter, see Use the SageMaker and Debugger Configuration API Operations to Create, Update, and Debug Your Training Job.

Contents

InstanceType

The instance type to deploy a Debugger custom rule for debugging a training job.

Type: String

Valid Values: ml.t3.medium | ml.t3.large | ml.t3.xlarge | ml.t3.2xlarge
| ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge
| ml.m4.16xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge
| ml.m4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge
| ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.c5.xlarge
| ml.m5.large | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge
| ml.m5.24xlarge | ml.r5.large | ml.r5.xlarge | ml.r5.2xlarge | ml.r5.4xlarge
| ml.r5.8xlarge | ml.r5.12xlarge | ml.r5.16xlarge | ml.r5.24xlarge
| ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge
| ml.g4dn.12xlarge | ml.g4dn.16xlarge

Required: No

LocalPath

Path to local storage location for output of rules. Defaults to /opt/ml/processing/output/rule/.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: .*

Required: No

RuleConfigurationName

The name of the rule configuration. It must be unique relative to other rule configuration names.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: .*

Required: Yes

RuleEvaluatorImage

The Amazon Elastic Container (ECR) Image for the managed rule evaluation.

Type: String

Length Constraints: Maximum length of 255.
Pattern: .*  
Required: Yes

**RuleParameters**

Runtime configuration for rule container.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Key Length Constraints: Minimum length of 1. Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 256.

Value Pattern: .*

Required: No

**S3OutputPath**

Path to Amazon S3 storage location for rules.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([\^/]+)/?([^/]*)$

Required: No

**VolumeSizeInGB**

The size, in GB, of the ML storage volume attached to the processing instance.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DebugRuleEvaluationStatus
Service: Amazon SageMaker Service
Information about the status of the rule evaluation.

Contents

LastModifiedTime
Timestamp when the rule evaluation status was last modified.
Type: Timestamp
Required: No

RuleConfigurationName
The name of the rule configuration.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: .*
Required: No

RuleEvaluationJobArn
The Amazon Resource Name (ARN) of the rule evaluation job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:processing-job/.*
Required: No

RuleEvaluationStatus
Status of the rule evaluation.
Type: String
Valid Values: InProgress | NoIssuesFound | IssuesFound | Error | Stopping | Stopped
Required: No

StatusDetails
Details from the rule evaluation.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: .*
Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DeployedImage

Service: Amazon SageMaker Service

Gets the Amazon EC2 Container Registry path of the docker image of the model that is hosted in this ProductionVariant (p. 1293).

If you used the `registry/repository[:tag]` form to specify the image path of the primary container when you created the model hosted in this ProductionVariant, the path resolves to a path of the form `registry/repository[@digest]`. A digest is a hash value that identifies a specific version of an image. For information about Amazon ECR paths, see Pulling an Image in the Amazon ECR User Guide.

Contents

ResolutionTime

The date and time when the image path for the model resolved to the ResolvedImage

Type: Timestamp

Required: No

ResolvedImage

The specific digest path of the image hosted in this ProductionVariant.

Type: String

Length Constraints: Maximum length of 255.

Pattern: [\S]+

Required: No

SpecifiedImage

The image path you specified when you created the model.

Type: String

Length Constraints: Maximum length of 255.

Pattern: [\S]+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DeploymentConfig

Service: Amazon SageMaker Service

The deployment configuration for an endpoint, which contains the desired deployment strategy and rollback configurations.

Contents

AutoRollbackConfiguration

Automatic rollback configuration for handling endpoint deployment failures and recovery.

Type: AutoRollbackConfig (p. 925) object

Required: No

BlueGreenUpdatePolicy

Update policy for a blue/green deployment. If this update policy is specified, SageMaker creates a new fleet during the deployment while maintaining the old fleet. SageMaker flips traffic to the new fleet according to the specified traffic routing configuration. Only one update policy should be used in the deployment configuration. If no update policy is specified, SageMaker uses a blue/green deployment strategy with all at once traffic shifting by default.

Type: BlueGreenUpdatePolicy (p. 930) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DesiredWeightAndCapacity
Service: Amazon SageMaker Service
Specifies weight and capacity values for a production variant.

Contents

DesiredInstanceCount
The variant's capacity.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

DesiredWeight
The variant's weight.
Type: Float
Valid Range: Minimum value of 0.
Required: No

VariantName
The name of the variant to update.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Device
Service: Amazon SageMaker Service
Information of a particular device.

Contents

Description
Description of the device.
Type: String
Pattern: [\S\s]+
Required: No

DeviceName
The name of the device.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$
Required: Yes

IotThingName
AWS Internet of Things (IoT) object name.
Type: String
Length Constraints: Maximum length of 128.
Pattern: [a-zA-Z0-9:_.-]+
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DeviceFleetSummary
Service: Amazon SageMaker Service
Summary of the device fleet.

Contents

CreationTime
Timestamp of when the device fleet was created.
Type: Timestamp
Required: No

DeviceFleetArn
Amazon Resource Name (ARN) of the device fleet.
Type: String
Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:device-fleet/?[a-zA-Z0-9=\+,@\-_/.]+$
Required: Yes

DeviceFleetName
Name of the device fleet.
Type: String
Pattern: ^[a-zA-Z0-9\-](\*\[a-zA-Z0-9\])\{0,62\}$
Required: Yes

LastModifiedTime
Timestamp of when the device fleet was last updated.
Type: Timestamp
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DeviceStats
Service: Amazon SageMaker Service
Status of devices.

Contents

ConnectedDeviceCount
The number of devices connected with a heartbeat.
Type: Long
Required: Yes

RegisteredDeviceCount
The number of registered devices.
Type: Long
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DeviceSummary

Service: Amazon SageMaker Service

Summary of the device.

Contents

AgentVersion

Edge Manager agent version.

Type: String


Pattern: [a-zA-Z0-9\ \._\-]+

Required: No

Description

A description of the device.

Type: String


Pattern: [\S\s]+

Required: No

DeviceArn

Amazon Resource Name (ARN) of the device.

Type: String


Pattern: ^arn:aws[a-zA-Z\-]*:[a-zA-Z\-]*:[a-zA-Z\-]*:[d(12):[a-zA-Z\-]*]?[a-zA-Z0-9_0-9+=,.@\-_/]+$\n
Required: Yes

DeviceFleetName

The name of the fleet the device belongs to.

Type: String


Pattern: ^[a-zA-Z0-9-](?=.*[a-zA-Z0-9])\{0,62}$

Required: No

DeviceName

The unique identifier of the device.

Type: String

DeviceSummary

Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9\-]\{0,62\}$
Required: Yes

IoTThingName
The AWS Internet of Things (IoT) object thing name associated with the device.
Type: String
Length Constraints: Maximum length of 128.
Pattern: [a-zA-Z0-9_:\-]+
Required: No

LatestHeartbeat
The last heartbeat received from the device.
Type: Timestamp
Required: No

Models
Models on the device.
Type: Array of EdgeModelSummary (p. 1007) objects
Required: No

RegistrationTime
The timestamp of the last registration or de-reregistration.
Type: Timestamp
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DomainDetails
Service: Amazon SageMaker Service

The domain's details.

Contents

CreationTime
The creation time.
Type: Timestamp
Required: No

DomainArn
The domain's Amazon Resource Name (ARN).
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:domain/.*
Required: No

DomainId
The domain ID.
Type: String
Length Constraints: Maximum length of 63.
Required: No

DomainName
The domain name.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9-\-]{0,62}$
Required: No

LastModifiedTime
The last modified time.
Type: Timestamp
Required: No

Status
The status.
Type: String
Valid Values: Deleting | Failed | InService | Pending | Updating | Update_Failed | Delete_Failed
Required: No

**Url**

The domain's URL.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DomainSettings

Service: Amazon SageMaker Service

A collection of settings that apply to the SageMaker Domain. These settings are specified through the CreateDomain API call.

Contents

RStudioServerProDomainSettings

A collection of settings that configure the RStudioServerPro Domain-level app.

Type: RStudioServerProDomainSettings (p. 1347) object

Required: No

SecurityGroupIds

The security groups for the Amazon Virtual Private Cloud that the Domain uses for communication between Domain-level apps and user apps.

Type: Array of strings

Array Members: Maximum number of 3 items.

Length Constraints: Maximum length of 32.

Pattern: [-0-9a-zA-Z]+ 

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DomainSettingsForUpdate

Service: Amazon SageMaker Service

A collection of domain configuration settings to update.

Contents

RStudioServerProDomainSettingsForUpdate

A collection of RStudioServerPro domain-level app settings to update.

Type: RStudioServerProDomainSettingsForUpdate (p. 1348) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DriftCheckBaselines

Service: Amazon SageMaker Service

 Represents the drift check baselines that can be used when the model monitor is set using the model package.

Contents

Bias

 Represents the drift check bias baselines that can be used when the model monitor is set using the model package.

 Type: DriftCheckBias (p. 999) object

 Required: No

Explainability

 Represents the drift check explainability baselines that can be used when the model monitor is set using the model package.

 Type: DriftCheckExplainability (p. 1000) object

 Required: No

ModelDataQuality

 Represents the drift check model data quality baselines that can be used when the model monitor is set using the model package.

 Type: DriftCheckModelDataQuality (p. 1001) object

 Required: No

ModelQuality

 Represents the drift check model quality baselines that can be used when the model monitor is set using the model package.

 Type: DriftCheckModelQuality (p. 1002) object

 Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DriftCheckBias

Service: Amazon SageMaker Service

Represents the drift check bias baselines that can be used when the model monitor is set using the model package.

Contents

**ConfigFile**

The bias config file for a model.

Type: `FileSource (p. 1042)` object

Required: No

**PostTrainingConstraints**

The post-training constraints.

Type: `MetricsSource (p. 1148)` object

Required: No

**PreTrainingConstraints**

The pre-training constraints.

Type: `MetricsSource (p. 1148)` object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DriftCheckExplainability

Service: Amazon SageMaker Service

Represents the drift check explainability baselines that can be used when the model monitor is set using the model package.

Contents

ConfigFile

The explainability config file for the model.

Type: FileSource (p. 1042) object

Required: No

Constraints

The drift check explainability constraints.

Type: MetricsSource (p. 1148) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DriftCheckModelDataQuality
Service: Amazon SageMaker Service

Represents the drift check data quality baselines that can be used when the model monitor is set using the model package.

Contents

Constraints

The drift check model data quality constraints.

Type: MetricsSource (p. 1148) object

Required: No

Statistics

The drift check model data quality statistics.

Type: MetricsSource (p. 1148) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
DriftCheckModelQuality

Service: Amazon SageMaker Service

Represents the drift check model quality baselines that can be used when the model monitor is set using the model package.

Contents

Constraints

The drift check model quality constraints.

Type: MetricsSource (p. 1148) object

Required: No

Statistics

The drift check model quality statistics.

Type: MetricsSource (p. 1148) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Edge

Service: Amazon SageMaker Service

A directed edge connecting two lineage entities.

Contents

AssociationType

The type of the Association(Edge) between the source and destination. For example ContributedTo, Produced, or DerivedFrom.

Type: String

Valid Values: ContributedTo | AssociatedWith | DerivedFrom | Produced

Required: No

DestinationArn

The Amazon Resource Name (ARN) of the destination lineage entity of the directed edge.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

Required: No

SourceArn

The Amazon Resource Name (ARN) of the source lineage entity of the directed edge.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
### EdgeModel

**Service:** Amazon SageMaker Service

The model on the edge device.

#### Contents

**LatestInference**

The timestamp of the last inference that was made.

Type: Timestamp

Required: No

**LatestSampleTime**

The timestamp of the last data sample taken.

Type: Timestamp

Required: No

**ModelName**

The name of the model.

Type: String


Pattern: `^[a-zA-Z0-9\-_\.*]{0,62}$`

Required: Yes

**ModelVersion**

The model version.

Type: String


Pattern: `^[a-zA-Z0-9\-_\.]+$`

Required: Yes

#### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EdgeModelStat
Service: Amazon SageMaker Service
Status of edge devices with this model.

Contents

ActiveDeviceCount
The number of devices that have this model version, a heart beat, and are currently running.
Type: Long
Required: Yes

ConnectedDeviceCount
The number of devices that have this model version and have a heart beat.
Type: Long
Required: Yes

ModelName
The name of the model.
Type: String
Pattern: ^[a-zA-Z0-9\-_\.]\{0,62}$
Required: Yes

ModelVersion
The model version.
Type: String
Pattern: [a-zA-Z0-9\-_\.]\+
Required: Yes

OfflineDeviceCount
The number of devices that have this model version and do not have a heart beat.
Type: Long
Required: Yes

SamplingDeviceCount
The number of devices with this model version and are producing sample data.
Type: Long
Required: Yes
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EdgeModelSummary
Service: Amazon SageMaker Service
Summary of model on edge device.

Contents

ModelName
The name of the model.
Type: String
Pattern: ^[a-zA-Z0-9\-\*_\_.]*[a-zA-Z0-9\-\*_\_.]{0,62}$
Required: Yes

ModelVersion
The version model.
Type: String
Pattern: [a-zA-Z0-9\-_\.]+$
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EdgeOutputConfig

Service: Amazon SageMaker Service

The output configuration.

Contents

KmsKeyId

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt data on the storage volume after compilation job. If you don't provide a KMS key ID, Amazon SageMaker uses the default KMS key for Amazon S3 for your role's account.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

PresetDeploymentConfig

The configuration used to create deployment artifacts. Specify configuration options with a JSON string. The available configuration options for each type are:

- ComponentName (optional) - Name of the GreenGrass V2 component. If not specified, the default name generated consists of "SagemakerEdgeManager" and the name of your SageMaker Edge Manager packaging job.
- ComponentDescription (optional) - Description of the component.
- ComponentVersion (optional) - The version of the component.

Note

AWS IoT Greengrass uses semantic versions for components. Semantic versions follow a major.minor.patch number system. For example, version 1.0.0 represents the first major release for a component. For more information, see the semantic version specification.

- PlatformOS (optional) - The name of the operating system for the platform. Supported platforms include Windows and Linux.
- PlatformArchitecture (optional) - The processor architecture for the platform.

Supported architectures Windows include: Windows32_x86, Windows64_x64.

Supported architectures for Linux include: Linux x86_64, Linux ARMV8.

Type: String

Required: No

PresetDeploymentType

The deployment type SageMaker Edge Manager will create. Currently only supports AWS IoT Greengrass Version 2 components.

Type: String

Valid Values: GreengrassV2Component

Required: No

S3OutputLocation

The Amazon Simple Storage (S3) bucket URI.
EdgeOutputConfig

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)/?([^/]*)$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EdgePackagingJobSummary

Service: Amazon SageMaker Service

Summary of edge packaging job.

Contents

CompilationJobName

The name of the SageMaker Neo compilation job.

Type: String


Pattern: ^[a-zA-Z0-9\-\]*[a-zA-Z0-9]\{0,62}$

Required: No

CreationTime

The timestamp of when the job was created.

Type: Timestamp

Required: No

EdgePackagingJobArn

The Amazon Resource Name (ARN) of the edge packaging job.

Type: String


Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z\-]*:\d{12}:edge-packaging-job/?[a-zA-Z0-9\-_+=,.@\-_/]+$

Required: Yes

EdgePackagingJobName

The name of the edge packaging job.

Type: String


Pattern: ^[a-zA-Z0-9\-\]*[a-zA-Z0-9]\{0,62}$

Required: Yes

EdgePackagingJobStatus

The status of the edge packaging job.

Type: String

Valid Values: STARTING | INPROGRESS | COMPLETED | FAILED | STOPPING | STOPPED

Required: Yes

LastModifiedTime

The timestamp of when the edge packaging job was last updated.
Type: Timestamp
Required: No

**ModelName**

The name of the model.

Type: String


Pattern: ^[a-zA-Z0-9-]*[a-zA-Z0-9]\(0,62\)$

Required: No

**ModelVersion**

The version of the model.

Type: String


Pattern: [a-zA-Z0-9\_\.]+$

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EdgePresetDeploymentOutput
Service: Amazon SageMaker Service
The output of a SageMaker Edge Manager deployable resource.

Contents

Artifact
The Amazon Resource Name (ARN) of the generated deployable resource.
Type: String
Required: No

Status
The status of the deployable resource.
Type: String
Valid Values: COMPLETED | FAILED
Required: No

StatusMessage
Returns a message describing the status of the deployed resource.
Type: String
Required: No

Type
The deployment type created by SageMaker Edge Manager. Currently only supports AWS IoT Greengrass Version 2 components.
Type: String
Valid Values: GreengrassV2Component
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EMRStepMetadata

Service: Amazon SageMaker Service

The configurations and outcomes of an Amazon EMR step execution.

Contents

ClusterId

The identifier of the EMR cluster.

Type: String

Length Constraints: Maximum length of 256.

Required: No

LogFilePath

The path to the log file where the cluster step's failure root cause is recorded.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

StepId

The identifier of the EMR cluster step.

Type: String

Length Constraints: Maximum length of 256.

Required: No

StepName

The name of the EMR cluster step.

Type: String

Length Constraints: Maximum length of 256.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Endpoint
Service: Amazon SageMaker Service
A hosted endpoint for real-time inference.

Contents

CreationTime
The time that the endpoint was created.
Type: Timestamp
Required: Yes

DataCaptureConfig
The currently active data capture configuration used by your Endpoint.
Type: DataCaptureConfigSummary (p. 968) object
Required: No

EndpointArn
The Amazon Resource Name (ARN) of the endpoint.
Type: String
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:endpoint/.*
Required: Yes

EndpointConfigName
The endpoint configuration associated with the endpoint.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](\*[a-zA-Z0-9])*\{0,62\}
Required: Yes

EndpointName
The name of the endpoint.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](\*[a-zA-Z0-9])*\{0,62\}
Required: Yes

EndpointStatus
The status of the endpoint.
Type: String
Valid Values: OutOfService | Creating | Updating | SystemUpdating | RollingBack | InService | Deleting | Failed

Required: Yes

**FailureReason**

If the endpoint failed, the reason it failed.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

**LastModifiedTime**

The last time the endpoint was modified.

Type: Timestamp

Required: Yes

**MonitoringSchedules**

A list of monitoring schedules for the endpoint. For information about model monitoring, see Amazon SageMaker Model Monitor.

Type: Array of MonitoringSchedule (p. 1212) objects

Required: No

**ProductionVariants**

A list of the production variants hosted on the endpoint. Each production variant is a model.

Type: Array of ProductionVariantSummary (p. 1299) objects

Array Members: Minimum number of 1 item.

Required: No

**Tags**

A list of the tags associated with the endpoint. For more information, see Tagging AWS resources in the AWS General Reference Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EndpointConfigSummary

Service: Amazon SageMaker Service

Provides summary information for an endpoint configuration.

Contents

CreationTime

A timestamp that shows when the endpoint configuration was created.

Type: Timestamp

Required: Yes

EndpointConfigArn

The Amazon Resource Name (ARN) of the endpoint configuration.

Type: String


Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:endpoint-config/.*

Required: Yes

EndpointConfigName

The name of the endpoint configuration.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9-\-]*[a-zA-Z0-9]{0,62}

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EndpointInput

Service: Amazon SageMaker Service

Input object for the endpoint

Contents

**EndpointName**

An endpoint in customer's account which has enabled DataCaptureConfig enabled.

- Type: String
- Length Constraints: Maximum length of 63.
- Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\]`
- Required: Yes

**EndTimeOffset**

If specified, monitoring jobs subtract this time from the end time. For information about using offsets for scheduling monitoring jobs, see Schedule Model Quality Monitoring Jobs.

- Type: String
- Pattern: `^\ ?P\ .*`
- Required: No

**FeaturesAttribute**

The attributes of the input data that are the input features.

- Type: String
- Required: No

**InferenceAttribute**

The attribute of the input data that represents the ground truth label.

- Type: String
- Required: No

**LocalPath**

Path to the filesystem where the endpoint data is available to the container.

- Type: String
- Length Constraints: Maximum length of 256.
- Pattern: `.*`
- Required: Yes

**ProbabilityAttribute**

In a classification problem, the attribute that represents the class probability.
**EndpointInput**

- **Type**: String
  - **Required**: No

**ProbabilityThresholdAttribute**

The threshold for the class probability to be evaluated as a positive result.

- **Type**: Double
  - **Required**: No

**S3DataDistributionType**

Whether input data distributed in Amazon S3 is fully replicated or sharded by an S3 key. Defaults to **FullyReplicated**

- **Type**: String
  - **Valid Values**: FullyReplicated | ShardedByS3Key
  - **Required**: No

**S3InputMode**

Whether the Pipe or File is used as the input mode for transferring data for the monitoring job. Pipe mode is recommended for large datasets. File mode is useful for small files that fit in memory. Defaults to File.

- **Type**: String
  - **Valid Values**: Pipe | File
  - **Required**: No

**StartTimeOffset**

If specified, monitoring jobs subtract this time from the start time. For information about using offsets for scheduling monitoring jobs, see [Schedule Model Quality Monitoring Jobs](#).

- **Type**: String
  - **Length Constraints**: Minimum length of 1. Maximum length of 15.
  - **Pattern**: ^\.?P.*
  - **Required**: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EndpointInputConfiguration

Service: Amazon SageMaker Service

The endpoint configuration for the load test.

Contents

EnvironmentParameterRanges

The parameter you want to benchmark against.

Type: EnvironmentParameterRanges (p. 1026) object

Required: No

InferenceSpecificationName

The inference specification name in the model package version.

Type: String


Pattern: \^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62\}$

Required: No

InstanceType

The instance types to use for the load test.

Type: String

Valid Values: ml.t2.medium | ml.t2.large | ml.t2.xlarge | ml.t2.2xlarge |
ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge |
ml.m4.16xlarge | ml.m5.large | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge |
ml.m5.12xlarge | ml.m5.24xlarge | ml.m5d.large | ml.m5d.xlarge |
ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.12xlarge | ml.m5d.24xlarge | ml.c4.large | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge |
ml.c4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge | ml.p3.2xlarge |
ml.p3.8xlarge | ml.p3.16xlarge | ml.c5.large | ml.c5.xlarge | ml.c5.2xlarge |
ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge | ml.c5d.large |
ml.c5d.xlarge | ml.c5d.2xlarge | ml.c5d.4xlarge | ml.c5d.9xlarge |
ml.c5d.18xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge |
ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.r5.large |
ml.r5.xlarge | ml.r5.2xlarge | ml.r5.4xlarge | ml.r5.12xlarge |
ml.r5.24xlarge | ml.r5d.large | ml.r5d.xlarge | ml.r5d.2xlarge |
ml.r5d.4xlarge | ml.r5d.12xlarge | ml.r5d.24xlarge | ml.inf1.xlarge |
ml.inf1.2xlarge | ml.inf1.6xlarge | ml.inf1.24xlarge

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EndpointOutputConfiguration

Service: Amazon SageMaker Service

The endpoint configuration made by Inference Recommender during a recommendation job.

Contents

**EndpointName**

The name of the endpoint made during a recommendation job.

Type: String

Required: Yes

**InitialInstanceCount**

The number of instances recommended to launch initially.

Type: Integer

Required: Yes

**InstanceType**

The instance type recommended by Amazon SageMaker Inference Recommender.

Type: String

Valid Values:

- ml.t2.medium
- ml.t2.large
- ml.t2.xlarge
- ml.t2.2xlarge
- ml.m4.xlarge
- ml.m4.2xlarge
- ml.m4.4xlarge
- ml.m4.10xlarge
- ml.m4.16xlarge
- ml.m5.large
- ml.m5.xlarge
- ml.m5.2xlarge
- ml.m5.4xlarge
- ml.m5.12xlarge
- ml.m5.24xlarge
- ml.m5d.large
- ml.m5d.xlarge
- ml.m5d.2xlarge
- ml.m5d.4xlarge
- ml.m5d.10xlarge
- ml.m5d.12xlarge
- ml.m5d.24xlarge
- ml.c4.large
- ml.c4.xlarge
- ml.c4.2xlarge
- ml.c4.4xlarge
- ml.c4.8xlarge
- ml.p2.xlarge
- ml.p2.8xlarge
- ml.p2.16xlarge
- ml.p3.2xlarge
- ml.p3.8xlarge
- ml.p3.16xlarge
- ml.c5.large
- ml.c5.xlarge
- ml.c5.2xlarge
- ml.c5.4xlarge
- ml.c5.9xlarge
- ml.c5.18xlarge
- ml.c5d.large
- ml.c5d.xlarge
- ml.c5d.2xlarge
- ml.c5d.4xlarge
- ml.c5d.9xlarge
- ml.c5d.18xlarge
- ml.g4dn.xlarge
- ml.g4dn.2xlarge
- ml.g4dn.4xlarge
- ml.g4dn.8xlarge
- ml.g4dn.12xlarge
- ml.g4dn.16xlarge
- ml.r5.large
- ml.r5.xlarge
- ml.r5.2xlarge
- ml.r5.4xlarge
- ml.r5.12xlarge
- ml.r5.24xlarge
- ml.r5d.large
- ml.r5d.xlarge
- ml.r5d.2xlarge
- ml.r5d.4xlarge
- ml.r5d.12xlarge
- ml.r5d.24xlarge
- ml.inf1.xlarge
- ml.inf1.2xlarge
- ml.inf1.6xlarge
- ml.inf1.24xlarge

Required: Yes

**VariantName**

The name of the production variant (deployed model) made during a recommendation job.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EndpointSummary

Service: Amazon SageMaker Service

Provides summary information for an endpoint.

Contents

CreationTime

A timestamp that shows when the endpoint was created.

Type: Timestamp

Required: Yes

EndpointArn

The Amazon Resource Name (ARN) of the endpoint.

Type: String


Pattern: arn:aws[a-z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:endpoint/.*

Required: Yes

EndpointName

The name of the endpoint.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}

Required: Yes

EndpointStatus

The status of the endpoint.

- **OutOfService**: Endpoint is not available to take incoming requests.
- **Creating**: CreateEndpoint (p. 69) is executing.
- **Updating**: UpdateEndpoint (p. 767) or UpdateEndpointWeightsAndCapacities (p. 771) is executing.
- **SystemUpdating**: Endpoint is undergoing maintenance and cannot be updated or deleted or re-scaled until it has completed. This maintenance operation does not change any customer-specified values such as VPC config, KMS encryption, model, instance type, or instance count.
- **RollingBack**: Endpoint fails to scale up or down or change its variant weight and is in the process of rolling back to its previous configuration. Once the rollback completes, endpoint returns to an **InService** status. This transitional status only applies to an endpoint that has autoscaling enabled and is undergoing variant weight or capacity changes as part of an UpdateEndpointWeightsAndCapacities (p. 771) call or when the UpdateEndpointWeightsAndCapacities (p. 771) operation is called explicitly.
- **InService**: Endpoint is available to process incoming requests.
- **Deleting**: DeleteEndpoint (p. 225) is executing.
- **Failed**: Endpoint could not be created, updated, or re-scaled. Use DescribeEndpoint:FailureReason (p. 341) for information about the failure. DeleteEndpoint (p. 225) is the only operation that can be performed on a failed endpoint.
EndpointSummary

To get a list of endpoints with a specified status, use the `ListEndpoints:StatusEquals (p. 551)` filter.

Type: String

Valid Values: OutOfService | Creating | Updating | SystemUpdating | RollingBack | InService | Deleting | Failed

Required: Yes

**LastModifiedTime**

A timestamp that shows when the endpoint was last modified.

Type: Timestamp

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EnvironmentParameter
Service: Amazon SageMaker Service

A list of environment parameters suggested by the Amazon SageMaker Inference Recommender.

Contents

**Key**
The environment key suggested by the Amazon SageMaker Inference Recommender.

Type: String
Required: Yes

**Value**
The value suggested by the Amazon SageMaker Inference Recommender.

Type: String
Required: Yes

**ValueType**
The value type suggested by the Amazon SageMaker Inference Recommender.

Type: String
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
EnvironmentParameterRanges
Service: Amazon SageMaker Service
Specifies the range of environment parameters

Contents

CategoricalParameterRanges
Specified a list of parameters for each category.
Type: Array of CategoricalParameter (p. 938) objects
Array Members: Minimum number of 1 item. Maximum number of 5 items.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Experiment
Service: Amazon SageMaker Service

The properties of an experiment as returned by the Search (p. 691) API.

Contents

CreatedBy
Who created the experiment.
Type: UserContext (p. 1449) object
Required: No

CreationTime
When the experiment was created.
Type: Timestamp
Required: No

Description
The description of the experiment.
Type: String
Length Constraints: Maximum length of 3072.
Pattern: .*
Required: No

DisplayName
The name of the experiment as displayed. If DisplayName isn't specified, ExperimentName is displayed.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9\-\-\ morally\-]*[a-zA-Z0-9\-\-\ morally\-]{0,119}$
Required: No

ExperimentArn
The Amazon Resource Name (ARN) of the experiment.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-zA-Z\-\-\ morally\-]*:sagemaker:[a-zA-Z0-9\-\-\ morally\-]:[0-9]{12}:experiment/.*
Required: No

ExperimentName
The name of the experiment.
Experiment

**Type**: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: No

**LastModifiedBy**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

Required: No

**LastModifiedTime**

When the experiment was last modified.

Type: Timestamp

Required: No

**Source**

The source of the experiment.

Type: ExperimentSource (p. 1031) object

Required: No

**Tags**

The list of tags that are associated with the experiment. You can use Search (p. 691) API to search on the tags.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ExperimentConfig

Service: Amazon SageMaker Service

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:

- CreateProcessingJob (p. 162)
- CreateTrainingJob (p. 173)
- CreateTransformJob (p. 182)

Contents

ExperimentName

The name of an existing experiment to associate the trial component with.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: No

TrialComponentDisplayName

The display name for the trial component. If this key isn't specified, the display name is the trial component name.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: No

TrialName

The name of an existing trial to associate the trial component with. If not specified, a new trial is created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for Ruby V3
ExperimentSource
Service: Amazon SageMaker Service

The source of the experiment.

Contents

SourceArn
The Amazon Resource Name (ARN) of the source.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}::*
Required: Yes

SourceType
The source type.
Type: String
Length Constraints: Maximum length of 128.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ExperimentSummary

Service: Amazon SageMaker Service

A summary of the properties of an experiment. To get the complete set of properties, call the DescribeExperiment (p. 346) API and provide the ExperimentName.

Contents

CreationTime

When the experiment was created.

Type: Timestamp

Required: No

DisplayName

The name of the experiment as displayed. If DisplayName isn't specified, ExperimentName is displayed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){$,119}

Required: No

ExperimentArn

The Amazon Resource Name (ARN) of the experiment.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment/.*

Required: No

ExperimentName

The name of the experiment.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){$,119}

Required: No

ExperimentSource

The source of the experiment.

Type: ExperimentSource (p. 1031) object

Required: No

LastModifiedTime

When the experiment was last modified.
ExperimentSummary

Type: Timestamp
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Explainability
Service: Amazon SageMaker Service
Contains explainability metrics for a model.

Contents

Report
The explainability report for a model.
Type: MetricsSource (p. 1148) object
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FailStepMetadata

Service: Amazon SageMaker Service

The container for the metadata for Fail step.

Contents

ErrorMessage

A message that you define and then is processed and rendered by the Fail step when the error occurs.

Type: String

Length Constraints: Maximum length of 3072.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FeatureDefinition
Service: Amazon SageMaker Service

A list of features. You must include FeatureName and FeatureType. Valid feature FeatureTypes are Integral, Fractional and String.

Contents

**FeatureName**

The name of a feature. The type must be a string. FeatureName cannot be any of the following: is_deleted, write_time, api_invocation_time.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9][-_]*[a-zA-Z0-9]{0,63}

Required: No

**FeatureType**

The value type of a feature. Valid values are Integral, Fractional, or String.

Type: String

Valid Values: Integral | Fractional | String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**FeatureGroup**

Service: Amazon SageMaker Service

Amazon SageMaker Feature Store stores features in a collection called Feature Group. A Feature Group can be visualized as a table which has rows, with a unique identifier for each row where each column in the table is a feature. In principle, a Feature Group is composed of features and values per features.

**Contents**

**CreationTime**

The time a FeatureGroup was created.

Type: Timestamp

Required: No

**Description**

A free form description of a FeatureGroup.

Type: String

Length Constraints: Maximum length of 128.

Required: No

**EventTimeFeatureName**

The name of the feature that stores the EventTime of a Record in a FeatureGroup.

A EventTime is point in time when a new event occurs that corresponds to the creation or update of a Record in FeatureGroup. All Records in the FeatureGroup must have a corresponding EventTime.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9][_-]*[a-zA-Z0-9]{0,63}$

Required: No

**FailureReason**

The reason that the FeatureGroup failed to be replicated in the OfflineStore. This is failure may be due to a failure to create a FeatureGroup in or delete a FeatureGroup from the OfflineStore.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

**FeatureDefinitions**

A list of Features. Each Feature must include a FeatureName and a FeatureType.

Valid FeatureTypes are Integral, Fractional and String.

FeatureNames cannot be any of the following: is_deleted, write_time, api_invocation_time.
You can create up to 2,500 FeatureDefinitions per FeatureGroup.

Type: Array of FeatureDefinition (p. 1036) objects

Array Members: Minimum number of 1 item. Maximum number of 2500 items.

Required: No

**FeatureGroupArn**

The Amazon Resource Name (ARN) of a FeatureGroup.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:feature-group/.*`

Required: No

**FeatureGroupName**

The name of the FeatureGroup.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^\[a-zA-Z0-9\](-*[a-zA-Z0-9]\{0,63\}`

Required: No

**FeatureGroupStatus**

A FeatureGroup status.

Type: String

Valid Values: Creating | Created | CreateFailed | Deleting | DeleteFailed

Required: No

**OfflineStoreConfig**

The configuration of an OfflineStore.

Provide an OfflineStoreConfig in a request to CreateFeatureGroup to create an OfflineStore.

To encrypt an OfflineStore using at rest data encryption, specify AWS Key Management Service (KMS) key ID, or KMSKeyId, in S3StorageConfig.

Type: OfflineStoreConfig (p. 1231) object

Required: No

**OfflineStoreStatus**

The status of OfflineStore.

Type: OfflineStoreStatus (p. 1232) object

Required: No
**OnlineStoreConfig**

Use this to specify the AWS Key Management Service (KMS) Key ID, or KMSKeyId, for at rest data encryption. You can turn OnlineStore on or off by specifying the EnableOnlineStore flag at General Assembly; the default value is False.

Type: OnlineStoreConfig (p. 1238) object

Required: No

**RecordIdentifierFeatureName**

The name of the Feature whose value uniquely identifies a Record defined in the FeatureGroup FeatureDefinitions.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: \^[a-zA-Z0-9]*[-_]*[a-zA-Z0-9]*\{0,63\}

Required: No

**RoleArn**

The Amazon Resource Name (ARN) of the IAM execution role used to create the feature group.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9+=,.@-_\//\]+$

Required: No

**Tags**

Tags used to define a FeatureGroup.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FeatureGroupSummary

Service: Amazon SageMaker Service

The name, Arn, CreationTime, FeatureGroup values, LastUpdatedTime and EnableOnlineStorage status of a FeatureGroup.

Contents

CreationTime

A timestamp indicating the time of creation time of the FeatureGroup.

Type: Timestamp

Required: Yes

FeatureGroupArn

Unique identifier for the FeatureGroup.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:feature-group/.*

Required: Yes

FeatureGroupName

The name of FeatureGroup.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^([a-zA-Z0-9\-]\d*[a-zA-Z0-9\-])\{0,63\}

Required: Yes

FeatureGroupStatus

The status of a FeatureGroup. The status can be any of the following: Creating, Created, CreateFail, Deleting or DeleteFail.

Type: String

Valid Values: Creating | Created | CreateFailed | Deleting | DeleteFailed

Required: No

OfflineStoreStatus

Notifies you if replicating data into the OfflineStore has failed. Returns either: Active or Blocked.

Type: OfflineStoreStatus (p. 1232) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
FileSource

Service: Amazon SageMaker Service

Contains details regarding the file source.

Contents

ContentDigest

The digest of the file source.

Type: String

Length Constraints: Maximum length of 72.

Pattern: ^[Ss][Hh][Aa]256:[0-9a-fA-F]{64}$

Required: No

ContentType

The type of content stored in the file source.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

S3Uri

The Amazon S3 URI for the file source.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://(/[^/]+)?(.*)$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FileSystemConfig
Service: Amazon SageMaker Service

The Amazon Elastic File System (EFS) storage configuration for a SageMaker image.

Contents

DefaultGid

The default POSIX group ID (GID). If not specified, defaults to 100.
Type: Integer
Valid Range: Minimum value of 0. Maximum value of 65535.
Required: No

DefaultUid

The default POSIX user ID (UID). If not specified, defaults to 1000.
Type: Integer
Valid Range: Minimum value of 0. Maximum value of 65535.
Required: No

MountPath

The path within the image to mount the user’s EFS home directory. The directory should be empty. If not specified, defaults to /home/sagemaker-user.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: ^\//.*
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FileSystemDataSource

Service: Amazon SageMaker Service

Specifies a file system data source for a channel.

Contents

DirectoryPath

The full path to the directory to associate with the channel.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: .*

Required: Yes

FileSystemAccessMode

The access mode of the mount of the directory associated with the channel. A directory can be mounted either in ro (read-only) or rw (read-write) mode.

Type: String

Valid Values: rw | ro

Required: Yes

FileSystemId

The file system id.

Type: String

Length Constraints: Minimum length of 11.

Pattern: .*

Required: Yes

FileSystemType

The file system type.

Type: String

Valid Values: EFS | FSxLustre

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Filter

Service: Amazon SageMaker Service

A conditional statement for a search expression that includes a resource property, a Boolean operator, and a value. Resources that match the statement are returned in the results from the Search (p. 691) API.

If you specify a Value, but not an Operator, Amazon SageMaker uses the equals operator.

In search, there are several property types:

Metrics

To define a metric filter, enter a value using the form "Metrics.<name>", where <name> is a metric name. For example, the following filter searches for training jobs with an "accuracy" metric greater than "0.9":

```json
{
    "Name": "Metrics.accuracy",
    "Operator": "GreaterThan",
    "Value": "0.9"
}
```

HyperParameters

To define a hyperparameter filter, enter a value with the form "HyperParameters.<name>". Decimal hyperparameter values are treated as a decimal in a comparison if the specified Value is also a decimal value. If the specified Value is an integer, the decimal hyperparameter values are treated as integers. For example, the following filter is satisfied by training jobs with a "learning_rate" hyperparameter that is less than "0.5":

```json
{
    "Name": "HyperParameters.learning_rate",
    "Operator": "LessThan",
    "Value": "0.5"
}
```

Tags

To define a tag filter, enter a value with the form Tags.<key>.

Contents

Name

A resource property name. For example, TrainingJobName. For valid property names, see SearchRecord (p. 1357). You must specify a valid property for the resource.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .+
Required: Yes

**Operator**

A Boolean binary operator that is used to evaluate the filter. The operator field contains one of the following values:

- **Equals**
  
  The value of Name equals Value.

- **NotEquals**
  
  The value of Name doesn't equal Value.

- **Exists**
  
  The Name property exists.

- **NotExists**
  
  The Name property does not exist.

- **GreaterThan**
  
  The value of Name is greater than Value. Not supported for text properties.

- **GreaterThanOrEqualTo**
  
  The value of Name is greater than or equal to Value. Not supported for text properties.

- **LessThan**
  
  The value of Name is less than Value. Not supported for text properties.

- **LessThanOrEqualTo**
  
  The value of Name is less than or equal to Value. Not supported for text properties.

- **In**
  
  The value of Name is one of the comma delimited strings in Value. Only supported for text properties.

- **Contains**
  
  The value of Name contains the string Value. Only supported for text properties.

A SearchExpression can include the Contains operator multiple times when the value of Name is one of the following:

- Experiment.DisplayName
- Experiment.ExperimentName
- Experiment.Tags
- Trial.DisplayName
- Trial.TrialName
- Trial.Tags
- TrialComponent.DisplayName
- TrialComponent.TrialComponentName
- TrialComponent.Tags
- TrialComponent.InputArtifacts
- TrialComponent.OutputArtifacts

A SearchExpression can include only one Contains operator for all other values of Name. In these cases, if you include multiple Contains operators in the SearchExpression, the result is the following error message: "'CONTAINS' operator usage limit of 1 exceeded."
Filter

**Type:** String

**Valid Values:**
- Equals
- NotEquals
- GreaterThan
- GreaterThanOrEqualTo
- LessThan
- LessThanOrEqualTo
- Contains
- Exists
- NotExists
- In

**Required:** No

**Value**

A value used with Name and Operator to determine which resources satisfy the filter's condition. For numerical properties, Value must be an integer or floating-point decimal. For timestamp properties, Value must be an ISO 8601 date-time string of the following format: `YYYY-mm-dd'T'HH:MM:SS`.

**Type:** String

**Length Constraints:** Minimum length of 1. Maximum length of 1024.

**Pattern:** .+

**Required:** No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FinalAutoMLJobObjectiveMetric
Service: Amazon SageMaker Service
The best candidate result from an AutoML training job.

Contents

MetricName
The name of the metric with the best result. For a description of the possible objective metrics, see AutoMLJobObjective:MetricName (p. 917).
Type: String
Valid Values: Accuracy | MSE | F1 | F1macro | AUC
Required: Yes

Type
The type of metric with the best result.
Type: String
Valid Values: Maximize | Minimize
Required: No

Value
The value of the metric with the best result.
Type: Float
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FinalHyperParameterTuningJobObjectiveMetric

Service: Amazon SageMaker Service

Shows the final value for the objective metric for a training job that was launched by a hyperparameter tuning job. You define the objective metric in the HyperParameterTuningJobObjective parameter of HyperParameterTuningJobConfig (p. 1089).

Contents

MetricName

The name of the objective metric.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .+

Required: Yes

Type

Whether to minimize or maximize the objective metric. Valid values are Minimize and Maximize.

Type: String

Valid Values: Maximize | Minimize

Required: No

Value

The value of the objective metric.

Type: Float

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FlowDefinitionOutputConfig
Service: Amazon SageMaker Service

Contains information about where human output will be stored.

Contents

KmsKeyId
The Amazon Key Management Service (KMS) key ID for server-side encryption.
Type: String
Length Constraints: Maximum length of 2048.
Pattern: . *
Required: No

S3OutputPath
The Amazon S3 path where the object containing human output will be made available.
To learn more about the format of Amazon A2I output data, see Amazon A2I Output Data.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: ^(https|s3):/(/[^/]+/)?(.*)$
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FlowDefinitionSummary
Service: Amazon SageMaker Service
Contains summary information about the flow definition.

Contents

CreationTime
The timestamp when SageMaker created the flow definition.
Type: Timestamp
Required: Yes

FailureReason
The reason why the flow definition creation failed. A failure reason is returned only when the flow
definition status is Failed.
Type: String
Length Constraints: Maximum length of 1024.
Required: No

FlowDefinitionArn
The Amazon Resource Name (ARN) of the flow definition.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:flow-definition/.*
Required: Yes

FlowDefinitionName
The name of the flow definition.
Type: String
Pattern: ^[a-z0-9](-*[a-z0-9])\{0,62\}
Required: Yes

FlowDefinitionStatus
The status of the flow definition. Valid values:
Type: String
Valid Values: Initializing | Active | Failed | Deleting
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
GitConfig
Service: Amazon SageMaker Service
Specifies configuration details for a Git repository in your AWS account.

Contents

Branch
The default branch for the Git repository.
Type: String
Pattern: [^ -^:?!\[ ]+
Required: No

RepositoryUrl
The URL where the Git repository is located.
Type: String
Pattern: ^https://([^/]+)/(.*)$
Required: Yes

SecretArn
The Amazon Resource Name (ARN) of the AWS Secrets Manager secret that contains the credentials used to access the git repository. The secret must have a staging label of AWSCURRENT and must be in the following format:

{"username": UserName, "password": Password}
Type: String
Length Constraints: Minimum length of 1. Maximum length of 2048.
Pattern: arn:aws[a-z-]*:secretsmanager:[a-z0-9-]*:[0-9]{12}:secret:.*
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
GitConfigForUpdate

Service: Amazon SageMaker Service

Specifies configuration details for a Git repository when the repository is updated.

Contents

SecretArn

The Amazon Resource Name (ARN) of the AWS Secrets Manager secret that contains the credentials used to access the git repository. The secret must have a staging label of AWSCURRENT and must be in the following format:

{"username": UserName, "password": Password}

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:secretsmanager:[a-z0-9\-]*:[0-9]{12}:secret:.*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HumanLoopActivationConditionsConfig

Service: Amazon SageMaker Service

Defines under what conditions SageMaker creates a human loop. Used within CreateFlowDefinition (p. 85). See HumanLoopActivationConditionsConfig (p. 1056) for the required format of activation conditions.

Contents

HumanLoopActivationConditions

JSON expressing use-case specific conditions declaratively. If any condition is matched, atomic tasks are created against the configured work team. The set of conditions is different for Rekognition and Textract. For more information about how to structure the JSON, see JSON Schema for Human Loop Activation Conditions in Amazon Augmented AI in the Amazon SageMaker Developer Guide.

Type: String

Length Constraints: Maximum length of 10240.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HumanLoopActivationConfig
Service: Amazon SageMaker Service

Provides information about how and under what conditions SageMaker creates a human loop. If HumanLoopActivationConfig is not given, then all requests go to humans.

Contents

HumanLoopActivationConditionsConfig

Container structure for defining under what conditions SageMaker creates a human loop.

Type: HumanLoopActivationConditionsConfig (p. 1056) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**HumanLoopConfig**

Service: Amazon SageMaker Service

Describes the work to be performed by human workers.

**Contents**

**HumanTaskUiArn**

The Amazon Resource Name (ARN) of the human task user interface.

You can use standard HTML and Crowd HTML Elements to create a custom worker task template. You use this template to create a human task UI.

To learn how to create a custom HTML template, see Create Custom Worker Task Template.

To learn how to create a human task UI, which is a worker task template that can be used in a flow definition, see Create and Delete a Worker Task Templates.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:human-task-ui/.*`

Required: Yes

**PublicWorkforceTaskPrice**

Defines the amount of money paid to an Amazon Mechanical Turk worker for each task performed.

Use one of the following prices for bounding box tasks. Prices are in US dollars and should be based on the complexity of the task; the longer it takes in your initial testing, the more you should offer.

- 0.036
- 0.048
- 0.060
- 0.072
- 0.120
- 0.240
- 0.360
- 0.480
- 0.600
- 0.720
- 0.840
- 0.960
- 1.080
- 1.200

Use one of the following prices for image classification, text classification, and custom tasks. Prices are in US dollars.

- 0.012
- 0.024
- 0.036
- 0.048
use one of the following prices for semantic segmentation tasks. prices are in us dollars.

- 0.840
- 0.960
- 1.080
- 1.200

use one of the following prices for textract analyze document important form key amazon augmented ai review tasks. prices are in us dollars.

- 2.400
- 2.280
- 2.160
- 2.040
- 1.920
- 1.800
- 1.680
- 1.560
- 1.440
- 1.320
- 1.200
- 1.080
- 0.960
- 0.840
- 0.720
- 0.600
- 0.480
- 0.360
- 0.240
- 0.120
- 0.072
- 0.060
- 0.048
- 0.036
- 0.024
- 0.012
Use one of the following prices for Rekognition DetectModerationLabels Amazon Augmented AI review tasks. Prices are in US dollars.

- 1.200
- 1.080
- 0.960
- 0.840
- 0.720
- 0.600
- 0.480
- 0.360
- 0.240
- 0.120
- 0.072
- 0.060
- 0.048
- 0.036
- 0.024
- 0.012

Use one of the following prices for Amazon Augmented AI custom human review tasks. Prices are in US dollars.

- 1.200
- 1.080
- 0.960
- 0.840
- 0.720
- 0.600
- 0.480
- 0.360
- 0.240
- 0.120
- 0.072
- 0.060
- 0.048
- 0.036
- 0.024
- 0.012

Type: `PublicWorkforceTaskPrice (p. 1317)` object

Required: No

**TaskAvailabilityLifetimeInSeconds**

The length of time that a task remains available for review by human workers.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 864000.

Required: No
TaskCount

The number of distinct workers who will perform the same task on each object. For example, if TaskCount is set to 3 for an image classification labeling job, three workers will classify each input image. Increasing TaskCount can improve label accuracy.

Type: Integer


Required: Yes

TaskDescription

A description for the human worker task.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .+

Required: Yes

TaskKeywords

Keywords used to describe the task so that workers can discover the task.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.


Pattern: ^[A-Za-z0-9]+( [A-Za-z0-9]+)*$

Required: No

TaskTimeLimitInSeconds

The amount of time that a worker has to complete a task. The default value is 3,600 seconds (1 hour).

Type: Integer


Required: No

TaskTitle

A title for the human worker task.

Type: String


Pattern: ^\[\t\n\r -\uD7FF\uE000-\uFFFD]*$

Required: Yes

WorkteamArn

Amazon Resource Name (ARN) of a team of workers. To learn more about the types of workforces and work teams you can create and use with Amazon A2I, see Create and Manage Workforces.
Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workteam/.*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HumanLoopRequestSource
Service: Amazon SageMaker Service

Container for configuring the source of human task requests.

Contents

AwsManagedHumanLoopRequestSource

Specifies whether Amazon Rekognition or Amazon Textract are used as the integration source. The
default field settings and JSON parsing rules are different based on the integration source. Valid values:

Type: String

Valid Values: AWS/Rekognition/DetectModerationLabels/Image/V3 | AWS/Textreat/
AnalyzeDocument/Forms/V1

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HumanTaskConfig
Service: Amazon SageMaker Service

Information required for human workers to complete a labeling task.

Contents

AnnotationConsolidationConfig

Configures how labels are consolidated across human workers.

Type: AnnotationConsolidationConfig (p. 876) object

Required: Yes

MaxConcurrentTaskCount

Defines the maximum number of data objects that can be labeled by human workers at the same time. Also referred to as batch size. Each object may have more than one worker at one time. The default value is 1000 objects.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NumberOfHumanWorkersPerDataObject

The number of human workers that will label an object.

Type: Integer


Required: Yes

PreHumanTaskLambdaArn

The Amazon Resource Name (ARN) of a Lambda function that is run before a data object is sent to a human worker. Use this function to provide input to a custom labeling job.

For built-in task types, use one of the following Amazon SageMaker Ground Truth Lambda function ARNs for PreHumanTaskLambdaArn. For custom labeling workflows, see Pre-annotation Lambda.

Bounding box - Finds the most similar boxes from different workers based on the Jaccard index of the boxes.

- arn:aws:lambda:us-east-1:432418664414:function:PRE-BoundingBox
- arn:aws:lambda:ca-central-1:918755190332:function:PRE-BoundingBox
- arn:aws:lambda:eu-west-1:568282634449:function:PRE-BoundingBox
- arn:aws:lambda:eu-west-2:487402164563:function:PRE-BoundingBox
- arn:aws:lambda:eu-central-1:203001061592:function:PRE-BoundingBox
- arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-BoundingBox
- arn:aws:lambda:ap-south-1:565803892007:function:PRE-BoundingBox
- arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-BoundingBox
• arn:aws:lambda:ap-southeast-2:454466003867:function:PRE-BoundingBox

**Image classification** - Uses a variant of the Expectation Maximization approach to estimate the true class of an image based on annotations from individual workers.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-ImageMultiClass
• arn:aws:lambda:ca-central-1:918755190332:function:PRE-ImageMultiClass
• arn:aws:lambda:eu-west-1:568282634449:function:PRE-ImageMultiClass
• arn:aws:lambda:eu-west-2:487402164563:function:PRE-ImageMultiClass
• arn:aws:lambda:eu-central-1:203001061592:function:PRE-ImageMultiClass
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-ImageMultiClass
• arn:aws:lambda:ap-south-1:377565633583:function:PRE-ImageMultiClass
• arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-ImageMultiClass

**Multi-label image classification** - Uses a variant of the Expectation Maximization approach to estimate the true classes of an image based on annotations from individual workers.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-ImageMultiClassMultiLabel
• arn:aws:lambda:eu-west-1:568282634449:function:PRE-ImageMultiClassMultiLabel
• arn:aws:lambda:eu-west-2:487402164563:function:PRE-ImageMultiClassMultiLabel
• arn:aws:lambda:eu-central-1:203001061592:function:PRE-ImageMultiClassMultiLabel
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-ImageMultiClassMultiLabel
• arn:aws:lambda:ap-south-1:1565803892007:function:PRE-ImageMultiClassMultiLabel
• arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-ImageMultiClassMultiLabel

**Semantic segmentation** - Treats each pixel in an image as a multi-class classification and treats pixel annotations from workers as "votes" for the correct label.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-SemanticSegmentation
Text classification - Uses a variant of the Expectation Maximization approach to estimate the true class of text based on annotations from individual workers.

- `arn:aws:lambda:us-east-1:432418664414:function:PRE-TextMultiClass`
- `arn:aws:lambda:eu-west-1:568282634449:function:PRE-TextMultiClass`

Multi-label text classification - Uses a variant of the Expectation Maximization approach to estimate the true classes of text based on annotations from individual workers.

- `arn:aws:lambda:eu-west-1:568282634449:function:PRE-TextMultiClassMultiLabel`
• arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-TextMultiClassMultiLabel

**Named entity recognition** - Groups similar selections and calculates aggregate boundaries, resolving to most-assigned label.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-NamedEntityRecognition
• arn:aws:lambda:eu-west-1:568282634449:function:PRE-NamedEntityRecognition

**Video Classification** - Use this task type when you need workers to classify videos using predefined labels that you specify. Workers are shown videos and are asked to choose one label for each video.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-VideoMultiClass
• arn:aws:lambda:us-east-2:266458841044:function:PRE-VideoMultiClass
• arn:aws:lambda:us-west-2:203001061592:function:PRE-VideoMultiClass
• arn:aws:lambda:eu-west-1:568282634449:function:PRE-VideoMultiClass
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-VideoMultiClass
• arn:aws:lambda:ap-northeast-2:845288260483:function:PRE-VideoMultiClass
• arn:aws:lambda:ap-south-1:565803892007:function:PRE-VideoMultiClass
• arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-VideoMultiClass
• arn:aws:lambda:ap-southeast-2:45446003867:function:PRE-VideoMultiClass
• arn:aws:lambda:eu-central-1:918755190332:function:PRE-VideoMultiClass

**Video Frame Object Detection** - Use this task type to have workers identify and locate objects in a sequence of video frames (images extracted from a video) using bounding boxes. For example, you can use this task to ask workers to identify and localize various objects in a series of video frames, such as cars, bikes, and pedestrians.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-VideoObjectDetection
• arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-VideoObjectDetection
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-VideoObjectDetection
• arn:aws:lambda:ap-south-1:565803892007:function:PRE-VideoObjectDetection
• arn:aws:lambda:eu-central-1:203001061592:function:PRE-VideoObjectDetection
• arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-VideoObjectDetection

Video Frame Object Tracking - Use this task type to have workers track the movement of objects in a sequence of video frames (images extracted from a video) using bounding boxes. For example, you can use this task to ask workers to track the movement of objects, such as cars, bikes, and pedestrians.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-VideoObjectTracking
• arn:aws:lambda:us-east-2:266458841044:function:PRE-VideoObjectTracking
• arn:aws:lambda:eu-west-2:568282634449:function:PRE-VideoObjectTracking
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-VideoObjectTracking
• arn:aws:lambda:ap-south-1:565803892007:function:PRE-VideoObjectTracking
• arn:aws:lambda:eu-central-1:203001061592:function:PRE-VideoObjectTracking
• arn:aws:lambda:eu-west-2:487402164563:function:PRE-VideoObjectTracking
• arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-VideoObjectTracking
• arn:aws:lambda:ca-central-1:918755190332:function:PRE-VideoObjectTracking

3D Point Cloud Modalities

Use the following pre-annotation lambdas for 3D point cloud labeling modality tasks. See 3D Point Cloud Task types to learn more.

3D Point Cloud Object Detection - Use this task type when you want workers to classify objects in a 3D point cloud by drawing 3D cuboids around objects. For example, you can use this task type to ask workers to identify different types of objects in a point cloud, such as cars, bikes, and pedestrians.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-3DPointCloudObjectDetection
• arn:aws:lambda:eu-west-2:568282634449:function:PRE-3DPointCloudObjectDetection
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-3DPointCloudObjectDetection
• arn:aws:lambda:ap-south-1:565803892007:function:PRE-3DPointCloudObjectDetection
• arn:aws:lambda:eu-central-1:203001061592:function:PRE-3DPointCloudObjectDetection
• arn:aws:lambda:eu-west-2:487402164563:function:PRE-3DPointCloudObjectDetection
• arn:aws:lambda:eu-southeast-1:377565635583:function:PRE-3DPointCloudObjectDetection
• arn:aws:lambda:ca-central-1:918755190332:function:PRE-3DPointCloudObjectDetection

3D Point Cloud Object Tracking - Use this task type when you want workers to draw 3D cuboids around objects that appear in a sequence of 3D point cloud frames. For example, you can use this task type to ask workers to track the movement of vehicles across multiple point cloud frames.
• arn:aws:lambda:us-east-1:432418664414:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:us-east-2:266458841044:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:us-west-2:081040179340:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:eu-west-1:568282634449:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:ap-southeast-2:454466003867:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:ap-south-1:565803892007:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:eu-central-1:203001061592:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:ap-northeast-2:845288260483:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:eu-west-2:487402164563:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:ap-southeast-1:377565635583:function:PRE-3DPointCloudObjectTracking
• arn:aws:lambda:ca-central-1:918755190332:function:PRE-3DPointCloudObjectTracking

3D Point Cloud Semantic Segmentation - Use this task type when you want workers to create a point-level semantic segmentation masks by painting objects in a 3D point cloud using different colors where each color is assigned to one of the classes you specify.
• arn:aws:lambda:us-east-1:432418664414:function:PRE-3DPointCloudSemanticSegmentation
Use the following ARNs for Label Verification and Adjustment Jobs

Use label verification and adjustment jobs to review and adjust labels. To learn more, see Verify and Adjust Labels.

**Bounding box verification** - Uses a variant of the Expectation Maximization approach to estimate the true class of verification judgement for bounding box labels based on annotations from individual workers.

- `arn:aws:lambda:eu-west-1:568282634449:function:PRE-3DPointCloudSemanticSegmentation`

**Bounding box adjustment** - Finds the most similar boxes from different workers based on the Jaccard index of the adjusted annotations.

- `arn:aws:lambda:us-east-1:432418664414:function:PRE-AdjustmentBoundingBox`
- `arn:aws:lambda:eu-west-1:568282634449:function:PRE-AdjustmentBoundingBox`
- `arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-AdjustmentBoundingBox`
- `arn:aws:lambda:ap-south-1:565803892007:function:PRE-AdjustmentBoundingBox`
- `arn:aws:lambda:eu-central-1:203001061592:function:PRE-AdjustmentBoundingBox`
- `arn:aws:lambda:eu-west-2:487402164563:function:PRE-AdjustmentBoundingBox`
- `arn:aws:lambda:ap-southeast-1:377565633583:function:PRE-AdjustmentBoundingBox`
- `arn:aws:lambda:ca-central-1:918755190332:function:PRE-AdjustmentBoundingBox`
• arn:aws:lambda:ca-central-1:918755190332:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:eu-west-1:568282634449:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:eu-west-2:487402164563:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:eu-central-1:203001061592:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:eu-west-1:568282634449:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:eu-west-2:487402164563:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:eu-central-1:203001061592:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:ap-northeast-2:845288260483:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:ap-south-1:565803892007:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:ap-southeast-1:37756563583:function:PRE-AdjustmentBoundingBox
• arn:aws:lambda:ap-southeast-2:454466003867:function:PRE-AdjustmentBoundingBox

Semantic segmentation verification - Uses a variant of the Expectation Maximization approach to estimate the true class of verification judgment for semantic segmentation labels based on annotations from individual workers.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-VerificationSemanticSegmentation
• arn:aws:lambda:eu-west-1:568282634449:function:PRE-VerificationSemanticSegmentation
• arn:aws:lambda:ap-northeast-1:477331159723:function:PRE-VerificationSemanticSegmentation
• arn:aws:lambda:ap-south-1:565803892007:function:PRE-VerificationSemanticSegmentation
• arn:aws:lambda:ap-southeast-1:37756563583:function:PRE-VerificationSemanticSegmentation

Semantic segmentation adjustment - Treats each pixel in an image as a multi-class classification and treats pixel adjusted annotations from workers as "votes" for the correct label.

• arn:aws:lambda:us-east-1:432418664414:function:PRE-AdjustmentSemanticSegmentation
Amazon SageMaker Amazon Sagemaker API Reference
HumanTaskConfig


**Video Frame Object Detection Adjustment** - Use this task type when you want workers to adjust bounding boxes that workers have added to video frames to classify and localize objects in a sequence of video frames.


**Video Frame Object Tracking Adjustment** - Use this task type when you want workers to adjust bounding boxes that workers have added to video frames to track object movement across a sequence of video frames.


3D point cloud object detection adjustment - Adjust 3D cuboids in a point cloud frame.


3D point cloud object tracking adjustment - Adjust 3D cuboids across a sequence of point cloud frames.
3D point cloud semantic segmentation adjustment - Adjust semantic segmentation masks in a 3D point cloud.


3D point cloud object tracking adjustment - Adjust object tracking in a 3D point cloud.

- arn:aws:lambda:eu-west-1:568282634449:function:PRE-Adjustment3DPointCloudObjectTracking
HumanTaskConfig

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:lambda:[a-z0-9\-]*:[0-9]{12}:function:*`

Required: Yes

PublicWorkforceTaskPrice

The price that you pay for each task performed by an Amazon Mechanical Turk worker.

Type: PublicWorkforceTaskPrice (p. 1317) object

Required: No

TaskAvailabilityLifetimeInSeconds

The length of time that a task remains available for labeling by human workers. The default and maximum values for this parameter depend on the type of workforce you use.

- If you choose the Amazon Mechanical Turk workforce, the maximum is 12 hours (43,200 seconds). The default is 6 hours (21,600 seconds).
- If you choose a private or vendor workforce, the default value is 30 days (2592,000 seconds) for non-AL mode. For most users, the maximum is also 30 days.

Type: Integer

Valid Range: Minimum value of 60.

Required: No

TaskDescription

A description of the task for your human workers.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `.+`

Required: Yes

TaskKeywords

Keywords used to describe the task so that workers on Amazon Mechanical Turk can discover the task.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.


Pattern: `^[A-Za-z0-9]+( [A-Za-z0-9]+)*$`

Required: No

TaskTimeLimitInSeconds

The amount of time that a worker has to complete a task.

If you create a custom labeling job, the maximum value for this parameter is 8 hours (28,800 seconds).
If you create a labeling job using a built-in task type the maximum for this parameter depends on the task type you use:

- For image and text labeling jobs, the maximum is 8 hours (28,800 seconds).
- For 3D point cloud and video frame labeling jobs, the maximum is 30 days (2952,000 seconds) for non-AL mode. For most users, the maximum is also 30 days.

**TaskTitle**

A title for the task for your human workers.

Type: String


**Pattern:** ^[\t\n\r -\uD7FF\uE000-\uFFFD]*$

Required: Yes

**UiConfig**

Information about the user interface that workers use to complete the labeling task.

Type: **UiConfig** (p. 1444) object

Required: Yes

**WorkteamArn**

The Amazon Resource Name (ARN) of the work team assigned to complete the tasks.

Type: String

Length Constraints: Maximum length of 256.

**Pattern:** arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workteam/.*

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HumanTaskUiSummary

Service: Amazon SageMaker Service

Container for human task user interface information.

Contents

CreationTime

A timestamp when SageMaker created the human task user interface.

Type: Timestamp

Required: Yes

HumanTaskUiArn

The Amazon Resource Name (ARN) of the human task user interface.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:human-task-ui/.*

Required: Yes

HumanTaskUiName

The name of the human task user interface.

Type: String


Pattern: ^[a-z0-9\-]([-\-*[a-z0-9]*)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HyperParameterAlgorithmSpecification

Service: Amazon SageMaker Service

Specifies which training algorithm to use for training jobs that a hyperparameter tuning job launches and the metrics to monitor.

Contents

AlgorithmName

The name of the resource algorithm to use for the hyperparameter tuning job. If you specify a value for this parameter, do not specify a value for TrainingImage.

Type: String


Pattern: (arn:aws[a-zA-Z-]*:sagemaker:[a-z0-9-]*:[0-9]{12}:[a-zA-Z-]*\/)?(\[a-zA-Z0-9-\])(0,62)(?<!-)$

Required: No

MetricDefinitions

An array of MetricDefinition (p. 1147) objects that specify the metrics that the algorithm emits.

Type: Array of MetricDefinition (p. 1147) objects

Array Members: Minimum number of 0 items. Maximum number of 40 items.

Required: No

TrainingImage

The registry path of the Docker image that contains the training algorithm. For information about Docker registry paths for built-in algorithms, see Algorithms Provided by Amazon SageMaker: Common Parameters. SageMaker supports both registry/repository[:tag] and registry/repository[@digest] image path formats. For more information, see Using Your Own Algorithms with Amazon SageMaker.

Type: String

Length Constraints: Maximum length of 255.

Pattern: .*

Required: No

TrainingInputMode

The training input mode that the algorithm supports. For more information about input modes, see Algorithms.

Pipe mode

If an algorithm supports Pipe mode, Amazon SageMaker streams data directly from Amazon S3 to the container.

File mode

If an algorithm supports File mode, SageMaker downloads the training data from S3 to the provisioned ML storage volume, and mounts the directory to the Docker volume for the training container.
You must provision the ML storage volume with sufficient capacity to accommodate the data downloaded from S3. In addition to the training data, the ML storage volume also stores the output model. The algorithm container uses the ML storage volume to also store intermediate information, if any.

For distributed algorithms, training data is distributed uniformly. Your training duration is predictable if the input data objects sizes are approximately the same. SageMaker does not split the files any further for model training. If the object sizes are skewed, training won't be optimal as the data distribution is also skewed when one host in a training cluster is overloaded, thus becoming a bottleneck in training.

**FastFile mode**

If an algorithm supports **FastFile** mode, SageMaker streams data directly from S3 to the container with no code changes, and provides file system access to the data. Users can author their training script to interact with these files as if they were stored on disk.

FastFile mode works best when the data is read sequentially. Augmented manifest files aren't supported. The startup time is lower when there are fewer files in the S3 bucket provided.

Type: String

Valid Values: Pipe | File | FastFile

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HyperParameterSpecification

Service: Amazon SageMaker Service

Defines a hyperparameter to be used by an algorithm.

Contents

DefaultValue

The default value for this hyperparameter. If a default value is specified, a hyperparameter cannot be required.

Type: String

Length Constraints: Maximum length of 2500.

Pattern: . *

Required: No

Description

A brief description of the hyperparameter.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: \[\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*

Required: No

IsRequired

Indicates whether this hyperparameter is required.

Type: Boolean

Required: No

IsTunable

Indicates whether this hyperparameter is tunable in a hyperparameter tuning job.

Type: Boolean

Required: No

Name

The name of this hyperparameter. The name must be unique.

Type: String

Length Constraints: Maximum length of 256.

Pattern: \[\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*

Required: Yes

Range

The allowed range for this hyperparameter.
Type: ParameterRange (p. 1249) object

Required: No

Type

The type of this hyperparameter. The valid types are Integer, Continuous, Categorical, and FreeText.

Type: String

Valid Values: Integer | Continuous | Categorical | FreeText

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HyperParameterTrainingJobDefinition

Service: Amazon SageMaker Service

Defines the training jobs launched by a hyperparameter tuning job.

Contents

AlgorithmSpecification

The HyperParameterAlgorithmSpecification (p. 1078) object that specifies the resource algorithm to use for the training jobs that the tuning job launches.

Type: HyperParameterAlgorithmSpecification (p. 1078) object

Required: Yes

CheckpointConfig

Contains information about the output location for managed spot training checkpoint data.

Type: CheckpointConfig (p. 945) object

Required: No

DefinitionName

The job definition name.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,63}$

Required: No

EnableInterContainerTrafficEncryption

To encrypt all communications between ML compute instances in distributed training, choose True. Encryption provides greater security for distributed training, but training might take longer. How long it takes depends on the amount of communication between compute instances, especially if you use a deep learning algorithm in distributed training.

Type: Boolean

Required: No

EnableManagedSpotTraining

A Boolean indicating whether managed spot training is enabled (True) or not (False).

Type: Boolean

Required: No

EnableNetworkIsolation

Isolates the training container. No inbound or outbound network calls can be made, except for calls between peers within a training cluster for distributed training. If network isolation is used for training jobs that are configured to use a VPC, SageMaker downloads and uploads customer data and model artifacts through the specified VPC, but the training container does not have network access.

Type: Boolean
Required: No

**HyperParameterRanges**

Specifies ranges of integer, continuous, and categorical hyperparameters that a hyperparameter tuning job searches. The hyperparameter tuning job launches training jobs with hyperparameter values within these ranges to find the combination of values that result in the training job with the best performance as measured by the objective metric of the hyperparameter tuning job.

**Note**

You can specify a maximum of 20 hyperparameters that a hyperparameter tuning job can search over. Every possible value of a categorical parameter range counts against this limit.

Type: `ParameterRanges (p. 1250)` object

Required: No

**InputDataConfig**

An array of `Channel (p. 941)` objects that specify the input for the training jobs that the tuning job launches.

Type: Array of `Channel (p. 941)` objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Required: No

**OutputDataConfig**

Specifies the path to the Amazon S3 bucket where you store model artifacts from the training jobs that the tuning job launches.

Type: `OutputDataConfig (p. 1244)` object

Required: Yes

**ResourceConfig**

The resources, including the compute instances and storage volumes, to use for the training jobs that the tuning job launches.

Storage volumes store model artifacts and incremental states. Training algorithms might also use storage volumes for scratch space. If you want SageMaker to use the storage volume to store the training data, choose File as the `TrainingInputMode` in the algorithm specification. For distributed training algorithms, specify an instance count greater than 1.

Type: `ResourceConfig (p. 1338)` object

Required: Yes

**RetryStrategy**

The number of times to retry the job when the job fails due to an `InternalServerError`.

Type: `RetryStrategy (p. 1344)` object

Required: No

**RoleArn**

The Amazon Resource Name (ARN) of the IAM role associated with the training jobs that the tuning job launches.

Type: String

Pattern: ^arn:aws[a-zA-z\-]*:iam::\d{12}:role/?[a-zA-z0-9+=,.@/-]+$

Required: Yes

**StaticHyperParameters**

Specifies the values of hyperparameters that do not change for the tuning job.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 2500.

Value Pattern: .*

Required: No

**StoppingCondition**

Specifies a limit to how long a model hyperparameter training job can run. It also specifies how long a managed spot training job has to complete. When the job reaches the time limit, SageMaker ends the training job. Use this API to cap model training costs.

Type: StoppingCondition (p. 1370) object

Required: Yes

**TuningObjective**

Defines the objective metric for a hyperparameter tuning job. Hyperparameter tuning uses the value of this metric to evaluate the training jobs it launches, and returns the training job that results in either the highest or lowest value for this metric, depending on the value you specify for the `Type` parameter.

Type: HyperParameterTuningJobObjective (p. 1091) object

Required: No

**VpcConfig**

The VpcConfig (p. 1456) object that specifies the VPC that you want the training jobs that this hyperparameter tuning job launches to connect to. Control access to and from your training container by configuring the VPC. For more information, see Protect Training Jobs by Using an Amazon Virtual Private Cloud.

Type: VpcConfig (p. 1456) object

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
HyperParameterTrainingJobSummary

Service: Amazon SageMaker Service

The container for the summary information about a training job.

Contents

CreationTime

The date and time that the training job was created.

Type: Timestamp

Required: Yes

FailureReason

The reason that the training job failed.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

FinalHyperParameterTuningJobObjectiveMetric

The FinalHyperParameterTuningJobObjectiveMetric (p. 1050) object that specifies the value of the objective metric of the tuning job that launched this training job.

Type: FinalHyperParameterTuningJobObjectiveMetric (p. 1050) object

Required: No

ObjectiveStatus

The status of the objective metric for the training job:

- Succeeded: The final objective metric for the training job was evaluated by the hyperparameter tuning job and used in the hyperparameter tuning process.
- Pending: The training job is in progress and evaluation of its final objective metric is pending.
- Failed: The final objective metric for the training job was not evaluated, and was not used in the hyperparameter tuning process. This typically occurs when the training job failed or did not emit an objective metric.

Type: String

Valid Values: Succeeded | Pending | Failed

Required: No

TrainingEndTime

Specifies the time when the training job ends on training instances. You are billed for the time interval between the value of TrainingStartTime and this time. For successful jobs and stopped jobs, this is the time after model artifacts are uploaded. For failed jobs, this is the time when SageMaker detects a job failure.

Type: Timestamp

Required: No
**TrainingJobArn**

The Amazon Resource Name (ARN) of the training job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:training-job/.*`

Required: Yes

**TrainingJobDefinitionName**

The training job definition name.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,63}`

Required: No

**TrainingJobName**

The name of the training job.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}`

Required: Yes

**TrainingJobStatus**

The status of the training job.

Type: String

Valid Values: `InProgress` | `Completed` | `Failed` | `Stopping` | `Stopped`

Required: Yes

**TrainingStartTime**

The date and time that the training job started.

Type: Timestamp

Required: No

**TunedHyperParameters**

A list of the hyperparameters for which you specified ranges to search.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: `.*`
Value Length Constraints: Maximum length of 2500.
Value Pattern: .*
Required: Yes

**TuningJobName**

The HyperParameter tuning job that launched the training job.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HyperParameterTuningJobConfig

Service: Amazon SageMaker Service

Configures a hyperparameter tuning job.

Contents

HyperParameterTuningJobObjective

The HyperParameterTuningJobObjective object that specifies the objective metric for this tuning job.

Type: HyperParameterTuningJobObjective object

Required: No

ParameterRanges

The ParameterRanges object that specifies the ranges of hyperparameters that this tuning job searches.

Type: ParameterRanges object

Required: No

ResourceLimits

The ResourceLimits object that specifies the maximum number of training jobs and parallel training jobs for this tuning job.

Type: ResourceLimits object

Required: Yes

Strategy

Specifies how hyperparameter tuning chooses the combinations of hyperparameter values to use for the training job it launches. To use the Bayesian search strategy, set this to Bayesian. To randomly search, set it to Random. For information about search strategies, see How Hyperparameter Tuning Works.

Type: String

Valid Values: Bayesian | Random

Required: Yes

TrainingJobEarlyStoppingType

Specifies whether to use early stopping for training jobs launched by the hyperparameter tuning job. This can be one of the following values (the default value is OFF):

OFF

Training jobs launched by the hyperparameter tuning job do not use early stopping.

AUTO

SageMaker stops training jobs launched by the hyperparameter tuning job when they are unlikely to perform better than previously completed training jobs. For more information, see Stop Training Jobs Early.

Type: String
Valid Values: Off | Auto

Required: No

**TuningJobCompletionCriteria**

The tuning job's completion criteria.

Type: TuningJobCompletionCriteria (p. 1442) object

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HyperParameterTuningJobObjective

Service: Amazon SageMaker Service

Defines the objective metric for a hyperparameter tuning job. Hyperparameter tuning uses the value of this metric to evaluate the training jobs it launches, and returns the training job that results in either the highest or lowest value for this metric, depending on the value you specify for the `Type` parameter.

Contents

MetricName

The name of the metric to use for the objective metric.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .+

Required: Yes

Type

Whether to minimize or maximize the objective metric.

Type: String

Valid Values: Maximize | Minimize

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HyperParameterTuningJobSummary

Service: Amazon SageMaker Service

Provides summary information about a hyperparameter tuning job.

Contents

CreationTime

The date and time that the tuning job was created.

Type: Timestamp

Required: Yes

HyperParameterTuningEndTime

The date and time that the tuning job ended.

Type: Timestamp

Required: No

HyperParameterTuningJobArn

The Amazon Resource Name (ARN) of the tuning job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:hyper-parameter-tuning-job/.*

Required: Yes

HyperParameterTuningJobName

The name of the tuning job.

Type: String


Pattern: ^[a-zA-Z0-9\-]*(\-[a-zA-Z0-9])\{0,31}$

Required: Yes

HyperParameterTuningJobStatus

The status of the tuning job.

Type: String

Valid Values: Completed | InProgress | Failed | Stopped | Stopping

Required: Yes

LastModifiedTime

The date and time that the tuning job was modified.

Type: Timestamp
Required: No

**ObjectiveStatusCounters**

The *ObjectiveStatusCounters* object that specifies the numbers of training jobs, categorized by objective metric status, that this tuning job launched.

Type: *ObjectiveStatusCounters* object

Required: Yes

**ResourceLimits**

The *ResourceLimits* object that specifies the maximum number of training jobs and parallel training jobs allowed for this tuning job.

Type: *ResourceLimits* object

Required: No

**Strategy**

Specifies the search strategy hyperparameter tuning uses to choose which hyperparameters to use for each iteration. Currently, the only valid value is Bayesian.

Type: String

Valid Values: Bayesian | Random

Required: Yes

**TrainingJobStatusCounters**

The *TrainingJobStatusCounters* object that specifies the numbers of training jobs, categorized by status, that this tuning job launched.

Type: *TrainingJobStatusCounters* object

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
HyperParameterTuningJobWarmStartConfig

Service: Amazon SageMaker Service

Specifies the configuration for a hyperparameter tuning job that uses one or more previous hyperparameter tuning jobs as a starting point. The results of previous tuning jobs are used to inform which combinations of hyperparameters to search over in the new tuning job.

All training jobs launched by the new hyperparameter tuning job are evaluated by using the objective metric, and the training job that performs the best is compared to the best training jobs from the parent tuning jobs. From these, the training job that performs the best as measured by the objective metric is returned as the overall best training job.

**Note**

All training jobs launched by parent hyperparameter tuning jobs and the new hyperparameter tuning jobs count against the limit of training jobs for the tuning job.

**Contents**

**ParentHyperParameterTuningJobs**

An array of hyperparameter tuning jobs that are used as the starting point for the new hyperparameter tuning job. For more information about warm starting a hyperparameter tuning job, see Using a Previous Hyperparameter Tuning Job as a Starting Point.

Hyperparameter tuning jobs created before October 1, 2018 cannot be used as parent jobs for warm start tuning jobs.

Type: Array of ParentHyperParameterTuningJob (p. 1252) objects

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Required: Yes

**WarmStartType**

Specifies one of the following:

**IDENTICAL_DATA_AND_ALGORITHM**

The new hyperparameter tuning job uses the same input data and training image as the parent tuning jobs. You can change the hyperparameter ranges to search and the maximum number of training jobs that the hyperparameter tuning job launches. You cannot use a new version of the training algorithm, unless the changes in the new version do not affect the algorithm itself. For example, changes that improve logging or adding support for a different data format are allowed. You can also change hyperparameters from tunable to static, and from static to tunable, but the total number of static plus tunable hyperparameters must remain the same as it is in all parent jobs. The objective metric for the new tuning job must be the same as for all parent jobs.

**TRANSFER_LEARNING**

The new hyperparameter tuning job can include input data, hyperparameter ranges, maximum number of concurrent training jobs, and maximum number of training jobs that are different than those of its parent hyperparameter tuning jobs. The training image can also be a different version from the version used in the parent hyperparameter tuning job. You can also change hyperparameters from tunable to static, and from static to tunable, but the total number of static plus tunable hyperparameters must remain the same as it is in all parent jobs. The objective metric for the new tuning job must be the same as for all parent jobs.

Type: String
Valid Values: IdenticalDataAndAlgorithm | TransferLearning

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Image
Service: Amazon SageMaker Service

A SageMaker image. A SageMaker image represents a set of container images that are derived from a common base container image. Each of these container images is represented by a SageMaker ImageVersion.

Contents

CreationTime
When the image was created.
Type: Timestamp
Required: Yes

Description
The description of the image.
Type: String
Pattern: .*
Required: No

DisplayName
The name of the image as displayed.
Type: String
Pattern: ^\S(.*\S)?$  
Required: No

FailureReason
When a create, update, or delete operation fails, the reason for the failure.
Type: String
Length Constraints: Maximum length of 1024.
Required: No

ImageArn
The Amazon Resource Name (ARN) of the image.
Type: String
Length Constraints: Maximum length of 256.
Pattern: ^arn:aws(-\[\w]+)*:sagemaker:+:[0-9]{12}:image/[a-z0-9]*\[.-]?[a-z0-9]*$  
Required: Yes
**ImageName**

The name of the image.

Type: String


Pattern: `^[a-zA-Z0-9][.-][a-zA-Z0-9]{0,62}$`

Required: Yes

**ImageStatus**

The status of the image.

Type: String

Valid Values: CREATING | CREATED | CREATE_FAILED | UPDATING | UPDATE_FAILED | DELETING | DELETE_FAILED

Required: Yes

**LastModifiedTime**

When the image was last modified.

Type: Timestamp

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**ImageConfig**

Service: Amazon SageMaker Service

Specifies whether the model container is in Amazon ECR or a private Docker registry accessible from your Amazon Virtual Private Cloud (VPC).

**Contents**

**RepositoryAccessMode**

Set this to one of the following values:

- **Platform** - The model image is hosted in Amazon ECR.
- **Vpc** - The model image is hosted in a private Docker registry in your VPC.

Type: String

Valid Values: Platform | Vpc

Required: Yes

**RepositoryAuthConfig**

(Optional) Specifies an authentication configuration for the private docker registry where your model image is hosted. Specify a value for this property only if you specified Vpc as the value for the RepositoryAccessMode field, and the private Docker registry where the model image is hosted requires authentication.

Type: RepositoryAuthConfig (p. 1336) object

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**ImageVersion**

*Service: Amazon SageMaker Service*

A version of a SageMaker Image. A version represents an existing container image.

**Contents**

**CreationTime**

When the version was created.

Type: Timestamp

Required: Yes

**FailureReason**

When a create or delete operation fails, the reason for the failure.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

**ImageArn**

The Amazon Resource Name (ARN) of the image the version is based on.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^arn:aws(-[\w]+)*:sagemaker:.+:([0-9]{12}):image/[a-z0-9]([\.-]?[a-z0-9])*$`

Required: Yes

**ImageVersionArn**

The ARN of the version.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^arn:aws(-[\w]+)*:sagemaker:.+:([0-9]{12}):image-version/[a-z0-9]([\.-]?[a-z0-9])*[/0-9]+$`

Required: Yes

**ImageVersionStatus**

The status of the version.

Type: String

Valid Values: CREATING | CREATED | CREATE_FAILED | DELETING | DELETE_FAILED

Required: Yes

**LastModifiedTime**

When the version was last modified.
Type: Timestamp
Required: Yes

Version
The version number.
Type: Integer
Valid Range: Minimum value of 0.
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**InferenceExecutionConfig**

Service: Amazon SageMaker Service

Specifies details about how containers in a multi-container endpoint are run.

**Contents**

**Mode**

How containers in a multi-container are run. The following values are valid.

- **SERIAL** - Containers run as a serial pipeline.
- **DIRECT** - Only the individual container that you specify is run.

Type: String

Valid Values: Serial | Direct

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
InferenceRecommendation
Service: Amazon SageMaker Service

A list of recommendations made by Amazon SageMaker Inference Recommender.

Contents

EndpointConfiguration
Defines the endpoint configuration parameters.
Type: EndpointOutputConfiguration (p. 1021) object
Required: Yes

Metrics
The metrics used to decide what recommendation to make.
Type: RecommendationMetrics (p. 1330) object
Required: Yes

ModelConfiguration
Defines the model configuration.
Type: ModelConfiguration (p. 1154) object
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
InferenceRecommendationsJob
Service: Amazon SageMaker Service
A structure that contains a list of recommendation jobs.

Contents

CompletionTime
A timestamp that shows when the job completed.
Type: Timestamp
Required: No

CreationTime
A timestamp that shows when the job was created.
Type: Timestamp
Required: Yes

FailureReason
If the job fails, provides information why the job failed.
Type: String
Length Constraints: Maximum length of 1024.
Required: No

JobArn
The Amazon Resource Name (ARN) of the recommendation job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:inference-recommendations-job/.*
Required: Yes

JobDescription
The job description.
Type: String
Length Constraints: Maximum length of 128.
Required: Yes

JobName
The name of the job.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 64.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,63}$

Required: Yes

**JobType**

The recommendation job type.

Type: String

Valid Values: Default | Advanced

Required: Yes

**LastModifiedTime**

A timestamp that shows when the job was last modified.

Type: Timestamp

Required: Yes

**RoleArn**

The Amazon Resource Name (ARN) of an IAM role that enables Amazon SageMaker to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z-Z_0-9+=,.@\-_/]+$

Required: Yes

**Status**

The status of the job.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | FAILED | STOPPING | STOPPED

Required: Yes

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
InferenceSpecification
Service: Amazon SageMaker Service

Defines how to perform inference generation after a training job is run.

Contents

Containers

The Amazon ECR registry path of the Docker image that contains the inference code.

Type: Array of ModelPackageContainerDefinition (p. 1173) objects

Array Members: Minimum number of 1 item. Maximum number of 15 items.

Required: Yes

SupportedContentTypes

The supported MIME types for the input data.

Type: Array of strings

Length Constraints: Maximum length of 256.

Pattern: .*

Required: Yes

SupportedRealtimeInferenceInstanceTypes

A list of the instance types that are used to generate inferences in real-time.

This parameter is required for unversioned models, and optional for versioned models.

Type: Array of strings

Valid Values:
- ml.t2.medium
- ml.t2.large
- ml.t2.xlarge
- ml.t2.2xlarge
- ml.m4.xlarge
- ml.m4.2xlarge
- ml.m4.4xlarge
- ml.m4.10xlarge
- ml.m4.16xlarge
- ml.m5.large
- ml.m5.xlarge
- ml.m5.2xlarge
- ml.m5.4xlarge
- ml.m5.12xlarge
- ml.m5.24xlarge
- ml.m5d.large
- ml.m5d.xlarge
- ml.m5d.2xlarge
- ml.m5d.4xlarge
- ml.m5d.12xlarge
- ml.m5d.24xlarge
- ml.c4.large
- ml.c4.xlarge
- ml.c4.2xlarge
- ml.c4.4xlarge
- ml.c4.8xlarge
- ml.p2.large
- ml.p2.xlarge
- ml.p2.8xlarge
- ml.p2.16xlarge
- ml.p3.large
- ml.p3.xlarge
- ml.p3.8xlarge
- ml.p3.16xlarge
- ml.c5.large
- ml.c5.xlarge
- ml.c5.2xlarge
- ml.c5.4xlarge
- ml.c5.9xlarge
- ml.c5.18xlarge
- ml.c5d.large
- ml.c5d.xlarge
- ml.c5d.2xlarge
- ml.c5d.4xlarge
- ml.c5d.9xlarge
- ml.c5d.18xlarge
- ml.g4dn.xlarge
- ml.g4dn.xlarge
- ml.g4dn.2xlarge
- ml.g4dn.4xlarge
- ml.g4dn.8xlarge
- ml.g4dn.16xlarge
- ml.r5.large
- ml.r5.xlarge
- ml.r5.2xlarge
- ml.r5.4xlarge
- ml.r5.8xlarge
- ml.r5.16xlarge
- ml.r5d.large
- ml.r5d.xlarge
- ml.r5d.2xlarge
- ml.r5d.4xlarge
- ml.r5d.8xlarge
- ml.r5d.16xlarge
- ml.inf1.xlarge
- ml.inf1.2xlarge
- ml.inf1.6xlarge
- ml.inf1.24xlarge

Required: No

SupportedResponseMIMETypes

The supported MIME types for the output data.

Type: Array of strings
Length Constraints: Maximum length of 1024.
Pattern: ^[-\w]+/.++$
Required: Yes

**SupportedTransformInstanceTypes**

A list of the instance types on which a transformation job can be run or on which an endpoint can be deployed.

This parameter is required for unversioned models, and optional for versioned models.

Type: Array of strings

Array Members: Minimum number of 1 item.

Valid Values:
- `ml.m4.xlarge`
- `ml.m4.2xlarge`
- `ml.m4.4xlarge`
- `ml.m4.10xlarge`
- `ml.m4.16xlarge`
- `ml.c4.xlarge`
- `ml.c4.2xlarge`
- `ml.c4.4xlarge`
- `ml.c4.8xlarge`
- `ml.p2.xlarge`
- `ml.p2.8xlarge`
- `ml.p2.16xlarge`
- `ml.p3.2xlarge`
- `ml.p3.8xlarge`
- `ml.p3.16xlarge`
- `ml.c5.xlarge`
- `ml.c5.2xlarge`
- `ml.c5.4xlarge`
- `ml.c5.9xlarge`
- `ml.c5.18xlarge`
- `ml.m5.xlarge`
- `ml.m5.2xlarge`
- `ml.m5.4xlarge`
- `ml.m5.12xlarge`
- `ml.m5.24xlarge`
- `ml.g4dn.xlarge`
- `ml.g4dn.2xlarge`
- `ml.g4dn.4xlarge`
- `ml.g4dn.8xlarge`
- `ml.g4dn.12xlarge`
- `ml.g4dn.16xlarge`

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
InputConfig
Service: Amazon SageMaker Service

Contains information about the location of input model artifacts, the name and shape of the expected data inputs, and the framework in which the model was trained.

Contents

DataInputConfig

Specifies the name and shape of the expected data inputs for your trained model with a JSON dictionary form. The data inputs are InputConfig:Framework (p. 1109) specific.

- **TensorFlow**: You must specify the name and shape (NHWC format) of the expected data inputs using a dictionary format for your trained model. The dictionary formats required for the console and CLI are different.
  - Examples for one input:
    - If using the console, {"input":[1,1024,1024,3]}
    - If using the CLI, {"input":[1,1024,1024,3]}
  - Examples for two inputs:
    - If using the console, {"data1": [1,28,28,1], "data2": [1,28,28,1]}
    - If using the CLI, {"data1": [1,28,28,1], "data2": [1,28,28,1]}

- **Keras**: You must specify the name and shape (NCHW format) of expected data inputs using a dictionary format for your trained model. Note that while Keras model artifacts should be uploaded in NHWC (channel-last) format, DataInputConfig should be specified in NCHW (channel-first) format. The dictionary formats required for the console and CLI are different.
  - Examples for one input:
    - If using the console, {"input_1": [1,3,224,224]}
    - If using the CLI, {"input_1": [1,3,224,224]}
  - Examples for two inputs:
    - If using the console, {"input_1": [1,3,224,224], "input_2": [1,3,224,224]}
    - If using the CLI, {"input_1": [1,3,224,224], "input_2": [1,3,224,224]}

- **MXNet/ONNX/DARKNET**: You must specify the name and shape (NCHW format) of the expected data inputs in order using a dictionary format for your trained model. The dictionary formats required for the console and CLI are different.
  - Examples for one input:
    - If using the console, {"data": [1,3,1024,1024]}
    - If using the CLI, {"data": [1,3,1024,1024]}
  - Examples for two inputs:
    - If using the console, {"var1": [1,1,28,28], "var2": [1,1,28,28]}
    - If using the CLI, {"var1": [1,1,28,28], "var2": [1,1,28,28]}

- **PyTorch**: You can either specify the name and shape (NCHW format) of expected data inputs in order using a dictionary format for your trained model or you can specify the shape only using a list format. The dictionary formats required for the console and CLI are different. The list formats for the console and CLI are the same.
  - Examples for one input in dictionary format:
    - If using the console, {"input0": [1,3,224,224]}
    - If using the CLI, {"input0": [1,3,224,224]}
  - Example for one input in list format: [[1,3,224,224]]
  - Examples for two inputs in dictionary format:
• If using the console, 
  ```json
  {"input0": [1,3,224,224], "input1": [1,3,224,224]}
  ```
• If using the CLI, 
  ```json
  \"input0\": [1,3,224,224], \"input1\": [1,3,224,224]
  ```
• Example for two inputs in list format: 
  ```json
  [[1,3,224,224], [1,3,224,224]]
  ```
• **XGBOOST**: input data name and shape are not needed.

**DataInputConfig** supports the following parameters for **CoreML**:

**OutputConfig**:TargetDevice (p. 1241) (ML Model format):

- **shape**: Input shape, for example 
  ```json
  "input_1": {"shape": [1,224,224,3]}
  ```

  In addition to static input shapes, CoreML converter supports Flexible input shapes:
  - **Range Dimension**: You can use the Range Dimension feature if you know the input shape will be within some specific interval in that dimension, for example:
    ```json
    "input_1": {"shape": ["1..10", 224, 224, 3]}
    ```
  - **Enumerated shapes**: Sometimes, the models are trained to work only on a select set of inputs. You can enumerate all supported input shapes, for example:
    ```json
    "input_1": {"shape": [[1, 224, 224, 3], [1, 160, 160, 3]]}
    ```

- **default_shape**: Default input shape. You can set a default shape during conversion for both Range Dimension and Enumerated Shapes. For example:
  ```json
  "input_1": {"shape": ["1..10", 224, 224, 3], "default_shape": [1, 224, 224, 3]}
  ```

- **type**: Input type. Allowed values: Image and Tensor. By default, the converter generates an ML Model with inputs of type Tensor (MultiArray). User can set input type to be Image. Image input type requires additional input parameters such as bias and scale.

- **bias**: If the input type is an Image, you need to provide the bias vector.

- **scale**: If the input type is an Image, you need to provide a scale factor.

**CoreML ClassifierConfig** parameters can be specified using **OutputConfig:CompilerOptions** (p. 1240). CoreML converter supports Tensorflow and PyTorch models. CoreML conversion examples:

- **Tensor type input**:
  ```json
  "DataInputConfig": {"input_1": {"shape": [[1,224,224,3], [1,160,160,3]],
  "default_shape": [1,224,224,3]}}
  ```

- **Tensor type input without input name (PyTorch)**:
  ```json
  "DataInputConfig": [{"shape": [[1,3,224,224], [1,3,160,160]],
  "default_shape": [1,3,224,224]}]
  ```

- **Image type input**:
  ```json
  "DataInputConfig": {"input_1": {"shape": [[1,224,224,3], [1,160,160,3]],
  "default_shape": [1,224,224,3],
  "type": "Image",
  "bias": [-1,-1,-1],
  "scale": 0.007843137255}}
  ```

  ```json
  "CompilerOptions": {"class_labels": "imagenet_labels_1000.txt"}
  ```

- **Image type input without input name (PyTorch)**:
  ```json
  "DataInputConfig": [{"shape": [[1,3,224,224], [1,3,160,160]],
  "default_shape": [1,3,224,224],
  "type": "Image",
  "bias": [-1,-1,-1],
  "scale": 0.007843137255}]
  ```

  ```json
  "CompilerOptions": {"class_labels": "imagenet_labels_1000.txt"}
  ```

Depending on the model format, **DataInputConfig** requires the following parameters for **ml_eia2 OutputConfig:TargetDevice**:

- For TensorFlow models saved in the SavedModel format, specify the input names from signature_def_key and the input model shapes for DataInputConfig. Specify the signature_def_key in **OutputConfig:CompilerOptions** if the model does not use TensorFlow's default signature def key. For example:
  ```json
  "DataInputConfig": {"inputs": [1, 224, 224, 3]}
  ```

  ```json
  "CompilerOptions": {"signature_def_key": "serving_custom"}
  ```
For TensorFlow models saved as a frozen graph, specify the input tensor names and shapes in `DataInputConfig` and the output tensor names for `output_names` in `OutputConfig`. For example:

- "DataInputConfig": {"input_tensor:0": [1, 224, 224, 3]}
- "CompilerOptions": {"output_names": ["output_tensor:0"]}

Type: String


Pattern: [\S\s]+

Required: Yes

**Framework**

Identifies the framework in which the model was trained. For example: TENSORFLOW.

Type: String

Valid Values: TENSORFLOW | KERAS | MXNET | ONNX | PYTORCH | XGBOOST | TFLITE | DARKNET | SKLEARN

Required: Yes

**FrameworkVersion**

Specifies the framework version to use. This API field is only supported for the PyTorch and TensorFlow frameworks.

For information about framework versions supported for cloud targets and edge devices, see Cloud Supported Instance Types and Frameworks and Edge Supported Frameworks.

Type: String


Pattern: [0-9]\.[A-Za-z0-9-9.]+

Required: No

**S3Uri**

The S3 path where the model artifacts, which result from model training, are stored. This path must point to a single gzip compressed tar archive (.tar.gz suffix).

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)//([^/]+)?(.*$)

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for Ruby V3
InstanceMetadataServiceConfiguration

Service: Amazon SageMaker Service

Information on the IMDS configuration of the notebook instance

Contents

MinimumInstanceMetadataServiceVersion

Indicates the minimum IMDS version that the notebook instance supports. When passed as part of CreateNotebookInstance, if no value is selected, then it defaults to IMDSv1. This means that both IMDSv1 and IMDSv2 are supported. If passed as part of UpdateNotebookInstance, there is no default.

Type: String

Length Constraints: Maximum length of 1.

Pattern: 1 | 2

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
IntegerParameterRange
Service: Amazon SageMaker Service

For a hyperparameter of the integer type, specifies the range that a hyperparameter tuning job searches.

Contents

MaxValue

The maximum value of the hyperparameter to search.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: Yes

MinValue

The minimum value of the hyperparameter to search.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: Yes

Name

The name of the hyperparameter to search.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: Yes

ScalingType

The scale that hyperparameter tuning uses to search the hyperparameter range. For information about choosing a hyperparameter scale, see Hyperparameter Scaling. One of the following values:

Auto

SageMaker hyperparameter tuning chooses the best scale for the hyperparameter.

Linear

Hyperparameter tuning searches the values in the hyperparameter range by using a linear scale.

Logarithmic

Hyperparameter tuning searches the values in the hyperparameter range by using a logarithmic scale.

Logarithmic scaling works only for ranges that have only values greater than 0.

Type: String
Valid Values: Auto | Linear | Logarithmic | ReverseLogarithmic

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
IntegerParameterRangeSpecification

Service: Amazon SageMaker Service

Defines the possible values for an integer hyperparameter.

Contents

MaxValue

The maximum integer value allowed.

Type: String

Length Constraints: Maximum length of 256.

Pattern: . *

Required: Yes

MinValue

The minimum integer value allowed.

Type: String

Length Constraints: Maximum length of 256.

Pattern: . *

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
JupyterServerAppSettings
Service: Amazon SageMaker Service

The JupyterServer app settings.

Contents

DefaultResourceSpec

The default instance type and the Amazon Resource Name (ARN) of the default SageMaker image used by the JupyterServer app. If you use the LifecycleConfigArns parameter, then this parameter is also required.

Type: ResourceSpec (p. 1341) object

Required: No

LifecycleConfigArns

The Amazon Resource Name (ARN) of the Lifecycle Configurations attached to the JupyterServerApp. If you use this parameter, the DefaultResourceSpec parameter is also required.

Note

To remove a Lifecycle Config, you must set LifecycleConfigArns to an empty list.

Type: Array of strings

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:studio-lifecycle-config/.*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
KernelGatewayAppSettings
Service: Amazon SageMaker Service
The KernelGateway app settings.

Contents

CustomImages
A list of custom SageMaker images that are configured to run as a KernelGateway app.
Type: Array of CustomImage (p. 965) objects
Array Members: Maximum number of 200 items.
Required: No

DefaultResourceSpec
The default instance type and the Amazon Resource Name (ARN) of the default SageMaker image used by the KernelGateway app.

Note
The Amazon SageMaker Studio UI does not use the default instance type value set here. The default instance type set here is used when Apps are created using the AWS Command Line Interface or AWS CloudFormation and the instance type parameter value is not passed.
Type: ResourceSpec (p. 1341) object
Required: No

LifecycleConfigArns
The Amazon Resource Name (ARN) of the Lifecycle Configurations attached to the the user profile or domain.

Note
To remove a Lifecycle Config, you must set LifecycleConfigArns to an empty list.
Type: Array of strings
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:studio-lifecycle-config/.*
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
KernelGatewayImageConfig

Service: Amazon SageMaker Service

The configuration for the file system and kernels in a SageMaker image running as a KernelGateway app.

Contents

FileSystemConfig

The Amazon Elastic File System (EFS) storage configuration for a SageMaker image.

Type: FileSystemConfig (p. 1043) object

Required: No

KernelSpecs

The specification of the Jupyter kernels in the image.

Type: Array of KernelSpec (p. 1118) objects

Array Members: Fixed number of 1 item.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**KernelSpec**

Service: Amazon SageMaker Service

The specification of a Jupyter kernel.

**Contents**

**DisplayName**

The display name of the kernel.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

**Name**

The name of the Jupyter kernel in the image. This value is case sensitive.

Type: String

Length Constraints: Maximum length of 1024.

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelCounters
Service: Amazon SageMaker Service
Provides a breakdown of the number of objects labeled.

Contents

FailedNonRetryableError
The total number of objects that could not be labeled due to an error.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

HumanLabeled
The total number of objects labeled by a human worker.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

MachineLabeled
The total number of objects labeled by automated data labeling.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

TotalLabeled
The total number of objects labeled.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

Unlabeled
The total number of objects not yet labeled.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
LabelCountersForWorkteam

Service: Amazon SageMaker Service

Provides counts for human-labeled tasks in the labeling job.

Contents

HumanLabeled

The total number of data objects labeled by a human worker.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PendingHuman

The total number of data objects that need to be labeled by a human worker.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Total

The total number of tasks in the labeling job.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobAlgorithmsConfig
Service: Amazon SageMaker Service

Provides configuration information for auto-labeling of your data objects. A LabelingJobAlgorithmsConfig object must be supplied in order to use auto-labeling.

Contents

InitialActiveLearningModelArn

At the end of an auto-label job Ground Truth sends the Amazon Resource Name (ARN) of the final model used for auto-labeling. You can use this model as the starting point for subsequent similar jobs by providing the ARN of the model here.

Type: String


Pattern: \arn:\aws[a-z\-]*:sagemaker:\[a-z0-9\-]*:[0-9]{12}:model/.*

Required: No

LabelingJobAlgorithmSpecificationArn

Specifies the Amazon Resource Name (ARN) of the algorithm used for auto-labeling. You must select one of the following ARNs:

- **Image classification**
  

- **Text classification**
  

- **Object detection**
  

- **Semantic Segmentation**
  

Type: String

Length Constraints: Maximum length of 2048.

Pattern: \arn:.*

Required: Yes

LabelingJobResourceConfig

Provides configuration information for a labeling job.

Type: LabelingJobResourceConfig (p. 1132) object

Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobDataAttributes
Service: Amazon SageMaker Service

Attributes of the data specified by the customer. Use these to describe the data to be labeled.

Contents

ContentClassifiers

Declares that your content is free of personally identifiable information or adult content. SageMaker may restrict the Amazon Mechanical Turk workers that can view your task based on this information.

Type: Array of strings

Array Members: Maximum number of 256 items.

Valid Values: FreeOfPersonallyIdentifiableInformation | FreeOfAdultContent

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**LabelingJobDataSource**

**Service:** Amazon SageMaker Service

Provides information about the location of input data.

You must specify at least one of the following: **S3DataSource** or **SnsDataSource**.

**S3DataSource**

The Amazon S3 location of the input data objects.

Type: **LabelingJobS3DataSource (p. 1133)** object

Required: No

**SnsDataSource**

An Amazon SNS data source used for streaming labeling jobs. To learn more, see Send Data to a Streaming Labeling Job.

Type: **LabelingJobSnsDataSource (p. 1134)** object

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobForWorkteamSummary
Service: Amazon SageMaker Service

Provides summary information for a work team.

Contents

CreationTime

The date and time that the labeling job was created.

Type: Timestamp

Required: Yes

JobReferenceCode

A unique identifier for a labeling job. You can use this to refer to a specific labeling job.

Type: String

Length Constraints: Minimum length of 1.

Pattern: .+

Required: Yes

LabelCounters

Provides information about the progress of a labeling job.

Type: LabelCountersForWorkteam (p. 1121) object

Required: No

LabelingJobName

The name of the labeling job that the work team is assigned to.

Type: String


Pattern: ^[a-zA-Z0-9-]*[a-zA-Z0-9]{0,62}$

Required: No

NumberOfHumanWorkersPerDataObject

The configured number of workers per data object.

Type: Integer


Required: No

WorkRequesterAccountId

The AWS account ID of the account used to start the labeling job.

Type: String

Pattern: ^\d+$
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobInputConfig
Service: Amazon SageMaker Service
Input configuration information for a labeling job.

Contents

DataAttributes
Attributes of the data specified by the customer.
Type: LabelingJobDataAttributes (p. 1124) object
Required: No

DataSource
The location of the input data.
Type: LabelingJobDataSource (p. 1125) object
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobOutput

Service: Amazon SageMaker Service

Specifies the location of the output produced by the labeling job.

Contents

**FinalActiveLearningModelArn**

The Amazon Resource Name (ARN) for the most recent SageMaker model trained as part of automated data labeling.

Type: String


Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model/.*`

Required: No

**OutputDatasetS3Uri**

The Amazon S3 bucket location of the manifest file for labeled data.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(https|s3)://([^/]+)/(.*$)`

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobOutputConfig
Service: Amazon SageMaker Service

Output configuration information for a labeling job.

Contents

KmsKeyId

The AWS Key Management Service ID of the key used to encrypt the output data, if any.

If you provide your own KMS key ID, you must add the required permissions to your KMS key described in Encrypt Output Data and Storage Volume with AWS KMS.

If you don’t provide a KMS key ID, Amazon SageMaker uses the default AWS KMS key for Amazon S3 for your role’s account to encrypt your output data.

If you use a bucket policy with an `s3:PutObject` permission that only allows objects with server-side encryption, set the condition key of `s3:x-amz-server-side-encryption` to "aws:kms". For more information, see KMS-Managed Encryption Keys in the Amazon Simple Storage Service Developer Guide.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

S3OutputPath

The Amazon S3 location to write output data.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3):/(/([^/]+)?([^\d]+))$?

Required: Yes

SnsTopicArn

An Amazon Simple Notification Service (Amazon SNS) output topic ARN. Provide a SnsTopicArn if you want to do real time chaining to another streaming job and receive an Amazon SNS notifications each time a data object is submitted by a worker.

If you provide an SnsTopicArn in OutputConfig, when workers complete labeling tasks, Ground Truth will send labeling task output data to the SNS output topic you specify here.

To learn more, see Receive Output Data from a Streaming Labeling Job.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sns:[a-z0-9\-]*:[0-9]{12}:[a-zA-Z0-9_.-]+*

Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobResourceConfig

Service: Amazon SageMaker Service

Configure encryption on the storage volume attached to the ML compute instance used to run automated data labeling model training and inference.

Contents

VolumeKmsKeyId

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt data on the storage volume attached to the ML compute instance(s) that run the training and inference jobs used for automated data labeling.

You can only specify a VolumeKmsKeyId when you create a labeling job with automated data labeling enabled using the API operation CreateLabelingJob. You cannot specify an AWS KMS key to encrypt the storage volume used for automated data labeling model training and inference when you create a labeling job using the console. To learn more, see Output Data and Storage Volume Encryption.

The VolumeKmsKeyId can be any of the following formats:

- KMS Key ID
  
  "1234abcd-12ab-34cd-56ef-1234567890ab"

- Amazon Resource Name (ARN) of a KMS Key

  "arn:aws:kms:us-west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobS3DataSource
Service: Amazon SageMaker Service

The Amazon S3 location of the input data objects.

Contents

ManifestS3Uri

The Amazon S3 location of the manifest file that describes the input data objects.

The input manifest file referenced in ManifestS3Uri must contain one of the following keys: source-ref or source. The value of the keys are interpreted as follows:
- source-ref: The source of the object is the Amazon S3 object specified in the value. Use this value when the object is a binary object, such as an image.
- source: The source of the object is the value. Use this value when the object is a text value.

If you are a new user of Ground Truth, it is recommended you review Use an Input Manifest File in the Amazon SageMaker Developer Guide to learn how to create an input manifest file.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3):/(/[^/]+/)?(.*$)

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**LabelingJobSnsDataSource**

Service: Amazon SageMaker Service

An Amazon SNS data source used for streaming labeling jobs.

**Contents**

**SnsTopicArn**

The Amazon SNS input topic Amazon Resource Name (ARN). Specify the ARN of the input topic you will use to send new data objects to a streaming labeling job.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sns:[a-z0-9\-]*:[0-9]{12}:a-zA-Z0-9_]+`

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Go](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)
LabelingJobStoppingConditions
Service: Amazon SageMaker Service

A set of conditions for stopping a labeling job. If any of the conditions are met, the job is automatically stopped. You can use these conditions to control the cost of data labeling.

**Note**
Labeling jobs fail after 30 days with an appropriate client error message.

**Contents**

MaxHumanLabeledObjectCount
The maximum number of objects that can be labeled by human workers.

Type: Integer
Valid Range: Minimum value of 1.
Required: No

MaxPercentageOfInputDatasetLabeled
The maximum number of input data objects that should be labeled.

Type: Integer
Valid Range: Minimum value of 1. Maximum value of 100.
Required: No

**See Also**
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LabelingJobSummary
Service: Amazon SageMaker Service
Provides summary information about a labeling job.

Contents

AnnotationConsolidationLambdaArn
The Amazon Resource Name (ARN) of the Lambda function used to consolidate the annotations from individual workers into a label for a data object. For more information, see Annotation Consolidation.
Type: String
Length Constraints: Maximum length of 2048.
Pattern: arn:aws[a-z\-]*:lambda:[a-z0-9\-]*:[0-9]{12}:function:.*
Required: No

CreationTime
The date and time that the job was created (timestamp).
Type: Timestamp
Required: Yes

FailureReason
If the LabelingJobStatus field is Failed, this field contains a description of the error.
Type: String
Length Constraints: Maximum length of 1024.
Required: No

InputConfig
Input configuration for the labeling job.
Type: LabelingJobInputConfig (p. 1128) object
Required: No

LabelCounters
Counts showing the progress of the labeling job.
Type: LabelCounters (p. 1119) object
Required: Yes

LabelingJobArn
The Amazon Resource Name (ARN) assigned to the labeling job when it was created.
Type: String
Length Constraints: Maximum length of 2048.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:labeling-job/.*
LabelingJobSummary

Required: Yes

LabelingJobName
The name of the labeling job.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$
Required: Yes

LabelingJobOutput
The location of the output produced by the labeling job.
Type: LabelingJobOutput (p. 1129) object
Required: No

LabelingJobStatus
The current status of the labeling job.
Type: String
Valid Values: Initializing | InProgress | Completed | Failed | Stopping | Stopped
Required: Yes

LastModifiedTime
The date and time that the job was last modified (timestamp).
Type: Timestamp
Required: Yes

PreHumanTaskLambdaArn
The Amazon Resource Name (ARN) of a Lambda function. The function is run before each data object is sent to a worker.
Type: String
Length Constraints: Maximum length of 2048.
Pattern: arn:aws[a-z\-]*:lambda:[a-z0-9\-]*:[0-9]{12}:function:.*
Required: Yes

WorkteamArn
The Amazon Resource Name (ARN) of the work team assigned to the job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workteam/.*
Required: Yes
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LambdaStepMetadata
Service: Amazon SageMaker Service

Metadata for a Lambda step.

Contents

Arn
The Amazon Resource Name (ARN) of the Lambda function that was run by this step execution.
Type: String
Length Constraints: Maximum length of 256.
Required: No

OutputParameters
A list of the output parameters of the Lambda step.
Type: Array of OutputParameter (p. 1246) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
LineageGroupSummary

Service: Amazon SageMaker Service

Lists a summary of the properties of a lineage group. A lineage group provides a group of shareable lineage entity resources.

Contents

CreationTime

The creation time of the lineage group summary.
Type: Timestamp
Required: No

DisplayName

The display name of the lineage group summary.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9]){0,119}$
Required: No

LastModifiedTime

The last modified time of the lineage group summary.
Type: Timestamp
Required: No

LineageGroupArn

The Amazon Resource Name (ARN) of the lineage group resource.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:lineage-group/.*
Required: No

LineageGroupName

The name or Amazon Resource Name (ARN) of the lineage group.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9]){0,119}$
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
MemberDefinition

Service: Amazon SageMaker Service

Defines an Amazon Cognito or your own OIDC IdP user group that is part of a work team.

Contents

CognitoMemberDefinition

The Amazon Cognito user group that is part of the work team.

Type: CognitoMemberDefinition (p. 951) object

Required: No

OidcMemberDefinition

A list user groups that exist in your OIDC Identity Provider (IdP). One to ten groups can be used to create a single private work team. When you add a user group to the list of Groups, you can add that user group to one or more private work teams. If you add a user group to a private work team, all workers in that user group are added to the work team.

Type: OidcMemberDefinition (p. 1237) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MetadataProperties
Service: Amazon SageMaker Service

Metadata properties of the tracking entity, trial, or trial component.

Contents

CommitId
The commit ID.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: . *
Required: No

GeneratedBy
The entity this entity was generated by.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: . *
Required: No

ProjectId
The project ID.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: . *
Required: No

Repository
The repository.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: . *
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
MetricData
Service: Amazon SageMaker Service

The name, value, and date and time of a metric that was emitted to Amazon CloudWatch.

Contents

MetricName
   The name of the metric.
   Type: String
   Length Constraints: Minimum length of 1. Maximum length of 255.
   Pattern: .+
   Required: No

Timestamp
   The date and time that the algorithm emitted the metric.
   Type: Timestamp
   Required: No

Value
   The value of the metric.
   Type: Float
   Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MetricDatum
Service: Amazon SageMaker Service
Information about the metric for a candidate produced by an AutoML job.

Contents

MetricName
The name of the metric.
Type: String
Valid Values: Accuracy | MSE | F1 | F1macro | AUC
Required: No

Set
The dataset split from which the AutoML job produced the metric.
Type: String
Valid Values: Train | Validation | Test
Required: No

StandardMetricName
The name of the standard metric.
Note
For definitions of the standard metrics, see Autopilot candidate metrics.
Type: String
Valid Values: Accuracy | MSE | F1 | F1macro | AUC | RMSE | MAE | R2 | BalancedAccuracy | Precision | PrecisionMacro | Recall | RecallMacro | LogLoss
Required: No

Value
The value of the metric.
Type: Float
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MetricDefinition
Service: Amazon SageMaker Service

Specifies a metric that the training algorithm writes to stderr or stdout. SageMaker hyperparameter tuning captures all defined metrics. You specify one metric that a hyperparameter tuning job uses as its objective metric to choose the best training job.

Contents

Name
The name of the metric.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 255.
Pattern: .+
Required: Yes

Regex
A regular expression that searches the output of a training job and gets the value of the metric. For more information about using regular expressions to define metrics, see Defining Objective Metrics.
Type: String
Pattern: .+
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MetricsSource
Service: Amazon SageMaker Service

Contents

ContentDigest

Type: String
Length Constraints: Maximum length of 72.
Pattern: ^[Ss][Hh][Aa]256:[0-9a-fA-F]{64}$
Required: No

ContentType

Type: String
Length Constraints: Maximum length of 256.
Pattern: .*
Required: Yes

S3Uri

Type: String
Length Constraints: Maximum length of 1024.
Pattern: ^(https|s3)://([^/]+)/(.*)$
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelArtifacts

Service: Amazon SageMaker Service

Provides information about the location that is configured for storing model artifacts.

Model artifacts are the output that results from training a model, and typically consist of trained parameters, a model definition that describes how to compute inferences, and other metadata.

Contents

S3ModelArtifacts

The path of the S3 object that contains the model artifacts. For example, s3://bucket-name/keynameprefix/model.tar.gz.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)/?(.*)$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelBiasAppSpecification

Service: Amazon SageMaker Service

Docker container image configuration object for the model bias job.

Contents

ConfigUri

JSON formatted S3 file that defines bias parameters. For more information on this JSON configuration file, see Configure bias parameters.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://(\[^/\]+)/(\^+)\?([^\s]*)$

Required: Yes

Environment

Sets the environment variables in the Docker container.

Type: String to string map

Map Entries: Maximum number of 50 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: [a-zA-Z_]\[a-zA-Z0-9_.]*

Value Length Constraints: Maximum length of 256.

Value Pattern: [\s\S]*

Required: No

ImageUri

The container image to be run by the model bias job.

Type: String

Length Constraints: Maximum length of 255.

Pattern: .*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelBiasBaselineConfig
Service: Amazon SageMaker Service
The configuration for a baseline model bias job.

Contents

BaselineJobName
The name of the baseline model bias job.
Type: String
Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]{0,62}
Required: No

ConstraintsResource
The constraints resource for a monitoring job.
Type: MonitoringConstraintsResource (p. 1198) object
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelBiasJobInput
Service: Amazon SageMaker Service
Inputs for the model bias job.

Contents

EndpointInput
  Input object for the endpoint
  Type: EndpointInput (p. 1017) object
  Required: Yes

GroundTruthS3Input
  Location of ground truth labels to use in model bias job.
  Type: MonitoringGroundTruthS3Input (p. 1201) object
  Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelClientConfig

Service: Amazon SageMaker Service

Configures the timeout and maximum number of retries for processing a transform job invocation.

Contents

**InvocationsMaxRetries**

The maximum number of retries when invocation requests are failing. The default value is 3.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 3.

Required: No

**InvocationsTimeoutInSeconds**

The timeout value in seconds for an invocation request. The default value is 600.

Type: Integer


Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelConfiguration
Service: Amazon SageMaker Service

Defines the model configuration. Includes the specification name and environment parameters.

Contents

EnvironmentParameters

Defines the environment parameters that includes key, value types, and values.

Type: Array of EnvironmentParameter (p. 1025) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Required: No

InferenceSpecificationName

The inference specification name in the model package version.

Type: String


Pattern: ^[a-zA-Z0-9][-][a-zA-Z0-9]{0,62}$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelDataQuality
Service: Amazon SageMaker Service
Data quality constraints and statistics for a model.

Contents

Constraints
Data quality constraints for a model.
Type: MetricsSource (p. 1148) object
Required: No

Statistics
Data quality statistics for a model.
Type: MetricsSource (p. 1148) object
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelDeployConfig
Service: Amazon SageMaker Service

Specifies how to generate the endpoint name for an automatic one-click Autopilot model deployment.

Contents

AutoGenerateEndpointName

Set to True to automatically generate an endpoint name for a one-click Autopilot model deployment; set to False otherwise. The default value is False.

Note
If you set AutoGenerateEndpointName to True, do not specify the EndpointName; otherwise a 400 error is thrown.

Type: Boolean

Required: No

EndpointName

Specifies the endpoint name to use for a one-click Autopilot model deployment if the endpoint name is not generated automatically.

Note
Specify the EndpointName if and only if you set AutoGenerateEndpointName to False; otherwise a 400 error is thrown.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62\}

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelDeployResult

Service: Amazon SageMaker Service

Provides information about the endpoint of the model deployment.

Contents

EndpointName

The name of the endpoint to which the model has been deployed.

Note
If model deployment fails, this field is omitted from the response.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelDigests
Service: Amazon SageMaker Service
Provides information to verify the integrity of stored model artifacts.

Contents
ArtifactDigest
Provides a hash value that uniquely identifies the stored model artifacts.
Type: String
Pattern: ^[\w: ]*$
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelExplainabilityAppSpecification

Service: Amazon SageMaker Service

Docker container image configuration object for the model explainability job.

Contents

ConfigUri

JSON formatted S3 file that defines explainability parameters. For more information on this JSON configuration file, see Configure model explainability parameters.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://(\[^/\]+)/?([^/]+)(\?.*)$

Required: Yes

Environment

Sets the environment variables in the Docker container.

Type: String to string map

Map Entries: Maximum number of 50 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: [a-zA-Z-_][a-zA-Z0-9-_]*

Value Length Constraints: Maximum length of 256.

Value Pattern: [\\s\s]*

Required: No

ImageUri

The container image to be run by the model explainability job.

Type: String

Length Constraints: Maximum length of 255.

Pattern: .*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelExplainabilityBaselineConfig
Service: Amazon SageMaker Service

The configuration for a baseline model explainability job.

Contents

BaseliningJobName

The name of the baseline model explainability job.

Type: String


Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9])\{0,62\}

Required: No

ConstraintsResource

The constraints resource for a monitoring job.

Type: MonitoringConstraintsResource (p. 1198) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelExplainabilityJobInput

Service: Amazon SageMaker Service

Inputs for the model explainability job.

Contents

EndpointInput

Input object for the endpoint

Type: EndpointInput (p. 1017) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelInput

Service: Amazon SageMaker Service

Input object for the model.

Contents

DataInputConfig

The input configuration object for the model.

Type: String


Pattern: \S+\n
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelLatencyThreshold

Service: Amazon SageMaker Service

The model latency threshold.

Contents

Percentile

The model latency percentile threshold.

Type: String

Length Constraints: Maximum length of 64.

Required: No

ValueInMilliseconds

The model latency percentile value in milliseconds.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelMetadataFilter
Service: Amazon SageMaker Service

Part of the search expression. You can specify the name and value (domain, task, framework, framework version, task, and model).

Contents

Name
The name of the model to filter by.
Type: String
Valid Values: Domain | Framework | Task | FrameworkVersion
Required: Yes

Value
The value to filter the model metadata.
Type: String
Length Constraints: Maximum length of 256.
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelMetadataSearchExpression
Service: Amazon SageMaker Service
One or more filters that searches for the specified resource or resources in a search. All resource objects that satisfy the expression's condition are included in the search results

Contents

Filters
A list of filter objects.
Type: Array of ModelMetadataFilter (p. 1164) objects
Array Members: Minimum number of 1 item. Maximum number of 4 items.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelMetadataSummary
Service: Amazon SageMaker Service
A summary of the model metadata.

Contents

Domain
The machine learning domain of the model.
Type: String
Required: Yes

Framework
The machine learning framework of the model.
Type: String
Required: Yes

FrameworkVersion
The framework version of the model.
Type: String
Required: Yes

Model
The name of the model.
Type: String
Required: Yes

Task
The machine learning task of the model.
Type: String
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelMetrics

Service: Amazon SageMaker Service

Contains metrics captured from a model.

Contents

Bias

Metrics that measure bias in a model.

Type: Bias (p. 929) object

Required: No

Explainability

Metrics that help explain a model.

Type: Explainability (p. 1034) object

Required: No

ModelDataQuality

Metrics that measure the quality of the input data for a model.

Type: ModelDataQuality (p. 1155) object

Required: No

ModelQuality

Metrics that measure the quality of a model.

Type: ModelQuality (p. 1186) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackage
Service: Amazon SageMaker Service
A versioned model that can be deployed for SageMaker inference.

Contents

AdditionalInferenceSpecifications
An array of additional Inference Specification objects.
Type: Array of AdditionalInferenceSpecificationDefinition (p. 863) objects
Array Members: Minimum number of 1 item. Maximum number of 15 items.
Required: No

ApprovalDescription
A description provided when the model approval is set.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: . *
Required: No

CertifyForMarketplace
Whether the model package is to be certified to be listed on AWS Marketplace. For information about listing model packages on AWS Marketplace, see List Your Algorithm or Model Package on AWS Marketplace.
Type: Boolean
Required: No

CreatedBy
Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
Type: UserContext (p. 1449) object
Required: No

CreationTime
The time that the model package was created.
Type: Timestamp
Required: No

CustomerMetadataProperties
The metadata properties for the model package.
Type: String to string map
Map Entries: Maximum number of 50 items.
Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: ^([\p{L}\p{Z}\p{N}_.:/=+-@]*)${1,128}

Value Length Constraints: Minimum length of 1. Maximum length of 256.

Value Pattern: ^([\p{L}\p{Z}\p{N}_.:/=+-@]*)${1,256}

Required: No

**Domain**

The machine learning domain of your model package and its components. Common machine learning domains include computer vision and natural language processing.

Type: String

Required: No

**DriftCheckBaselines**

Represents the drift check baselines that can be used when the model monitor is set using the model package.

Type: DriftCheckBaselines (p. 998) object

Required: No

**InferenceSpecification**

Defines how to perform inference generation after a training job is run.

Type: InferenceSpecification (p. 1105) object

Required: No

**LastModifiedBy**

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

Required: No

**LastModifiedTime**

The last time the model package was modified.

Type: Timestamp

Required: No

**MetadataProperties**

Metadata properties of the tracking entity, trial, or trial component.

Type: MetadataProperties (p. 1143) object

Required: No

**ModelApprovalStatus**

The approval status of the model. This can be one of the following values.

- APPROVED - The model is approved
- REJECTED - The model is rejected
• `PENDING_MANUAL_APPROVAL` - The model is waiting for manual approval.

  Type: String

  Valid Values: Approved | Rejected | PendingManualApproval

  Required: No

**ModelMetrics**

  Metrics for the model.

  Type: `ModelMetrics` (p. 1167) object

  Required: No

**ModelPackageArn**

  The Amazon Resource Name (ARN) of the model package.

  Type: String

  Length Constraints: Minimum length of 1. Maximum length of 2048.

  Pattern: `arn:aws[\da-z\-]*:sagemaker:[:a-z0-9\-]*:\[0-9\]{12}:model-package/.*`

  Required: No

**ModelPackageDescription**

  The description of the model package.

  Type: String

  Length Constraints: Maximum length of 1024.

  Pattern: `[^\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*`

  Required: No

**ModelPackageGroupName**

  The model group to which the model belongs.

  Type: String


  Pattern: `^[a-zA-Z0-9\-\]+(\-[a-zA-Z0-9\-]+\{0,62}\$`

  Required: No

**ModelPackageName**

  The name of the model.

  Type: String


  Pattern: `^[a-zA-Z0-9\-\]+(\-[a-zA-Z0-9\-]+\{0,62}\$`

  Required: No

**ModelPackageStatus**

  The status of the model package. This can be one of the following values.
ModelPackage

- **PENDING** - The model package is pending being created.
- **IN_PROGRESS** - The model package is in the process of being created.
- **COMPLETED** - The model package was successfully created.
- **FAILED** - The model package failed.
- **DELETING** - The model package is in the process of being deleted.

Type: String

Valid Values: Pending | InProgress | Completed | Failed | Deleting

Required: No

**ModelPackageStatusDetails**

Specifies the validation and image scan statuses of the model package.

Type: ModelPackageStatusDetails (p. 1180) object

Required: No

**ModelPackageVersion**

The version number of a versioned model.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

**SamplePayloadUrl**

The Amazon Simple Storage Service path where the sample payload are stored. This path must point to a single gzip compressed tar archive (.tar.gz suffix).

Type: String

Required: No

**SourceAlgorithmSpecification**

A list of algorithms that were used to create a model package.

Type: SourceAlgorithmSpecification (p. 1368) object

Required: No

**Tags**

A list of the tags associated with the model package. For more information, see Tagging AWS resources in the AWS General Reference Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**Task**

The machine learning task your model package accomplishes. Common machine learning tasks include object detection and image classification.

Type: String
Required: No

**ValidationSpecification**

Specifies batch transform jobs that SageMaker runs to validate your model package.

Type: `ModelPackageValidationSpecification (p. 1185)` object

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackageContainerDefinition

Service: Amazon SageMaker Service

Describes the Docker container for the model package.

Contents

ContainerHostname

The DNS host name for the Docker container.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]{0,62}){0,62}

Required: No

Environment

The environment variables to set in the Docker container. Each key and value in the Environment string to string map can have length of up to 1024. We support up to 16 entries in the map.

Type: String to string map

Map Entries: Maximum number of 16 items.

Key Length Constraints: Maximum length of 1024.

Key Pattern: [a-zA-Z_][a-zA-Z0-9_]*

Value Length Constraints: Maximum length of 1024.

Value Pattern: [\S\s]*

Required: No

Framework

The machine learning framework of the model package container image.

Type: String

Required: No

FrameworkVersion

The framework version of the Model Package Container image.

Type: String


Pattern: [0-9]\.[A-Za-z0-9.]*

Required: No

Image

The Amazon EC2 Container Registry (Amazon ECR) path where inference code is stored.

If you are using your own custom algorithm instead of an algorithm provided by SageMaker, the inference code must meet SageMaker requirements. SageMaker supports both registry/
ModelPackageContainerDefinition

repository[:tag] and registry/repository[@digest] image path formats. For more information, see Using Your Own Algorithms with Amazon SageMaker.

Type: String
Length Constraints: Maximum length of 255.
Pattern: \[\S]+
Required: Yes

ImageDigest
An MD5 hash of the training algorithm that identifies the Docker image used for training.

Type: String
Length Constraints: Maximum length of 72.
Pattern: ^[Ss][Hh][Aa]256:0-9a-fA-F]{64}$
Required: No

ModelDataUrl
The Amazon S3 path where the model artifacts, which result from model training, are stored. This path must point to a single gzip compressed tar archive (.tar.gz suffix).

Note
The model artifacts must be in an S3 bucket that is in the same region as the model package.

Type: String
Length Constraints: Maximum length of 1024.
Pattern: ^(https|s3://[^/]+)/(.*))$  
Required: No

ModelInput
A structure with Model Input details.

Type: ModelInput (p. 1162) object
Required: No

NearestModelName
The name of a pre-trained machine learning benchmarked by Amazon SageMaker Inference Recommender model that matches your model. You can find a list of benchmarked models by calling ListModelMetadata.

Type: String
Required: No

ProductId
The AWS Marketplace product ID of the model package.

Type: String
Length Constraints: Maximum length of 256.
Pattern: ^[a-zA-Z0-9](\-[a-zA-Z0-9])*\$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackageGroup
Service: Amazon SageMaker Service
A group of versioned models in the model registry.

Contents

CreatedBy
Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
Type: UserContext (p. 1449) object
Required: No

CreationTime
The time that the model group was created.
Type: Timestamp
Required: No

ModelPackageGroupArn
The Amazon Resource Name (ARN) of the model group.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 2048.
Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package-group/.*`
Required: No

ModelPackageGroupDescription
The description for the model group.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: `[^\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*`
Required: No

ModelPackageGroupName
The name of the model group.
Type: String
Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62\}$`
Required: No

ModelPackageGroupStatus
The status of the model group. This can be one of the following values.
• **PENDING** - The model group is pending being created.
• **IN_PROGRESS** - The model group is in the process of being created.
• **COMPLETED** - The model group was successfully created.
• **FAILED** - The model group failed.
• **DELETING** - The model group is in the process of being deleted.
• **DELETE_FAILED** - SageMaker failed to delete the model group.

Type: String

**Valid Values:** Pending | InProgress | Completed | Failed | Deleting | DeleteFailed

Required: No

**Tags**

A list of the tags associated with the model group. For more information, see [Tagging AWS resources](https://docs.aws.amazon.com/AmazonS3/latest/userguide/tagging.html) in the *AWS General Reference Guide*.

Type: Array of Tag objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackageGroupSummary

Service: Amazon SageMaker Service

Summary information about a model group.

Contents

CreationTime

The time that the model group was created.

Type: Timestamp

Required: Yes

ModelPackageGroupArn

The Amazon Resource Name (ARN) of the model group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: \[a-zA-Z0-9\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]\{12\}:model-package-group/.*

Required: Yes

ModelPackageGroupDescription

A description of the model group.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*

Required: No

ModelPackageGroupName

The name of the model group.

Type: String


Pattern: ^[a-zA-Z0-9\-](\-[a-zA-Z0-9\-])\{0,62}\$

Required: Yes

ModelPackageGroupStatus

The status of the model group.

Type: String

Valid Values: Pending | InProgress | Completed | Failed | Deleting | DeleteFailed

Required: Yes
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackageStatusDetails
Service: Amazon SageMaker Service

Specifies the validation and image scan statuses of the model package.

Contents

ImageScanStatuses
The status of the scan of the Docker image container for the model package.
Type: Array of ModelPackageStatusItem (p. 1181) objects
Required: No

ValidationStatuses
The validation status of the model package.
Type: Array of ModelPackageStatusItem (p. 1181) objects
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackageStatusItem

Service: Amazon SageMaker Service

Represents the overall status of a model package.

Contents

FailureReason

if the overall status is Failed, the reason for the failure.

Type: String

Required: No

Name

The name of the model package for which the overall status is being reported.

Type: String


Pattern: ^[a-zA-Z0-9][-]*[a-zA-Z0-9]{0,62}$

Required: Yes

Status

The current status.

Type: String

Valid Values: NotStarted | InProgress | Completed | Failed

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackageSummary
Service: Amazon SageMaker Service
Provides summary information about a model package.

Contents

CreationTime
A timestamp that shows when the model package was created.
Type: Timestamp
Required: Yes

ModelApprovalStatus
The approval status of the model. This can be one of the following values.
- APPROVED - The model is approved
- REJECTED - The model is rejected.
- PENDING_MANUAL_APPROVAL - The model is waiting for manual approval.
Type: String
Valid Values: Approved | Rejected | PendingManualApproval
Required: No

ModelPackageArn
The Amazon Resource Name (ARN) of the model package.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 2048.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package/.*
Required: Yes

ModelPackageDescription
A brief description of the model package.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*
Required: No

ModelPackageGroupName
If the model package is a versioned model, the model group that the versioned model belongs to.
Type: String
Pattern: ^[a-zA-Z0-9-]*(\-[a-zA-Z0-9]+)?[0,62]$
Required: No
ModelPackageName

The name of the model package.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

ModelPackageStatus

The overall status of the model package.

Type: String

Valid Values: Pending | InProgress | Completed | Failed | Deleting

Required: Yes

ModelPackageVersion

If the model package is a versioned model, the version of the model.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackageValidationProfile

Service: Amazon SageMaker Service

Contains data, such as the inputs and targeted instance types that are used in the process of validating the model package.

The data provided in the validation profile is made available to your buyers on AWS Marketplace.

Contents

ProfileName

The name of the profile for the model package.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: Yes

TransformJobDefinition

The TransformJobDefinition object that describes the transform job used for the validation of the model package.

Type: TransformJobDefinition (p. 1410) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelPackageValidationSpecification
Service: Amazon SageMaker Service

Specifies batch transform jobs that SageMaker runs to validate your model package.

Contents

ValidationProfiles

An array of ModelPackageValidationProfile objects, each of which specifies a batch transform job that SageMaker runs to validate your model package.

Type: Array of ModelPackageValidationProfile (p. 1184) objects

Array Members: Fixed number of 1 item.

Required: Yes

ValidationRole

The IAM roles to be used for the validation of the model package.

Type: String


Pattern: ^arn:aws[\w\-]*:iam::\d{12}:role/?[a-zA-Z0-9+=,.@\-_]+$ 

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelQuality
Service: Amazon SageMaker Service

Model quality statistics and constraints.

Contents

Constraints

Model quality constraints.

Type: MetricsSource (p. 1148) object

Required: No

Statistics

Model quality statistics.

Type: MetricsSource (p. 1148) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelQualityAppSpecification

Service: Amazon SageMaker Service

Container image configuration object for the monitoring job.

Contents

**ContainerArguments**

An array of arguments for the container used to run the monitoring job.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 50 items.

Length Constraints: Maximum length of 256.

Pattern: . *

Required: No

**ContainerEntrypoint**

Specifies the entrypoint for a container that the monitoring job runs.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Maximum length of 256.

Pattern: . *

Required: No

**Environment**

Sets the environment variables in the container that the monitoring job runs.

Type: String to string map

Map Entries: Maximum number of 50 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: [a-zA-Z_][a-zA-Z0-9_]*

Value Length Constraints: Maximum length of 256.

Value Pattern: \S\s*

Required: No

**ImageUri**

The address of the container image that the monitoring job runs.

Type: String

Length Constraints: Maximum length of 255.

Pattern: . *
Required: Yes

**PostAnalyticsProcessorSourceUri**

An Amazon S3 URI to a script that is called after analysis has been performed. Applicable only for the built-in (first party) containers.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(https|s3)://([^/]+)/?([^/]+$)

Required: No

**ProblemType**

The machine learning problem type of the model that the monitoring job monitors.

Type: String

Valid Values: BinaryClassification | MulticlassClassification | Regression

Required: No

**RecordPreprocessorSourceUri**

An Amazon S3 URI to a script that is called per row prior to running analysis. It can base64 decode the payload and convert it into a flatted json so that the built-in container can use the converted data. Applicable only for the built-in (first party) containers.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(https|s3)://([^/]+)/?([^/]+$)

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelQualityBaselineConfig
Service: Amazon SageMaker Service

Configuration for monitoring constraints and monitoring statistics. These baseline resources are compared against the results of the current job from the series of jobs scheduled to collect data periodically.

Contents

BaseliningJobName
The name of the job that performs baselining for the monitoring job.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$
Required: No

ConstraintsResource
The constraints resource for a monitoring job.
Type: MonitoringConstraintsResource (p. 1198) object
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelQualityJobInput

Service: Amazon SageMaker Service

The input for the model quality monitoring job. Currently endpoints are supported for input for model quality monitoring jobs.

Contents

EndpointInput

Input object for the endpoint

Type: EndpointInput (p. 1017) object

Required: Yes

GroundTruthS3Input

The ground truth label provided for the model.

Type: MonitoringGroundTruthS3Input (p. 1201) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelStepMetadata

Service: Amazon SageMaker Service

Metadata for Model steps.

Contents

Arn

The Amazon Resource Name (ARN) of the created model.

Type: String

Length Constraints: Maximum length of 256.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ModelSummary
Service: Amazon SageMaker Service

Provides summary information about a model.

Contents

CreationTime
A timestamp that indicates when the model was created.
Type: Timestamp
Required: Yes

ModelArn
The Amazon Resource Name (ARN) of the model.
Type: String
Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model/.*`
Required: Yes

ModelName
The name of the model that you want a summary for.
Type: String
Length Constraints: Maximum length of 63.
Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*`
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringAppSpecification

Service: Amazon SageMaker Service

Container image configuration object for the monitoring job.

Contents

**ContainerArguments**

An array of arguments for the container used to run the monitoring job.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 50 items.

Length Constraints: Maximum length of 256.

Pattern: . *

Required: No

**ContainerEntrypoint**

Specifies the entrypoint for a container used to run the monitoring job.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Maximum length of 256.

Pattern: . *

Required: No

**ImageUri**

The container image to be run by the monitoring job.

Type: String

Length Constraints: Maximum length of 255.

Pattern: . *

Required: Yes

**PostAnalyticsProcessorSourceUri**

An Amazon S3 URI to a script that is called after analysis has been performed. Applicable only for the built-in (first party) containers.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)/(.*$)

Required: No

**RecordPreprocessorSourceUri**

An Amazon S3 URI to a script that is called per row prior to running analysis. It can base64 decode the payload and convert it into a flatted json so that the built-in container can use the converted data. Applicable only for the built-in (first party) containers.
Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^((https|s3):/(/[^/]+)?(/.*)$)

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringBaselineConfig

Service: Amazon SageMaker Service

Configuration for monitoring constraints and monitoring statistics. These baseline resources are compared against the results of the current job from the series of jobs scheduled to collect data periodically.

Contents

**BaseliningJobName**

The name of the job that performs baselining for the monitoring job.

Type: String


Pattern: `^[a-zA-Z0-9](?=.*[a-zA-Z0-9])\{0,62\}`

Required: No

**ConstraintsResource**

The baseline constraint file in Amazon S3 that the current monitoring job should validated against.

Type: MonitoringConstraintsResource (p. 1198) object

Required: No

**StatisticsResource**

The baseline statistics file in Amazon S3 that the current monitoring job should be validated against.

Type: MonitoringStatisticsResource (p. 1218) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**MonitoringClusterConfig**

Service: Amazon SageMaker Service

Configuration for the cluster used to run model monitoring jobs.

**Contents**

**InstanceCount**

The number of ML compute instances to use in the model monitoring job. For distributed processing jobs, specify a value greater than 1. The default value is 1.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: Yes

**InstanceType**

The ML compute instance type for the processing job.

Type: String

Valid Values: ml.t3.medium | ml.t3.large | ml.t3.xlarge | ml.t3.2xlarge |
| ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge |
| ml.m4.16xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge |
| ml.c4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge |
| ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.c5.xlarge |
| ml.c5.2xlarge | ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge |
| ml.m5.large | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge |
| ml.m5.12xlarge | ml.m5.24xlarge | ml.r5.large | ml.r5.xlarge |
| ml.r5.2xlarge | ml.r5.4xlarge | ml.r5.8xlarge | ml.r5.12xlarge |
| ml.r5.16xlarge | ml.r5.24xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge |
| ml.g4dn.4xlarge | ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge |

Required: Yes

**VolumeKmsKeyId**

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt data on the storage volume attached to the ML compute instance(s) that run the model monitoring job.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

**VolumeSizeInGB**

The size of the ML storage volume, in gigabytes, that you want to provision. You must specify sufficient ML storage for your scenario.

Type: Integer


Required: Yes
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringConstraintsResource

Service: Amazon SageMaker Service

The constraints resource for a monitoring job.

Contents

S3Uri

The Amazon S3 URI for the constraints resource.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)/(.*)$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringExecutionSummary

Service: Amazon SageMaker Service

Summary of information about the last monitoring job to run.

Contents

CreationTime

The time at which the monitoring job was created.

Type: Timestamp

Required: Yes

EndpointName

The name of the endpoint used to run the monitoring job.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}$

Required: No

FailureReason

Contains the reason a monitoring job failed, if it failed.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

LastModifiedTime

A timestamp that indicates the last time the monitoring job was modified.

Type: Timestamp

Required: Yes

MonitoringExecutionStatus

The status of the monitoring job.

Type: String

Valid Values: Pending | Completed | CompletedWithViolations | InProgress | Failed | Stopping | Stopped

Required: Yes

MonitoringJobDefinitionName

The name of the monitoring job.

Type: String

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$

Required: No

**MonitoringScheduleName**

The name of the monitoring schedule.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\$

Required: Yes

**MonitoringType**

The type of the monitoring job.

Type: String

Valid Values: DataQuality | ModelQuality | ModelBias | ModelExplainability

Required: No

**ProcessingJobArn**

The Amazon Resource Name (ARN) of the monitoring job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]\{12\}:processing-job/.*

Required: No

**ScheduledTime**

The time the monitoring job was scheduled.

Type: Timestamp

Required: Yes

---

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringGroundTruthS3Input
Service: Amazon SageMaker Service

The ground truth labels for the dataset used for the monitoring job.

Contents

S3Uri

The address of the Amazon S3 location of the ground truth labels.

Type: String

Length Constraints: Maximum length of 512.

Pattern: ^(https|s3)://([^/]+)/?(.*)$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringInput
Service: Amazon SageMaker Service

The inputs for a monitoring job.

Contents

EndpointInput

  The endpoint for a monitoring job.

  Type: EndpointInput (p. 1017) object

  Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringJobDefinition
Service: Amazon SageMaker Service

Defines the monitoring job.

Contents

BaselineConfig
Baseline configuration used to validate that the data conforms to the specified constraints and statistics

Type: MonitoringBaselineConfig (p. 1195) object

Required: No

Environment
Sets the environment variables in the Docker container.

Type: String to string map

Map Entries: Maximum number of 50 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: [a-zA-Z_][a-zA-Z0-9_]*

Value Length Constraints: Maximum length of 256.

Value Pattern: [\S\s]*

Required: No

MonitoringAppSpecification
Configures the monitoring job to run a specified Docker container image.

Type: MonitoringAppSpecification (p. 1193) object

Required: Yes

MonitoringInputs
The array of inputs for the monitoring job. Currently we support monitoring an Amazon SageMaker Endpoint.

Type: Array of MonitoringInput (p. 1202) objects

Array Members: Fixed number of 1 item.

Required: Yes

MonitoringOutputConfig
The array of outputs from the monitoring job to be uploaded to Amazon Simple Storage Service (Amazon S3).

Type: MonitoringOutputConfig (p. 1209) object

Required: Yes
MonitoringResources

Identifies the resources, ML compute instances, and ML storage volumes to deploy for a monitoring job. In distributed processing, you specify more than one instance.

Type: MonitoringResources (p. 1210) object

Required: Yes

NetworkConfig

Specifies networking options for a monitoring job.

Type: NetworkConfig (p. 1223) object

Required: No

RoleArn

The Amazon Resource Name (ARN) of an IAM role that Amazon SageMaker can assume to perform tasks on your behalf.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_/]+$ 

Required: Yes

StoppingCondition

Specifies a time limit for how long the monitoring job is allowed to run.

Type: MonitoringStoppingCondition (p. 1219) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringJobDefinitionSummary
Service: Amazon SageMaker Service

Summary information about a monitoring job.

Contents

CreationTime
The time that the monitoring job was created.
Type: Timestamp
Required: Yes

EndpointName
The name of the endpoint that the job monitors.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\$
Required: Yes

MonitoringJobDefinitionArn
The Amazon Resource Name (ARN) of the monitoring job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: .*
Required: Yes

MonitoringJobDefinitionName
The name of the monitoring job.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])**
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringNetworkConfig

Service: Amazon SageMaker Service

The networking configuration for the monitoring job.

Contents

EnableInterContainerTrafficEncryption

Whether to encrypt all communications between the instances used for the monitoring jobs. Choose True to encrypt communications. Encryption provides greater security for distributed jobs, but the processing might take longer.

Type: Boolean

Required: No

EnableNetworkIsolation

Whether to allow inbound and outbound network calls to and from the containers used for the monitoring job.

Type: Boolean

Required: No

VpcConfig

Specifies a VPC that your training jobs and hosted models have access to. Control access to and from your training and model containers by configuring the VPC. For more information, see Protect Endpoints by Using an Amazon Virtual Private Cloud and Protect Training Jobs by Using an Amazon Virtual Private Cloud.

Type: VpcConfig (p. 1456) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringOutput

Service: Amazon SageMaker Service

The output object for a monitoring job.

Contents

S3Output

The Amazon S3 storage location where the results of a monitoring job are saved.

Type: MonitoringS3Output (p. 1211) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringOutputConfig
Service: Amazon SageMaker Service
The output configuration for monitoring jobs.

Contents

KmsKeyId
The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt the model artifacts at rest using Amazon S3 server-side encryption.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: . *

Required: No

MonitoringOutputs
Monitoring outputs for monitoring jobs. This is where the output of the periodic monitoring jobs is uploaded.

Type: Array of MonitoringOutput (p. 1208) objects

Array Members: Fixed number of 1 item.

Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringResources
Service: Amazon SageMaker Service

Identifies the resources to deploy for a monitoring job.

Contents

ClusterConfig

The configuration for the cluster resources used to run the processing job.

Type: MonitoringClusterConfig (p. 1196) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringS3Output
Service: Amazon SageMaker Service

Information about where and how you want to store the results of a monitoring job.

Contents

LocalPath

The local path to the Amazon S3 storage location where Amazon SageMaker saves the results of a monitoring job. LocalPath is an absolute path for the output data.

Type: String

Length Constraints: Maximum length of 256.

Pattern: . *

Required: Yes

S3UploadMode

Whether to upload the results of the monitoring job continuously or after the job completes.

Type: String

Valid Values: Continuous | EndOfJob

Required: No

S3Uri

A URI that identifies the Amazon S3 storage location where Amazon SageMaker saves the results of a monitoring job.

Type: String

Length Constraints: Maximum length of 512.

Pattern: ^(https|s3)://([^/]+)/?(.*$)

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringSchedule
Service: Amazon SageMaker Service

A schedule for a model monitoring job. For information about model monitor, see Amazon SageMaker Model Monitor.

Contents

CreationTime
The time that the monitoring schedule was created.
Type: Timestamp
Required: No

EndpointName
The endpoint that hosts the model being monitored.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}
Required: No

FailureReason
If the monitoring schedule failed, the reason it failed.
Type: String
Length Constraints: Maximum length of 1024.
Required: No

LastModifiedTime
The last time the monitoring schedule was changed.
Type: Timestamp
Required: No

LastMonitoringExecutionSummary
Summary of information about the last monitoring job to run.
Type: MonitoringExecutionSummary (p. 1199) object
Required: No

MonitoringScheduleArn
The Amazon Resource Name (ARN) of the monitoring schedule.
Type: String
Length Constraints: Maximum length of 256.
Pattern: .*
**MonitoringSchedule**

Required: No

**MonitoringScheduleConfig**

Configures the monitoring schedule and defines the monitoring job.

Type: MonitoringScheduleConfig (p. 1215) object

Required: No

**MonitoringScheduleName**

The name of the monitoring schedule.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\$`

Required: No

**MonitoringScheduleStatus**

The status of the monitoring schedule. This can be one of the following values.

- **PENDING** - The schedule is pending being created.
- **FAILED** - The schedule failed.
- **SCHEDULED** - The schedule was successfully created.
- **STOPPED** - The schedule was stopped.

Type: String

Valid Values: Pending | Failed | Scheduled | Stopped

Required: No

**MonitoringType**

The type of the monitoring job definition to schedule.

Type: String

Valid Values: DataQuality | ModelQuality | ModelBias | ModelExplainability

Required: No

**Tags**

A list of the tags associated with the monitoring schedule. For more information, see Tagging AWS resources in the AWS General Reference Guide.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringScheduleConfig
Service: Amazon SageMaker Service
Configures the monitoring schedule and defines the monitoring job.

Contents

MonitoringJobDefinition
  Defines the monitoring job.
  Type: MonitoringJobDefinition (p. 1203) object
  Required: No

MonitoringJobDefinitionName
  The name of the monitoring job definition to schedule.
  Type: String
  Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$
  Required: No

MonitoringType
  The type of the monitoring job definition to schedule.
  Type: String
  Valid Values: DataQuality | ModelQuality | ModelBias | ModelExplainability
  Required: No

ScheduleConfig
  Configures the monitoring schedule.
  Type: ScheduleConfig (p. 1354) object
  Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringScheduleSummary
Service: Amazon SageMaker Service
Summarizes the monitoring schedule.

Contents

**CreationTime**

The creation time of the monitoring schedule.

Type: Timestamp

Required: Yes

**EndpointName**

The name of the endpoint using the monitoring schedule.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$

Required: No

**LastModifiedTime**

The last time the monitoring schedule was modified.

Type: Timestamp

Required: Yes

**MonitoringJobDefinitionName**

The name of the monitoring job definition that the schedule is for.

Type: String


Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])*$

Required: No

**MonitoringScheduleArn**

The Amazon Resource Name (ARN) of the monitoring schedule.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: Yes

**MonitoringScheduleName**

The name of the monitoring schedule.

Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$
Required: Yes

**MonitoringScheduleStatus**

The status of the monitoring schedule.
Type: String
Valid Values: Pending | Failed | Scheduled | Stopped
Required: Yes

**MonitoringType**

The type of the monitoring job definition that the schedule is for.
Type: String
Valid Values: DataQuality | ModelQuality | ModelBias | ModelExplainability
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringStatisticsResource

Service: Amazon SageMaker Service

The statistics resource for a monitoring job.

Contents

S3Uri

The Amazon S3 URI for the statistics resource.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)//([^/]+)([^/]+)$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
MonitoringStoppingCondition

Service: Amazon SageMaker Service

A time limit for how long the monitoring job is allowed to run before stopping.

Contents

MaxRuntimeInSeconds

The maximum runtime allowed in seconds.

Note

The MaxRuntimeInSeconds cannot exceed the frequency of the job. For data quality and model explainability, this can be up to 3600 seconds for an hourly schedule. For model bias and model quality hourly schedules, this can be up to 1800 seconds.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 86400.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**MultiModelConfig**

Service: Amazon SageMaker Service

Specifies additional configuration for hosting multi-model endpoints.

**Contents**

**ModelCacheSetting**

Whether to cache models for a multi-model endpoint. By default, multi-model endpoints cache models so that a model does not have to be loaded into memory each time it is invoked. Some use cases do not benefit from model caching. For example, if an endpoint hosts a large number of models that are each invoked infrequently, the endpoint might perform better if you disable model caching. To disable model caching, set the value of this parameter to `Disabled`.

Type: String

Valid Values: `Enabled` | `Disabled`

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
NeoVpcConfig

Service: Amazon SageMaker Service

The VpcConfig (p. 1456) configuration object that specifies the VPC that you want the compilation jobs to connect to. For more information on controlling access to your Amazon S3 buckets used for compilation job, see Give Amazon SageMaker Compilation Jobs Access to Resources in Your Amazon VPC.

Contents

SecurityGroupIds

The VPC security group IDs. IDs have the form of sg-xxxxxxxxxx. Specify the security groups for the VPC that is specified in the Subnets field.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Maximum length of 32.

Pattern: [-0-9a-zA-Z]+

Required: Yes

Subnets

The ID of the subnets in the VPC that you want to connect the compilation job to for accessing the model in Amazon S3.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 16 items.

Length Constraints: Maximum length of 32.

Pattern: [-0-9a-zA-Z]+

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**NestedFilters**

Service: Amazon SageMaker Service

A list of nested Filter (p. 1046) objects. A resource must satisfy the conditions of all filters to be included in the results returned from the Search (p. 691) API.

For example, to filter on a training job's InputDataConfig property with a specific channel name and S3Uri prefix, define the following filters:

- '{Name:"InputDataConfig.ChannelName", "Operator":"Equals", "Value":"train"}',
- '{Name:"InputDataConfig.DataSource.S3DataSource.S3Uri", "Operator":"Contains", "Value":"mybucket/catdata"}'

**Contents**

**Filters**

A list of filters. Each filter acts on a property. Filters must contain at least one Filters value. For example, a NestedFilters call might include a filter on the PropertyName parameter of the InputDataConfig property: InputDataConfig.DataSource.S3DataSource.S3Uri.

Type: Array of Filter (p. 1046) objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Required: Yes

**NestedPropertyName**

The name of the property to use in the nested filters. The value must match a listed property name, such as InputDataConfig.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .+

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
NetworkConfig
Service: Amazon SageMaker Service

Networking options for a job, such as network traffic encryption between containers, whether to allow inbound and outbound network calls to and from containers, and the VPC subnets and security groups to use for VPC-enabled jobs.

Contents

EnableInterContainerTrafficEncryption
Whether to encrypt all communications between distributed processing jobs. Choose True to encrypt communications. Encryption provides greater security for distributed processing jobs, but the processing might take longer.

Type: Boolean
Required: No

EnableNetworkIsolation
Whether to allow inbound and outbound network calls to and from the containers used for the processing job.

Type: Boolean
Required: No

VpcConfig
Specifies a VPC that your training jobs and hosted models have access to. Control access to and from your training and model containers by configuring the VPC. For more information, see Protect Endpoints by Using an Amazon Virtual Private Cloud and Protect Training Jobs by Using an Amazon Virtual Private Cloud.

Type: VpcConfig (p. 1456) object
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
NotebookInstanceLifecycleConfigSummary
Service: Amazon SageMaker Service

Provides a summary of a notebook instance lifecycle configuration.

Contents

CreationTime
A timestamp that tells when the lifecycle configuration was created.
Type: Timestamp
Required: No

LastModifiedTime
A timestamp that tells when the lifecycle configuration was last modified.
Type: Timestamp
Required: No

NotebookInstanceLifecycleConfigArn
The Amazon Resource Name (ARN) of the lifecycle configuration.
Type: String
Length Constraints: Maximum length of 256.
Required: Yes

NotebookInstanceLifecycleConfigName
The name of the lifecycle configuration.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9][a-zA-Z0-9]*$^a-zA-Z0-9\][a-zA-Z0-9]*
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**NotebookInstanceLifecycleHook**

Service: Amazon SageMaker Service

Contains the notebook instance lifecycle configuration script.

Each lifecycle configuration script has a limit of 16384 characters.

The value of the `PATH` environment variable that is available to both scripts is `/sbin:bin:/usr/sbin:/usr/bin`.


Lifecycle configuration scripts cannot run for longer than 5 minutes. If a script runs for longer than 5 minutes, it fails and the notebook instance is not created or started.

For information about notebook instance lifestyle configurations, see Step 2.1: (Optional) Customize a Notebook Instance.

**Contents**

**Content**

A base64-encoded string that contains a shell script for a notebook instance lifecycle configuration.

Type: String


Pattern: `[\S\s]+`

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
NotebookInstanceSummary

Service: Amazon SageMaker Service

Provides summary information for an SageMaker notebook instance.

Contents

AdditionalCodeRepositories

An array of up to three Git repositories associated with the notebook instance. These can be either the names of Git repositories stored as resources in your account, or the URL of Git repositories in AWS CodeCommit or in any other Git repository. These repositories are cloned at the same level as the default repository of your notebook instance. For more information, see Associating Git Repositories with SageMaker Notebook Instances.

Type: Array of strings

Array Members: Maximum number of 3 items.


Pattern: ^https://([^/]+)/(.*)|^\[a-zA-Z0-9\](-*[a-zA-Z0-9]*)*

Required: No

CreationTime

A timestamp that shows when the notebook instance was created.

Type: Timestamp

Required: No

DefaultCodeRepository

The Git repository associated with the notebook instance as its default code repository. This can be either the name of a Git repository stored as a resource in your account, or the URL of a Git repository in AWS CodeCommit or in any other Git repository. When you open a notebook instance, it opens in the directory that contains this repository. For more information, see Associating Git Repositories with SageMaker Notebook Instances.

Type: String


Pattern: ^https://([^/]+)/(.*)|^\[a-zA-Z0-9\](-*[a-zA-Z0-9]*)*

Required: No

InstanceType

The type of ML compute instance that the notebook instance is running on.

Type: String

Valid Values: ml.t2.medium | ml.t2.large | ml.t2.xlarge | ml.t2.2xlarge | ml.t3.medium | ml.t3.large | ml.t3.xlarge | ml.t3.2xlarge | ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge | ml.m4.16xlarge | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge | ml.m5.24xlarge | ml.m5d.large | ml.m5d.xlarge | ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.8xlarge | ml.m5d.12xlarge | ml.m5d.16xlarge
<table>
<thead>
<tr>
<th>ml.m5d.24xlarge</th>
<th>ml.c4.xlarge</th>
<th>ml.c4.2xlarge</th>
<th>ml.c4.4xlarge</th>
<th>ml.c4.8xlarge</th>
<th>ml.c5.xlarge</th>
<th>ml.c5.2xlarge</th>
<th>ml.c5.4xlarge</th>
<th>ml.c5.9xlarge</th>
</tr>
</thead>
<tbody>
<tr>
<td>ml.c5.18xlarge</td>
<td>ml.c5d.xlarge</td>
<td>ml.c5d.2xlarge</td>
<td>ml.c5d.4xlarge</td>
<td>ml.p2.xlarge</td>
<td>ml.p2.8xlarge</td>
<td>ml.p3.2xlarge</td>
<td>ml.p3.8xlarge</td>
<td>ml.p3.16xlarge</td>
</tr>
<tr>
<td>ml.p3dn.24xlarge</td>
<td>ml.g4dn.xlarge</td>
<td>ml.g4dn.2xlarge</td>
<td>ml.g4dn.4xlarge</td>
<td>ml.g4dn.8xlarge</td>
<td>ml.g4dn.12xlarge</td>
<td>ml.g4dn.16xlarge</td>
<td>ml.r5.large</td>
<td></td>
</tr>
<tr>
<td>ml.r5.xlarge</td>
<td>ml.r5.2xlarge</td>
<td>ml.r5.4xlarge</td>
<td>ml.r5.8xlarge</td>
<td>ml.r5.12xlarge</td>
<td>ml.r5.16xlarge</td>
<td>ml.r5.24xlarge</td>
<td>ml.g5.xlarge</td>
<td></td>
</tr>
<tr>
<td>ml.g5.xlarge</td>
<td>ml.g5.2xlarge</td>
<td>ml.g5.4xlarge</td>
<td>ml.g5.8xlarge</td>
<td>ml.g5.16xlarge</td>
<td>ml.g5.12xlarge</td>
<td>ml.g5.24xlarge</td>
<td>ml.g5.48xlarge</td>
<td></td>
</tr>
</tbody>
</table>

**LastModifiedTime**

A timestamp that shows when the notebook instance was last modified.

Type: Timestamp

Required: No

**NotebookInstanceArn**

The Amazon Resource Name (ARN) of the notebook instance.

Type: String

Length Constraints: Maximum length of 256.

Required: Yes

**NotebookInstanceLifecycleConfigName**

The name of a notebook instance lifecycle configuration associated with this notebook instance.

For information about notebook instance lifestyle configurations, see Step 2.1: (Optional) Customize a Notebook Instance.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9])*`  

Required: No

**NotebookInstanceName**

The name of the notebook instance that you want a summary for.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9])*`  

Required: Yes

**NotebookInstanceStatus**

The status of the notebook instance.

Type: String
Valid Values: Pending | InService | Stopping | Stopped | Failed | Deleting | Updating

Required: No

**Url**

The URL that you use to connect to the Jupyter notebook running in your notebook instance.

Type: String

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
NotificationConfiguration

Service: Amazon SageMaker Service

Configures Amazon SNS notifications of available or expiring work items for work teams.

Contents

NotificationTopicArn

The ARN for the Amazon SNS topic to which notifications should be published.

Type: String

Pattern: `arn:aws[a-z\-]*:sns:[a-z0-9\-]*:[0-9]{12}:[a-zA-Z0-9_]\-*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ObjectiveStatusCounters

Service: Amazon SageMaker Service

Specifies the number of training jobs that this hyperparameter tuning job launched, categorized by the status of their objective metric. The objective metric status shows whether the final objective metric for the training job has been evaluated by the tuning job and used in the hyperparameter tuning process.

Contents

Failed

The number of training jobs whose final objective metric was not evaluated and used in the hyperparameter tuning process. This typically occurs when the training job failed or did not emit an objective metric.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Pending

The number of training jobs that are in progress and pending evaluation of their final objective metric.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Succeeded

The number of training jobs whose final objective metric was evaluated by the hyperparameter tuning job and used in the hyperparameter tuning process.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OfflineStoreConfig

Service: Amazon SageMaker Service

The configuration of an OfflineStore.

Provide an OfflineStoreConfig in a request to CreateFeatureGroup to create an OfflineStore.

To encrypt an OfflineStore using at rest data encryption, specify AWS Key Management Service (KMS) key ID, or KMSKeyId, in S3StorageConfig.

Contents

DataCatalogConfig

The meta data of the Glue table that is autogenerated when an OfflineStore is created.

Type: DataCatalogConfig (p. 970) object

Required: No

DisableGlueTableCreation

Set to True to disable the automatic creation of an AWS Glue table when configuring an OfflineStore.

Type: Boolean

Required: No

S3StorageConfig

The Amazon Simple Storage (Amazon S3) location of OfflineStore.

Type: S3StorageConfig (p. 1352) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OfflineStoreStatus

Service: Amazon SageMaker Service

The status of OfflineStore.

Contents

BlockedReason

The justification for why the OfflineStoreStatus is Blocked (if applicable).

Type: String

Length Constraints: Maximum length of 1024.

Required: No

Status

An OfflineStore status.

Type: String

Valid Values: Active | Blocked | Disabled

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OidcConfig

Contents

AuthorizationEndpoint

The OIDC IdP authorization endpoint used to configure your private workforce.
Type: String
Length Constraints: Maximum length of 500.
Pattern: https://\S+
Required: Yes

ClientId

The OIDC IdP client ID used to configure your private workforce.
Type: String
Pattern: [ -~]+
Required: Yes

ClientSecret

The OIDC IdP client secret used to configure your private workforce.
Type: String
Pattern: [ -~]+
Required: Yes

Issuer

The OIDC IdP issuer used to configure your private workforce.
Type: String
Length Constraints: Maximum length of 500.
Pattern: https://\S+
Required: Yes

JwksUri

The OIDC IdP JSON Web Key Set (Jwks) URI used to configure your private workforce.
Type: String
Length Constraints: Maximum length of 500.
Pattern: https://\S+
**OidcConfig**

**Required:** Yes

**LogoutEndpoint**

The OIDC IdP logout endpoint used to configure your private workforce.

- **Type:** String
- **Length Constraints:** Maximum length of 500.
- **Pattern:** https://\S+

**Required:** Yes

**TokenEndpoint**

The OIDC IdP token endpoint used to configure your private workforce.

- **Type:** String
- **Length Constraints:** Maximum length of 500.
- **Pattern:** https://\S+

**Required:** Yes

**UserInfoEndpoint**

The OIDC IdP user information endpoint used to configure your private workforce.

- **Type:** String
- **Length Constraints:** Maximum length of 500.
- **Pattern:** https://\S+

**Required:** Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OidcConfigForResponse
Service: Amazon SageMaker Service
Your OIDC IdP workforce configuration.

Contents

AuthorizationEndpoint
The OIDC IdP authorization endpoint used to configure your private workforce.
Type: String
Length Constraints: Maximum length of 500.
Pattern: https://\S+
 REQUIRED: No

ClientId
The OIDC IdP client ID used to configure your private workforce.
Type: String
Pattern: [ -~]+
 REQUIRED: No

Issuer
The OIDC IdP issuer used to configure your private workforce.
Type: String
Length Constraints: Maximum length of 500.
Pattern: https://\S+
 REQUIRED: No

JwksUri
The OIDC IdP JSON Web Key Set (Jwks) URI used to configure your private workforce.
Type: String
Length Constraints: Maximum length of 500.
Pattern: https://\S+
 REQUIRED: No

LogoutEndpoint
The OIDC IdP logout endpoint used to configure your private workforce.
Type: String
Length Constraints: Maximum length of 500.
Pattern: https://\S+
**Required:** No

**TokenEndpoint**

The OIDC IdP token endpoint used to configure your private workforce.

Type: String
Length Constraints: Maximum length of 500.

Pattern: `https://\S+`

**Required:** No

**UserInfoEndpoint**

The OIDC IdP user information endpoint used to configure your private workforce.

Type: String
Length Constraints: Maximum length of 500.

Pattern: `https://\S+`

**Required:** No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OidcMemberDefinition

Service: Amazon SageMaker Service

A list of user groups that exist in your OIDC Identity Provider (IdP). One to ten groups can be used to create a single private work team. When you add a user group to the list of Groups, you can add that user group to one or more private work teams. If you add a user group to a private work team, all workers in that user group are added to the work team.

Contents

Groups

A list of comma seperated strings that identifies user groups in your OIDC IdP. Each user group is made up of a group of private workers.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 10 items.


Pattern: [\p{L}\p{M}\p{S}\p{N}\p{P}]+

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**OnlineStoreConfig**

Service: Amazon SageMaker Service

Use this to specify the AWS Key Management Service (KMS) Key ID, or `KMSKeyId`, for at-rest data encryption. You can turn `OnlineStore` on or off by specifying the `EnableOnlineStore` flag at General Assembly; the default value is `False`.

**Contents**

**EnableOnlineStore**

Turn `OnlineStore` off by specifying `False` for the `EnableOnlineStore` flag. Turn `OnlineStore` on by specifying `True` for the `EnableOnlineStore` flag.

The default value is `False`.

Type: Boolean

Required: No

**SecurityConfig**

Use to specify KMS Key ID (`KMSKeyId`) for at-rest encryption of your `OnlineStore`.

Type: `OnlineStoreSecurityConfig (p. 1239)` object

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OnlineStoreSecurityConfig

Service: Amazon SageMaker Service

The security configuration for OnlineStore.

Contents

KmsKeyId

The ID of the AWS Key Management Service (AWS KMS) key that SageMaker Feature Store uses to encrypt the Amazon S3 objects at rest using Amazon S3 server-side encryption.

The caller (either IAM user or IAM role) of CreateFeatureGroup must have below permissions to the OnlineStore KmsKeyId:

• "kms:Encrypt"
• "kms:Decrypt"
• "kms:DescribeKey"
• "kms:CreateGrant"
• "kms:RetireGrant"
• "kms:ReEncryptFrom"
• "kms:ReEncryptTo"
• "kms:GenerateDataKey"
• "kms:ListAliases"
• "kms:ListGrants"
• "kms:RevokeGrant"

The caller (either IAM user or IAM role) to all DataPlane operations (PutRecord, GetRecord, DeleteRecord) must have the following permissions to the KmsKeyId:

• "kms:Decrypt"

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OutputConfig

Service: Amazon SageMaker Service

Contains information about the output location for the compiled model and the target device that the model runs on. TargetDevice and TargetPlatform are mutually exclusive, so you need to choose one between the two to specify your target device or platform. If you cannot find your device you want to use from the TargetDevice list, use TargetPlatform to describe the platform of your edge device and CompilerOptions if there are specific settings that are required or recommended to use for particular TargetPlatform.

Contents

CompilerOptions

Specifies additional parameters for compiler options in JSON format. The compiler options are TargetPlatform specific. It is required for NVIDIA accelerators and highly recommended for CPU compilations. For any other cases, it is optional to specify CompilerOptions.

- **DTYPE**: Specifies the data type for the input. When compiling for ml_* (except for ml_inf) instances using PyTorch framework, provide the data type (dtype) of the model's input. "float32" is used if "DTYPE" is not specified. Options for data type are:
  - float32: Use either "float" or "float32".
  - int64: Use either "int64" or "long".

  For example, 
  ```json
  {"dtype": "float32"}
  ```

- **CPU**: Compilation for CPU supports the following compiler options.
  - mcpu: CPU micro-architecture. For example, 
    ```json
    {'mcpu': 'skylake-avx512'}
    ```
  - mattr: CPU flags. For example, 
    ```json
    {'mattr': ['+neon', '+vfpv4']}
    ```

- **ARM**: Details of ARM CPU compilations.
  - NEON: NEON is an implementation of the Advanced SIMD extension used in ARMv7 processors.

    For example, add 
    ```json
    {'mattr': ['+neon']}
    ```
    to the compiler options if compiling for ARM 32-bit platform with the NEON support.

- **NVIDIA**: Compilation for NVIDIA GPU supports the following compiler options.
  - gpu_code: Specifies the targeted architecture.
  - trt-ver: Specifies the TensorRT versions in x.y.z format.
  - cuda-ver: Specifies the CUDA version in x.y format.

    For example, 
    ```json
    {'gpu-code': 'sm_72', 'trt-ver': '6.0.1', 'cuda-ver': '10.1'}
    ```

- **ANDROID**: Compilation for the Android OS supports the following compiler options:
  - ANDROID_PLATFORM: Specifies the Android API levels. Available levels range from 21 to 29. For example, 
    ```json
    {'ANDROID_PLATFORM': 28}
    ```
  - mattr: Add 
    ```json
    {'mattr': ['+neon']}
    ```
    to compiler options if compiling for ARM 32-bit platform with NEON support.

- **INFERENTIA**: Compilation for target ml_inf1 uses compiler options passed in as a JSON string. For example, 
  ```json
  "CompilerOptions": "--verbose 1 --num-neuroncores 2 -O2"
  ```

  For information about supported compiler options, see Neuron Compiler CLI.

- **CoreML**: Compilation for the CoreML OutputConfig:TargetDevice (p. 1241) supports the following compiler options:
  - class_labels: Specifies the classification labels file name inside input tar.gz file. For example, 
    ```json
    {"class_labels": "imagenet_labels_1000.txt"}
    ```
    Labels inside the txt file should be separated by newlines.
• **EIA**: Compilation for the Elastic Inference Accelerator supports the following compiler options:
  • **precision_mode**: Specifies the precision of compiled artifacts. Supported values are "FP16" and "FP32". Default is "FP32".
  • **signature_def_key**: Specifies the signature to use for models in SavedModel format. Defaults is TensorFlow's default signature def key.
  • **output_names**: Specifies a list of output tensor names for models in FrozenGraph format. Set at most one API field, either: signature_def_key or output_names.

For example: `{"precision_mode": "FP32", "output_names": ["output:0"]}

Type: String


Pattern: .*

Required: No

**KmsKeyId**

The AWS Key Management Service key (AWS KMS) that Amazon SageMaker uses to encrypt your output models with Amazon S3 server-side encryption after compilation job. If you don't provide a KMS key ID, Amazon SageMaker uses the default KMS key for Amazon S3 for your role's account. For more information, see KMS-Managed Encryption Keys in the Amazon Simple Storage Service Developer Guide.

The KmsKeyId can be any of the following formats:
  • Key ID: 1234abcd-12ab-34cd-56ef-1234567890ab
  • Key ARN: arn:aws:kms:us-west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab
  • Alias name: alias/ExampleAlias
  • Alias name ARN: arn:aws:kms:us-west-2:111122223333:alias/ExampleAlias

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

**S3OutputLocation**

Identifies the S3 bucket where you want Amazon SageMaker to store the model artifacts. For example, `s3://bucket-name/key-name-prefix`.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^https|s3://([^/]*)/\?\.*$`

Required: Yes

**TargetDevice**

Identifies the target device or the machine learning instance that you want to run your model on after the compilation has completed. Alternatively, you can specify OS, architecture, and accelerator using **TargetPlatform** (p. 1378) fields. It can be used instead of TargetPlatform.

Type: String
Valid Values: lambda | ml_m4 | ml_m5 | ml_c4 | ml_c5 | ml_p2 | ml_p3 | ml_g4dn
| ml_inf1 | ml_eia2 | jetson_tx1 | jetson_tx2 | jetson_nano | jetson_xavier
| rasp3b | imx8qm | deeplens | rk3399 | rk3288 | aisage | sbe_c | qcs605
| qcs603 | sitara_am57x | amba_cv2 | amba_cv25 | x86_win32 | x86_win64 | coreml | jacinto_tda4vm | imx8mplus

Required: No

**TargetPlatform**

Contains information about a target platform that you want your model to run on, such as OS, architecture, and accelerators. It is an alternative of TargetDevice.

The following examples show how to configure the TargetPlatform and CompilerOptions JSON strings for popular target platforms:

- **Raspberry Pi 3 Model B+**
  
  "TargetPlatform": {"Os": "LINUX", "Arch": "ARM_EABIHF"},
  
  "CompilerOptions": {'mattr': ['+neon']}

- **Jetson TX2**
  
  "TargetPlatform": {"Os": "LINUX", "Arch": "ARM64", "Accelerator": "NVIDIA"},
  
  "CompilerOptions": {'gpu-code': 'sm_62', 'trt-ver': '6.0.1', 'cuda-ver': '10.0'}

- **EC2 m5.2xlarge instance OS**
  
  "TargetPlatform": {"Os": "LINUX", "Arch": "X86_64", "Accelerator": "NVIDIA"},
  
  "CompilerOptions": {'mcpu': 'skylake-avx512'}

- **RK3399**
  
  "TargetPlatform": {"Os": "LINUX", "Arch": "ARM64", "Accelerator": "MALI"}

- **ARMv7 phone (CPU)**
  
  "TargetPlatform": {"Os": "ANDROID", "Arch": "ARM_EABI"},
  
  "CompilerOptions": {'ANDROID_PLATFORM': 25, 'mattr': ['+neon']}

- **ARMv8 phone (CPU)**
  
  "TargetPlatform": {"Os": "ANDROID", "Arch": "ARM64"},
  
  "CompilerOptions": {'ANDROID_PLATFORM': 29}

Type: **TargetPlatform (p. 1378) object**

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OutputDataConfig

Service: Amazon SageMaker Service

Provides information about how to store model training results (model artifacts).

Contents

KmsKeyId

The AWS Key Management Service (AWS KMS) key that SageMaker uses to encrypt the model artifacts at rest using Amazon S3 server-side encryption. The KmsKeyId can be any of the following formats:

- // KMS Key ID
  "1234abcd-12ab-34cd-56ef-1234567890ab"
- // Amazon Resource Name (ARN) of a KMS Key
  "arn:aws:kms:us-west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
- // KMS Key Alias
  "alias/ExampleAlias"
- // Amazon Resource Name (ARN) of a KMS Key Alias

If you use a KMS key ID or an alias of your KMS key, the SageMaker execution role must include permissions to call kms:Encrypt. If you don't provide a KMS key ID, SageMaker uses the default KMS key for Amazon S3 for your role's account. SageMaker uses server-side encryption with KMS-managed keys for OutputDataConfig. If you use a bucket policy with an s3:PutObject permission that only allows objects with server-side encryption, set the condition key of s3:x-amz-server-side-encryption to "aws:kms". For more information, see KMS-Managed Encryption Keys in the Amazon Simple Storage Service Developer Guide.

The KMS key policy must grant permission to the IAM role that you specify in your CreateTrainingJob, CreateTransformJob, or CreateHyperParameterTuningJob requests. For more information, see Using Key Policies in AWS KMS in the AWS Key Management Service Developer Guide.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

S3OutputPath

Identifies the S3 path where you want SageMaker to store the model artifacts. For example, s3://bucket-name/key-name-prefix.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)/?([^/]+$)

Required: Yes
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
OutputParameter

Service: Amazon SageMaker Service

An output parameter of a pipeline step.

Contents

Name

The name of the output parameter.

Type: String

Length Constraints: Maximum length of 256.

Required: Yes

Value

The value of the output parameter.

Type: String

Length Constraints: Maximum length of 1024.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ParallelismConfiguration
Service: Amazon SageMaker Service

Configuration that controls the parallelism of the pipeline. By default, the parallelism configuration specified applies to all executions of the pipeline unless overridden.

Contents

MaxParallelExecutionSteps

The max number of steps that can be executed in parallel.

Type: Integer

Valid Range: Minimum value of 1.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Parameter
Service: Amazon SageMaker Service
Assigns a value to a named Pipeline parameter.

Contents

Name
The name of the parameter to assign a value to. This parameter name must match a named parameter in the pipeline definition.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,255}\$
Required: Yes

Value
The literal value for the parameter.
Type: String
Length Constraints: Maximum length of 1024.
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ParameterRange
Service: Amazon SageMaker Service

Defines the possible values for categorical, continuous, and integer hyperparameters to be used by an algorithm.

Contents

CategoricalParameterRangeSpecification

A CategoricalParameterRangeSpecification object that defines the possible values for a categorical hyperparameter.

Type: CategoricalParameterRangeSpecification (p. 940) object

Required: No

ContinuousParameterRangeSpecification

A ContinuousParameterRangeSpecification object that defines the possible values for a continuous hyperparameter.

Type: ContinuousParameterRangeSpecification (p. 964) object

Required: No

IntegerParameterRangeSpecification

A IntegerParameterRangeSpecification object that defines the possible values for an integer hyperparameter.

Type: IntegerParameterRangeSpecification (p. 1114) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ParameterRanges
Service: Amazon SageMaker Service

Specifies ranges of integer, continuous, and categorical hyperparameters that a hyperparameter tuning job searches. The hyperparameter tuning job launches training jobs with hyperparameter values within these ranges to find the combination of values that result in the training job with the best performance as measured by the objective metric of the hyperparameter tuning job.

**Note**
You can specify a maximum of 20 hyperparameters that a hyperparameter tuning job can search over. Every possible value of a categorical parameter range counts against this limit.

**Contents**

**CategoricalParameterRanges**

The array of CategoricalParameterRange (p. 939) objects that specify ranges of categorical hyperparameters that a hyperparameter tuning job searches.

Type: Array of CategoricalParameterRange (p. 939) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

Required: No

**ContinuousParameterRanges**

The array of ContinuousParameterRange (p. 962) objects that specify ranges of continuous hyperparameters that a hyperparameter tuning job searches.

Type: Array of ContinuousParameterRange (p. 962) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

Required: No

**IntegerParameterRanges**

The array of IntegerParameterRange (p. 1112) objects that specify ranges of integer hyperparameters that a hyperparameter tuning job searches.

Type: Array of IntegerParameterRange (p. 1112) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Parent
Service: Amazon SageMaker Service

The trial that a trial component is associated with and the experiment the trial is part of. A component might not be associated with a trial. A component can be associated with multiple trials.

Contents

ExperimentName

The name of the experiment.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^(a-zA-Z0-9)(-[a-zA-Z0-9]{0,119})

Required: No

TrialName

The name of the trial.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^(a-zA-Z0-9)(-[a-zA-Z0-9]{0,119})

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ParentHyperParameterTuningJob
Service: Amazon SageMaker Service

A previously completed or stopped hyperparameter tuning job to be used as a starting point for a new hyperparameter tuning job.

Contents

HyperParameterTuningJobName

The name of the hyperparameter tuning job to be used as a starting point for a new hyperparameter tuning job.

Type: String


Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9]\{0,31\}

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PendingDeploymentSummary
Service: Amazon SageMaker Service

The summary of an in-progress deployment when an endpoint is creating or updating with a new endpoint configuration.

Contents

EndpointConfigName

The name of the endpoint configuration used in the deployment.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,62}\}

Required: Yes

ProductionVariants

List of PendingProductionVariantSummary objects.

Type: Array of PendingProductionVariantSummary (p. 1254) objects

Array Members: Minimum number of 1 item.

Required: No

StartTime

The start time of the deployment.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PendingProductionVariantSummary

Service: Amazon SageMaker Service

The production variant summary for a deployment when an endpoint is creating or updating with the CreateEndpoint (p. 69) or UpdateEndpoint (p. 767) operations. Describes the VariantStatus, weight and capacity for a production variant associated with an endpoint.

Contents

**AcceleratorType**

The size of the Elastic Inference (EI) instance to use for the production variant. EI instances provide on-demand GPU computing for inference. For more information, see Using Elastic Inference in Amazon SageMaker.

Type: String

Valid Values: ml.eia1.medium | ml.eia1.large | ml.eia1.xlarge | ml.eia2.medium | ml.eia2.large | ml.eia2.xlarge

Required: No

**CurrentInstanceCount**

The number of instances associated with the variant.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**CurrentServerlessConfig**

The serverless configuration for the endpoint.

Type: ProductionVariantServerlessConfig (p. 1297) object

Required: No

**CurrentWeight**

The weight associated with the variant.

Type: Float

Valid Range: Minimum value of 0.

Required: No

**DeployedImages**

An array of DeployedImage objects that specify the Amazon EC2 Container Registry paths of the inference images deployed on instances of this ProductionVariant.

Type: Array of DeployedImage (p. 986) objects

Required: No

**DesiredInstanceCount**

The number of instances requested in this deployment, as specified in the endpoint configuration for the endpoint. The value is taken from the request to the CreateEndpointConfig (p. 73) operation.
Type: Integer

Valid Range: Minimum value of 0.

Required: No

**DesiredServerlessConfig**

The serverless configuration requested for this deployment, as specified in the endpoint configuration for the endpoint.

Type: `ProductionVariantServerlessConfig (p. 1297)` object

Required: No

**DesiredWeight**

The requested weight for the variant in this deployment, as specified in the endpoint configuration for the endpoint. The value is taken from the request to the `CreateEndpointConfig (p. 73)` operation.

Type: Float

Valid Range: Minimum value of 0.

Required: No

**InstanceType**

The type of instances associated with the variant.

Type: String

Valid Values: `ml.t2.medium | ml.t2.large | ml.t2.xlarge | ml.t2.2xlarge | ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge | ml.m4.16xlarge | ml.m5.large | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge | ml.m5.24xlarge | ml.m5d.large | ml.m5d.xlarge | ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.12xlarge | ml.m5d.24xlarge | ml.c4.large | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge | ml.c4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge | ml.p3.xlarge | ml.p3.16xlarge | ml.p3.8xlarge | ml.c5.large | ml.c5.xlarge | ml.c5.2xlarge | ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge | ml.c5d.large | ml.c5d.xlarge | ml.c5d.2xlarge | ml.c5d.4xlarge | ml.c5d.9xlarge | ml.c5d.18xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.r5.large | ml.r5.xlarge | ml.r5.2xlarge | ml.r5.4xlarge | ml.r5.12xlarge | ml.r5.24xlarge | ml.r5d.large | ml.r5d.xlarge | ml.r5d.2xlarge | ml.r5d.4xlarge | ml.r5d.12xlarge | ml.r5d.24xlarge | ml.inf1.xlarge | ml.inf1.2xlarge | ml.inf1.6xlarge | ml.inf1.24xlarge

Required: No

**VariantName**

The name of the variant.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9-]{0,62}$`

Required: Yes
**VariantStatus**

The endpoint variant status which describes the current deployment stage status or operational status.

Type: Array of `ProductionVariantStatus` objects

Array Members: Minimum number of 0 items. Maximum number of 5 items.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](https://aws.amazon.com/sdk-for-cpp/)
- [AWS SDK for Go](https://golang.org/)
- [AWS SDK for Java V2](https://aws.amazon.com/sdk-for-java/)
- [AWS SDK for Ruby V3](https://aws.amazon.com/sdk-for-ruby/)
Phase
Service: Amazon SageMaker Service
Defines the traffic pattern.

Contents

**DurationInSeconds**

Specifies how long traffic phase should be.
Type: Integer
Valid Range: Minimum value of 1.
Required: No

**InitialNumberOfUsers**

Specifies how many concurrent users to start with.
Type: Integer
Valid Range: Minimum value of 1.
Required: No

**SpawnRate**

Specified how many new users to spawn in a minute.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Pipeline
Service: Amazon SageMaker Service

A SageMaker Model Building Pipeline instance.

Contents

CreatedBy
Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
Type: UserContext (p. 1449) object
Required: No

CreationTime
The creation time of the pipeline.
Type: Timestamp
Required: No

LastModifiedBy
Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
Type: UserContext (p. 1449) object
Required: No

LastModifiedTime
The time that the pipeline was last modified.
Type: Timestamp
Required: No

LastRunTime
The time when the pipeline was last run.
Type: Timestamp
Required: No

ParallelismConfiguration
The parallelism configuration applied to the pipeline.
Type: ParallelismConfiguration (p. 1247) object
Required: No

PipelineArn
The Amazon Resource Name (ARN) of the pipeline.
Type: String
Length Constraints: Maximum length of 256.
Pipeline

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*

Required: No

**PipelineDescription**

The description of the pipeline.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 3072.

Pattern: .*

Required: No

**PipelineDisplayName**

The display name of the pipeline.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9])\{0,255\}

Required: No

**PipelineName**

The name of the pipeline.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9])\{0,255\}

Required: No

**PipelineStatus**

The status of the pipeline.

Type: String

Valid Values: Active

Required: No

**RoleArn**

The Amazon Resource Name (ARN) of the role that created the pipeline.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9\-\+=,.@\-_\/%]*$\n
Required: No

**Tags**

A list of tags that apply to the pipeline.

Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PipelineDefinitionS3Location

Service: Amazon SageMaker Service

The location of the pipeline definition stored in Amazon S3.

Contents

**Bucket**

Name of the S3 bucket.

Type: String


Pattern: [a-z0-9][\.-a-z0-9]{1,61}[a-z0-9]

Required: Yes

**ObjectKey**

The object key (or key name) uniquely identifies the object in an S3 bucket.

Type: String


Pattern: .+

Required: Yes

**VersionId**

Version Id of the pipeline definition file. If not specified, Amazon SageMaker will retrieve the latest version.

Type: String


Pattern: .+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PipelineExecution

Service: Amazon SageMaker Service

An execution of a pipeline.

Contents

CreatedBy

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

Required: No

CreationTime

The creation time of the pipeline execution.

Type: Timestamp

Required: No

FailureReason

If the execution failed, a message describing why.

Type: String

Length Constraints: Maximum length of 1300.

Pattern: . *

Required: No

LastModifiedBy

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

Required: No

LastModifiedTime

The time that the pipeline execution was last modified.

Type: Timestamp

Required: No

ParallelismConfiguration

The parallelism configuration applied to the pipeline execution.

Type: ParallelismConfiguration (p. 1247) object

Required: No

PipelineArn

The Amazon Resource Name (ARN) of the pipeline that was executed.
PipelineExecution

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*

Required: No

PipelineExecutionArn

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*\/*\/*

Required: No

PipelineExecutionDescription

The description of the pipeline execution.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 3072.

Pattern: .*

Required: No

PipelineExecutionDisplayName

The display name of the pipeline execution.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 82.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,81}

Required: No

PipelineExecutionStatus

The status of the pipeline status.

Type: String

Valid Values: Executing | Stopping | Stopped | Failed | Succeeded

Required: No

PipelineExperimentConfig

Specifies the names of the experiment and trial created by a pipeline.

Type: PipelineExperimentConfig (p. 1272) object

Required: No

PipelineParameters

Contains a list of pipeline parameters. This list can be empty.
Type: Array of Parameter (p. 1248) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**PipelineExecutionStep**

Service: Amazon SageMaker Service

An execution of a step in a pipeline.

**Contents**

**AttemptCount**

The current attempt of the execution step. For more information, see [Retry Policy for SageMaker Pipelines steps](#).

Type: Integer  
Required: No

**CacheHitResult**

If this pipeline execution step was cached, details on the cache hit.

Type: `CacheHitResult` (p. 931) object  
Required: No

**EndTime**

The time that the step stopped executing.

Type: Timestamp  
Required: No

**FailureReason**

The reason why the step failed execution. This is only returned if the step failed its execution.

Type: String  
Length Constraints: Maximum length of 1024.  
Required: No

**Metadata**

Metadata for the step execution.

Type: `PipelineExecutionStepMetadata` (p. 1267) object  
Required: No

**StartTime**

The time that the step started executing.

Type: Timestamp  
Required: No

**StepDescription**

The description of the step.

Type: String  
Length Constraints: Minimum length of 0. Maximum length of 3072.
Pattern: . *
Required: No
**StepDisplayName**
The display name of the step.
Type: String
Length Constraints: Minimum length of 0. Maximum length of 256.
Pattern: . *
Required: No
**StepName**
The name of the step that is executed.
Type: String
Length Constraints: Maximum length of 64.
Pattern: ^[A-Za-z0-9\-\_]*$
Required: No
**StepStatus**
The status of the step execution.
Type: String
Valid Values: Starting | Executing | Stopping | Stopped | Failed | Succeeded
Required: No

**See Also**
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**PipelineExecutionStepMetadata**

Service: Amazon SageMaker Service

Metadata for a step execution.

**Contents**

**Callback**

The URL of the Amazon SQS queue used by this step execution, the pipeline generated token, and a list of output parameters.

Type: `CallbackStepMetadata (p. 932)` object

Required: No

**ClarifyCheck**

Container for the metadata for a Clarify check step. The configurations and outcomes of the check step execution. This includes:

- The type of the check conducted,
- The Amazon S3 URIs of baseline constraints and statistics files to be used for the drift check.
- The Amazon S3 URIs of newly calculated baseline constraints and statistics.
- The model package group name provided.
- The Amazon S3 URI of the violation report if violations detected.
- The Amazon Resource Name (ARN) of check processing job initiated by the step execution.
- The boolean flags indicating if the drift check is skipped.
- If step property `BaselineUsedForDriftCheck` is set the same as `CalculatedBaseline`.

Type: `ClarifyCheckStepMetadata (p. 946)` object

Required: No

**Condition**

The outcome of the condition evaluation that was run by this step execution.

Type: `ConditionStepMetadata (p. 955)` object

Required: No

**EMR**

The configurations and outcomes of an EMR step execution.

Type: `EMRStepMetadata (p. 1013)` object

Required: No

**Fail**

The configurations and outcomes of a Fail step execution.

Type: `FailStepMetadata (p. 1035)` object

Required: No

**Lambda**

The Amazon Resource Name (ARN) of the Lambda function that was run by this step execution and a list of output parameters.
PipelineExecutionStepMetadata

Type: LambdaStepMetadata (p. 1139) object

Required: No

Model

The Amazon Resource Name (ARN) of the model that was created by this step execution.

Type: ModelStepMetadata (p. 1191) object

Required: No

ProcessingJob

The Amazon Resource Name (ARN) of the processing job that was run by this step execution.

Type: ProcessingJobStepMetadata (p. 1283) object

Required: No

QualityCheck

The configurations and outcomes of the check step execution. This includes:
- The type of the check conducted,
- The Amazon S3 URIs of baseline constraints and statistics files to be used for the drift check.
- The Amazon S3 URIs of newly calculated baseline constraints and statistics.
- The model package group name provided.
- The Amazon S3 URI of the violation report if violations detected.
- The Amazon Resource Name (ARN) of check processing job initiated by the step execution.
- The boolean flags indicating if the drift check is skipped.
- If step property `BaselineUsedForDriftCheck` is set the same as `CalculatedBaseline`.

Type: QualityCheckStepMetadata (p. 1320) object

Required: No

RegisterModel

The Amazon Resource Name (ARN) of the model package the model was registered to by this step execution.

Type: RegisterModelStepMetadata (p. 1333) object

Required: No

TrainingJob

The Amazon Resource Name (ARN) of the training job that was run by this step execution.

Type: TrainingJobStepMetadata (p. 1397) object

Required: No

TransformJob

The Amazon Resource Name (ARN) of the transform job that was run by this step execution.

Type: TransformJobStepMetadata (p. 1412) object

Required: No

TuningJob

The Amazon Resource Name (ARN) of the tuning job that was run by this step execution.
Type: TuningJobStepMetaData (p. 1443) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PipelineExecutionSummary

Service: Amazon SageMaker Service

A pipeline execution summary.

Contents

PipelineExecutionArn

The Amazon Resource Name (ARN) of the pipeline execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline\/.*/\ execution\/.*$

Required: No

PipelineExecutionDescription

The description of the pipeline execution.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 3072.

Pattern: .*

Required: No

PipelineExecutionDisplayName

The display name of the pipeline execution.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 82.

Pattern: ^[a-zA-Z0-9\-]*(\-[a-zA-Z0-9\-])\{0,81}$

Required: No

PipelineExecutionFailureReason

A message generated by SageMaker Pipelines describing why the pipeline execution failed.

Type: String

Length Constraints: Maximum length of 3072.

Required: No

PipelineExecutionStatus

The status of the pipeline execution.

Type: String

Valid Values: Executing | Stopping | Stopped | Failed | Succeeded

Required: No
StartTime

The start time of the pipeline execution.
Type: Timestamp
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PipelineExperimentConfig
Service: Amazon SageMaker Service

Specifies the names of the experiment and trial created by a pipeline.

Contents

**ExperimentName**

The name of the experiment.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,119\}

Required: No

**TrialName**

The name of the trial.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,119\}

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PipelineSummary
Service: Amazon SageMaker Service
A summary of a pipeline.

Contents

CreationTime
The creation time of the pipeline.
Type: Timestamp
Required: No

LastExecutionTime
The last time that a pipeline execution began.
Type: Timestamp
Required: No

LastModifiedTime
The time that the pipeline was last modified.
Type: Timestamp
Required: No

PipelineArn
The Amazon Resource Name (ARN) of the pipeline.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:pipeline/.*
Required: No

PipelineDescription
The description of the pipeline.
Type: String
Length Constraints: Minimum length of 0. Maximum length of 3072.
Pattern: .*
Required: No

PipelineDisplayName
The display name of the pipeline.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: ^[a-zA-Z0-9]([-*[a-zA-Z0-9])\{0,255\}
PipelineSummary

Required: No

**PipelineName**

The name of the pipeline.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,255}$`

Required: No

**RoleArn**

The Amazon Resource Name (ARN) that the pipeline used to execute.

Type: String


Pattern: `^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9=+.@\-_\/]\+$`

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingClusterConfig

Service: Amazon SageMaker Service

Configuration for the cluster used to run a processing job.

Contents

InstanceCount

The number of ML compute instances to use in the processing job. For distributed processing jobs, specify a value greater than 1. The default value is 1.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: Yes

InstanceType

The ML compute instance type for the processing job.

Type: String

Valid Values: ml.t3.medium | ml.t3.large | ml.t3.xlarge | ml.t3.2xlarge
| ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge
| ml.m4.16xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge
| ml.c4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge
| ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.c5.xlarge
| ml.c5.2xlarge | ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge | ml.m5.large
| ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge
| ml.m5.24xlarge | ml.r5.large | ml.r5.xlarge | ml.r5.2xlarge | ml.r5.4xlarge
| ml.r5.8xlarge | ml.r5.12xlarge | ml.r5.16xlarge | ml.r5.24xlarge
| ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge | ml.g4dn.12xlarge
| ml.g4dn.16xlarge

Required: Yes

VolumeKmsKeyId

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt data on the storage volume attached to the ML compute instance(s) that run the processing job.

Note

Certain Nitro-based instances include local storage, dependent on the instance type. Local storage volumes are encrypted using a hardware module on the instance. You can't request a VolumeKmsKeyId when using an instance type with local storage.

For a list of instance types that support local instance storage, see Instance Store Volumes.

For more information about local instance storage encryption, see SSD Instance Store Volumes.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

VolumeSizeInGB

The size of the ML storage volume in gigabytes that you want to provision. You must specify sufficient ML storage for your scenario.
Note
Certain Nitro-based instances include local storage with a fixed total size, dependent on the instance type. When using these instances for processing, Amazon SageMaker mounts the local instance storage instead of Amazon EBS gp2 storage. You can't request a VolumeSizeInGB greater than the total size of the local instance storage. For a list of instance types that support local instance storage, including the total size per instance type, see Instance Store Volumes.

Type: Integer
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingFeatureStoreOutput
Service: Amazon SageMaker Service

Configuration for processing job outputs in Amazon SageMaker Feature Store.

Contents

FeatureGroupName

The name of the Amazon SageMaker FeatureGroup to use as the destination for processing job output. Note that your processing script is responsible for putting records into your Feature Store.

Type: String
Length Constraints: Minimum length of 1. Maximum length of 64.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9])\{0,63\}
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**ProcessingInput**

Service: Amazon SageMaker Service

The inputs for a processing job. The processing input must specify exactly one of either `S3Input` or `DatasetDefinition` types.

**Contents**

*AppManaged*

When `True`, input operations such as data download are managed natively by the processing job application. When `False` (default), input operations are managed by Amazon SageMaker.

Type: Boolean

Required: No

*DatasetDefinition*

Configuration for a Dataset Definition input.

Type: `DatasetDefinition (p. 977)` object

Required: No

*InputName*

The name for the processing job input.

Type: String

Required: Yes

*S3Input*

Configuration for downloading input data from Amazon S3 into the processing container.

Type: `ProcessingS3Input (p. 1289)` object

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingJob

Service: Amazon SageMaker Service

An Amazon SageMaker processing job that is used to analyze data and evaluate models. For more information, see Process Data and Evaluate Models.

Contents

AppSpecification

Configuration to run a processing job in a specified container image.

Type: AppSpecification (p. 891) object

Required: No

AutoMLJobArn

The Amazon Resource Name (ARN) of the AutoML job associated with this processing job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*

Required: No

CreationTime

The time the processing job was created.

Type: Timestamp

Required: No

Environment

Sets the environment variables in the Docker container.

Type: String to string map

Map Entries: Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: [a-zA-Z_][a-zA-Z0-9_]*

Value Length Constraints: Maximum length of 256.

Value Pattern: [\S\s]*

Required: No

ExitMessage

A string, up to one KB in size, that contains metadata from the processing container when the processing job exits.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: [\S\s]*
Required: No

**ExperimentConfig**

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:

- `CreateProcessingJob (p. 162)`
- `CreateTrainingJob (p. 173)`
- `CreateTransformJob (p. 182)`

Type: `ExperimentConfig (p. 1029)` object

Required: No

**FailureReason**

A string, up to one KB in size, that contains the reason a processing job failed, if it failed.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

**LastModifiedTime**

The time the processing job was last modified.

Type: Timestamp

Required: No

**MonitoringScheduleArn**

The ARN of a monitoring schedule for an endpoint associated with this processing job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `.*`

Required: No

**NetworkConfig**

Networking options for a job, such as network traffic encryption between containers, whether to allow inbound and outbound network calls to and from containers, and the VPC subnets and security groups to use for VPC-enabled jobs.

Type: `NetworkConfig (p. 1223)` object

Required: No

**ProcessingEndTime**

The time that the processing job ended.

Type: Timestamp

Required: No

**ProcessingInputs**

List of input configurations for the processing job.
Type: Array of `ProcessingInput` (p. 1278) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Required: No

**ProcessingJobArn**

The ARN of the processing job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:processing-job/.*`

Required: No

**ProcessingJobName**

The name of the processing job.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}`

Required: No

**ProcessingJobStatus**

The status of the processing job.

Type: String

Valid Values: `InProgress` | `Completed` | `Failed` | `Stopping` | `Stopped`

Required: No

**ProcessingOutputConfig**

Configuration for uploading output from the processing container.

Type: `ProcessingOutputConfig` (p. 1287) object

Required: No

**ProcessingResources**

Identifies the resources, ML compute instances, and ML storage volumes to deploy for a processing job. In distributed training, you specify more than one instance.

Type: `ProcessingResources` (p. 1288) object

Required: No

**ProcessingStartTime**

The time that the processing job started.

Type: Timestamp

Required: No

**RoleArn**

The ARN of the role used to create the processing job.
ProcessingJob

Type: String
Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/?[a-zA-Z0-9+=,.@\-_]/+\$
Required: No

StoppingCondition

Configures conditions under which the processing job should be stopped, such as how long the processing job has been running. After the condition is met, the processing job is stopped.

Type: ProcessingStoppingCondition (p. 1292) object
Required: No

Tags

An array of key-value pairs. For more information, see Using Cost Allocation Tags in the AWS Billing and Cost Management User Guide.

Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

TrainingJobArn

The ARN of the training job associated with this processing job.

Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-zA-Z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:training-job/.*
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingJobStepMetadata
Service: Amazon SageMaker Service
Metadata for a processing job step.

Contents

Arn
The Amazon Resource Name (ARN) of the processing job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:processing-job/.*
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingJobSummary
Service: Amazon SageMaker Service
Summary of information about a processing job.

Contents

CreationTime
The time at which the processing job was created.
Type: Timestamp
Required: Yes

ExitMessage
An optional string, up to one KB in size, that contains metadata from the processing container when the processing job exits.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: [\S\s]*
Required: No

FailureReason
A string, up to one KB in size, that contains the reason a processing job failed, if it failed.
Type: String
Length Constraints: Maximum length of 1024.
Required: No

LastModifiedTime
A timestamp that indicates the last time the processing job was modified.
Type: Timestamp
Required: No

ProcessingEndTime
The time at which the processing job completed.
Type: Timestamp
Required: No

ProcessingJobArn
The Amazon Resource Name (ARN) of the processing job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:processing-job/.*
Required: Yes

**ProcessingJobName**

The name of the processing job.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}`

Required: Yes

**ProcessingJobStatus**

The status of the processing job.

Type: String

Valid Values: InProgress | Completed | Failed | Stopping | Stopped

Required: Yes

### See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingOutput
Service: Amazon SageMaker Service

Describes the results of a processing job. The processing output must specify exactly one of either S3Output or FeatureStoreOutput types.

Contents

AppManaged
When True, output operations such as data upload are managed natively by the processing job application. When False (default), output operations are managed by Amazon SageMaker.

Type: Boolean
Required: No

FeatureStoreOutput
Configuration for processing job outputs in Amazon SageMaker Feature Store. This processing output type is only supported when AppManaged is specified.

Type: ProcessingFeatureStoreOutput (p. 1277) object
Required: No

OutputName
The name for the processing job output.

Type: String
Required: Yes

S3Output
Configuration for processing job outputs in Amazon S3.

Type: ProcessingS3Output (p. 1291) object
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingOutputConfig

Service: Amazon SageMaker Service

Configuration for uploading output from the processing container.

Contents

KmsKeyId

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt the processing job output. KmsKeyId can be an ID of a KMS key, ARN of a KMS key, alias of a KMS key, or alias of a KMS key. The KmsKeyId is applied to all outputs.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

Outputs

An array of outputs configuring the data to upload from the processing container.

Type: Array of ProcessingOutput (p. 1286) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingResources

Service: Amazon SageMaker Service

Identifies the resources, ML compute instances, and ML storage volumes to deploy for a processing job. In distributed training, you specify more than one instance.

Contents

ClusterConfig

The configuration for the resources in a cluster used to run the processing job.

Type: ProcessingClusterConfig (p. 1275) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingS3Input
Service: Amazon SageMaker Service

Contents

LocalPath
The local path in your container where you want Amazon SageMaker to write input data to.
LocalPath is an absolute path to the input data and must begin with /opt/ml/processing/
LocalPath is a required parameter when AppManaged is False (default).

Type: String
Length Constraints: Maximum length of 256.
Pattern: .*
Required: No

S3CompressionType
Whether to GZIP-decompress the data in Amazon S3 as it is streamed into the processing container.
Gzip can only be used when Pipe mode is specified as the S3InputMode. In Pipe mode, Amazon SageMaker streams input data from the source directly to your container without using the EBS volume.

Type: String
Valid Values: None | Gzip
Required: No

S3DataDistributionType
Whether to distribute the data from Amazon S3 to all processing instances with FullyReplicated, or whether the data from Amazon S3 is shared by Amazon S3 key, downloading one shard of data to each processing instance.

Type: String
Valid Values: FullyReplicated | ShardedByS3Key
Required: No

S3DataType
Whether you use an S3Prefix or a ManifestFile for the data type. If you choose S3Prefix, S3Uri identifies a key name prefix. Amazon SageMaker uses all objects with the specified key name prefix for the processing job. If you choose ManifestFile, S3Uri identifies an object that is a manifest file containing a list of object keys that you want Amazon SageMaker to use for the processing job.

Type: String
Valid Values: ManifestFile | S3Prefix
Required: Yes

S3InputMode
Whether to use File or Pipe input mode. In File mode, Amazon SageMaker copies the data from the input source onto the local ML storage volume before starting your processing container. This is
the most commonly used input mode. In **Pipe** mode, Amazon SageMaker streams input data from the source directly to your processing container into named pipes without using the ML storage volume.

**Type:** String

**Valid Values:** Pipe | File

**Required:** No

**S3Uri**

The URI of the Amazon S3 prefix Amazon SageMaker downloads data required to run a processing job.

**Type:** String

**Length Constraints:** Maximum length of 1024.

**Pattern:** `^(https|s3)://([^/]+)/?(.*)$`

**Required:** Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingS3Output
Service: Amazon SageMaker Service

Configuration for uploading output data to Amazon S3 from the processing container.

Contents

LocalPath

The local path of a directory where you want Amazon SageMaker to upload its contents to Amazon S3. LocalPath is an absolute path to a directory containing output files. This directory will be created by the platform and exist when your container's entrypoint is invoked.

Type: String
Length Constraints: Maximum length of 256.
Pattern: .*
Required: Yes

S3UploadMode

Whether to upload the results of the processing job continuously or after the job completes.

Type: String
Valid Values: Continuous | EndOfJob
Required: Yes

S3Uri

A URI that identifies the Amazon S3 bucket where you want Amazon SageMaker to save the results of a processing job.

Type: String
Length Constraints: Maximum length of 1024.
Pattern: \b(https|s3)://(([^/]+)\/?(\.*))\b
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProcessingStoppingCondition

Service: Amazon SageMaker Service

Configures conditions under which the processing job should be stopped, such as how long the processing job has been running. After the condition is met, the processing job is stopped.

Contents

MaxRuntimeInSeconds

  Specifies the maximum runtime in seconds.

  Type: Integer


  Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProductionVariant
Service: Amazon SageMaker Service

Identifies a model that you want to host and the resources chosen to deploy for hosting it. If you are deploying multiple models, tell SageMaker how to distribute traffic among the models by specifying variant weights.

Contents

AcceleratorType

The size of the Elastic Inference (EI) instance to use for the production variant. EI instances provide on-demand GPU computing for inference. For more information, see Using Elastic Inference in Amazon SageMaker.

Type: String

Valid Values: ml.eia1.medium | ml.eia1.large | ml.eia1.xlarge | ml.eia2.medium | ml.eia2.large | ml.eia2.xlarge

Required: No

CoreDumpConfig

Specifies configuration for a core dump from the model container when the process crashes.

Type: ProductionVariantCoreDumpConfig (p. 1295) object

Required: No

InitialInstanceCount

Number of instances to launch initially.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

InitialVariantWeight

Determines initial traffic distribution among all of the models that you specify in the endpoint configuration. The traffic to a production variant is determined by the ratio of the VariantWeight to the sum of all VariantWeight values across all ProductionVariants. If unspecified, it defaults to 1.0.

Type: Float

Valid Range: Minimum value of 0.

Required: No

InstanceType

The ML compute instance type.

Type: String

Valid Values: ml.t2.medium | ml.t2.large | ml.t2.xlarge | ml.t2.2xlarge | ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge | ml.m4.16xlarge | ml.m5.large | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge
| ml.m5.12xlarge | ml.m5.24xlarge | ml.m5d.large | ml.m5d.xlarge | ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.12xlarge | ml.m5d.24xlarge | ml.c4.large | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge | ml.c4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge | ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.c5.large | ml.c5.xlarge | ml.c5.2xlarge | ml.c5d.xlarge | ml.c5d.2xlarge | ml.c5d.4xlarge | ml.c5d.12xlarge | ml.c5d.24xlarge | ml.c5d.18xlarge | ml.c5d.9xlarge | ml.c5d.4xlarge | ml.c5d.9xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.r5.large | ml.r5.xlarge | ml.r5.2xlarge | ml.r5.4xlarge | ml.r5.12xlarge | ml.r5.24xlarge | ml.r5d.large | ml.r5d.xlarge | ml.r5d.2xlarge | ml.r5d.4xlarge | ml.r5d.12xlarge | ml.r5d.24xlarge | ml.inf1.xlarge | ml.inf1.2xlarge | ml.inf1.6xlarge | ml.inf1.24xlarge |

**Required:** No

**ModelName**

The name of the model that you want to host. This is the name that you specified when creating the model.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*$`

Required: Yes

**ServerlessConfig**

The serverless configuration for an endpoint. Specifies a serverless endpoint configuration instead of an instance-based endpoint configuration.

Type: `ProductionVariantServerlessConfig (p. 1297)` object

Required: No

**VariantName**

The name of the production variant.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}`

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProductionVariantCoreDumpConfig

Service: Amazon SageMaker Service

Specifies configuration for a core dump from the model container when the process crashes.

Contents

DestinationS3Uri

The Amazon S3 bucket to send the core dump to.

Type: String

Length Constraints: Maximum length of 512.

Pattern: ^(https|s3)://([^/])?/(.*)$

Required: Yes

KmsKeyId

The AWS Key Management Service (AWS KMS) key that SageMaker uses to encrypt the core dump data at rest using Amazon S3 server-side encryption. The KmsKeyId can be any of the following formats:

• // KMS Key ID

"1234abcd-12ab-34cd-56ef-1234567890ab"

• // Amazon Resource Name (ARN) of a KMS Key

"arn:aws:kms:us-west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"

• // KMS Key Alias

"alias/ExampleAlias"

• // Amazon Resource Name (ARN) of a KMS Key Alias


If you use a KMS key ID or an alias of your KMS key, the SageMaker execution role must include permissions to call kms:Encrypt. If you don't provide a KMS key ID, SageMaker uses the default KMS key for Amazon S3 for your role's account. SageMaker uses server-side encryption with KMS-managed keys for OutputDataConfig. If you use a bucket policy with an s3:PutObject permission that only allows objects with server-side encryption, set the condition key of s3:x-amz-server-side-encryption to "aws:kms". For more information, see KMS-Managed Encryption Keys in the Amazon Simple Storage Service Developer Guide.

The KMS key policy must grant permission to the IAM role that you specify in your CreateEndpoint and UpdateEndpoint requests. For more information, see Using Key Policies in AWS KMS in the AWS Key Management Service Developer Guide.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProductionVariantServerlessConfig
Service: Amazon SageMaker Service
Specifies the serverless configuration for an endpoint variant.

Contents

MaxConcurrency

The maximum number of concurrent invocations your serverless endpoint can process.

Type: Integer


Required: Yes

MemorySizeInMB

The memory size of your serverless endpoint. Valid values are in 1 GB increments: 1024 MB, 2048 MB, 3072 MB, 4096 MB, 5120 MB, or 6144 MB.

Type: Integer


Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProductionVariantStatus

Service: Amazon SageMaker Service

Describes the status of the production variant.

Contents

StartTime

The start time of the current status change.

Type: Timestamp

Required: No

Status

The endpoint variant status which describes the current deployment stage status or operational status.

- Creating: Creating inference resources for the production variant.
- Deleting: Terminating inference resources for the production variant.
- Updating: Updating capacity for the production variant.
- ActivatingTraffic: Turning on traffic for the production variant.
- Baking: Waiting period to monitor the CloudWatch alarms in the automatic rollback configuration.

Type: String

Valid Values: Creating | Updating | Deleting | ActivatingTraffic | Baking

Required: Yes

StatusMessage

A message that describes the status of the production variant.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProductionVariantSummary

Service: Amazon SageMaker Service

Describes weight and capacities for a production variant associated with an endpoint. If you sent a request to the UpdateEndpointWeightsAndCapacities API and the endpoint status is Updating, you get different desired and current values.

Contents

CurrentInstanceCount

The number of instances associated with the variant.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

CurrentServerlessConfig

The serverless configuration for the endpoint.
Type: ProductionVariantServerlessConfig (p. 1297) object
Required: No

CurrentWeight

The weight associated with the variant.
Type: Float
Valid Range: Minimum value of 0.
Required: No

DeployedImages

An array of DeployedImage objects that specify the Amazon EC2 Container Registry paths of the inference images deployed on instances of this ProductionVariant.
Type: Array of DeployedImage (p. 986) objects
Required: No

DesiredInstanceCount

The number of instances requested in the UpdateEndpointWeightsAndCapacities request.
Type: Integer
Valid Range: Minimum value of 0.
Required: No

DesiredServerlessConfig

The serverless configuration requested for the endpoint update.
Type: ProductionVariantServerlessConfig (p. 1297) object
Required: No
**DesiredWeight**

The requested weight, as specified in the `UpdateEndpointWeightsAndCapacities` request.

Type: Float

Valid Range: Minimum value of 0.

Required: No

**VariantName**

The name of the variant.

Type: String

Length Constraints: Maximum length of 63.

Pattern: `^[a-zA-Z0-9](\-[a-zA-Z0-9])*([a-zA-Z0-9])\{0,62\}`

Required: Yes

**VariantStatus**

The endpoint variant status which describes the current deployment stage status or operational status.

Type: Array of `ProductionVariantStatus` objects

Array Members: Minimum number of 0 items. Maximum number of 5 items.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProfilerConfig

Service: Amazon SageMaker Service

Configuration information for Debugger system monitoring, framework profiling, and storage paths.

Contents

ProfilingIntervalInMilliseconds

A time interval for capturing system metrics in milliseconds. Available values are 100, 200, 500, 1000 (1 second), 5000 (5 seconds), and 60000 (1 minute) milliseconds. The default value is 500 milliseconds.

Type: Long

Required: No

ProfilingParameters

Configuration information for capturing framework metrics. Available key strings for different profiling options are DetailedProfilingConfig, PythonProfilingConfig, and DataLoaderProfilingConfig. The following codes are configuration structures for the ProfilingParameters parameter. To learn more about how to configure the ProfilingParameters parameter, see Use the SageMaker and Debugger Configuration API Operations to Create, Update, and Debug Your Training Job.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 20 items.

Key Length Constraints: Minimum length of 1. Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 256.

Value Pattern: .*

Required: No

S3OutputPath

Path to Amazon S3 storage location for system and framework metrics.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3):/\/([^/]+\/)+(.*)$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for Ruby V3
ProfilerConfigForUpdate
Service: Amazon SageMaker Service

Configuration information for updating the Debugger profile parameters, system and framework metrics configurations, and storage paths.

Contents

DisableProfiler

To disable Debugger monitoring and profiling, set to True.
Type: Boolean
Required: No

ProfilingIntervalInMilliseconds

A time interval for capturing system metrics in milliseconds. Available values are 100, 200, 500, 1000 (1 second), 5000 (5 seconds), and 60000 (1 minute) milliseconds. The default value is 500 milliseconds.
Type: Long
Required: No

ProfilingParameters

Configuration information for capturing framework metrics. Available key strings for different profiling options are DetailedProfilingConfig, PythonProfilingConfig, and DataLoaderProfilingConfig. The following codes are configuration structures for the ProfilingParameters parameter. To learn more about how to configure the ProfilingParameters parameter, see Use the SageMaker and Debugger Configuration API Operations to Create, Update, and Debug Your Training Job.
Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 20 items.
Key Length Constraints: Minimum length of 1. Maximum length of 256.
Key Pattern: .*
Value Length Constraints: Maximum length of 256.
Value Pattern: .*
Required: No

S3OutputPath

Path to Amazon S3 storage location for system and framework metrics.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: ^(https|s3):/\\([^/\^]+)?(\.*$)
Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProfilerRuleConfiguration
Service: Amazon SageMaker Service

Configuration information for profiling rules.

Contents

InstanceType

The instance type to deploy a Debugger custom rule for profiling a training job.

Type: String

Valid Values: ml.t3.medium | ml.t3.large | ml.t3.xlarge | ml.t3.2xlarge
| ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge
| ml.m4.16xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.m5.2xlarge
| ml.c4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge
| ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.c5.2xlarge
| ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge | ml.m5.large
| ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge
| ml.m5.24xlarge | ml.r5.large | ml.r5.xlarge | ml.r5.2xlarge
| ml.r5.8xlarge | ml.r5.12xlarge | ml.r5.16xlarge | ml.r5.24xlarge
| ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge
| ml.g4dn.12xlarge | ml.g4dn.16xlarge

Required: No

LocalPath

Path to local storage location for output of rules. Defaults to /opt/ml/processing/output/rule/.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: .*

Required: No

RuleConfigurationName

The name of the rule configuration. It must be unique relative to other rule configuration names.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: .*

Required: Yes

RuleEvaluatorImage

The Amazon Elastic Container (ECR) Image for the managed rule evaluation.

Type: String

Length Constraints: Maximum length of 255.

Pattern: .*
Required: Yes

RuleParameters

Runtime configuration for rule container.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Key Length Constraints: Minimum length of 1. Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 256.

Value Pattern: .*

Required: No

S3OutputPath

Path to Amazon S3 storage location for rules.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://(/[^/]+/)?([^/]+)$

Required: No

VolumeSizeInGB

The size, in GB, of the ML storage volume attached to the processing instance.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProfilerRuleEvaluationStatus

Service: Amazon SageMaker Service

Information about the status of the rule evaluation.

Contents

**LastModifiedTime**

Timestamp when the rule evaluation status was last modified.

Type: Timestamp

Required: No

**RuleConfigurationName**

The name of the rule configuration.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: .*

Required: No

**RuleEvaluationJobArn**

The Amazon Resource Name (ARN) of the rule evaluation job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:processing-job/.*

Required: No

**RuleEvaluationStatus**

Status of the rule evaluation.

Type: String

Valid Values: InProgress | NoIssuesFound | IssuesFound | Error | Stopping | Stopped

Required: No

**StatusDetails**

Details from the rule evaluation.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: .*

Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Project
Service: Amazon SageMaker Service

The properties of a project as returned by the Search API.

Contents

CreatedBy
Who created the project.
Type: UserContext (p. 1449) object
Required: No

CreationTime
A timestamp specifying when the project was created.
Type: Timestamp
Required: No

LastModifiedBy
Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
Type: UserContext (p. 1449) object
Required: No

LastModifiedTime
A timestamp container for when the project was last modified.
Type: Timestamp
Required: No

ProjectArn
The Amazon Resource Name (ARN) of the project.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 2048.
Pattern: arn:aws[a-zA-Z\-]*:sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:project:.*
Required: No

ProjectDescription
The description of the project.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: [\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]*
Required: No
**ProjectId**

The ID of the project.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9])*`  

Required: No

**ProjectName**

The name of the project.

Type: String


Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,31}`

Required: No

**ProjectStatus**

The status of the project.

Type: String

Valid Values: `Pending` | `CreateInProgress` | `CreateCompleted` | `CreateFailed` | `DeleteInProgress` | `DeleteFailed` | `DeleteCompleted` | `UpdateInProgress` | `UpdateCompleted` | `UpdateFailed`

Required: No

**ServiceCatalogProvisionedProductDetails**

Details of a provisioned service catalog product. For information about service catalog, see [What is AWS Service Catalog](https://aws.amazon.com/servicecatalog/).

Type: `ServiceCatalogProvisionedProductDetails (p. 1361)` object

Required: No

**ServiceCatalogProvisioningDetails**

Details that you specify to provision a service catalog product. For information about service catalog, see [What is AWS Service Catalog](https://aws.amazon.com/servicecatalog/).

Type: `ServiceCatalogProvisioningDetails (p. 1362)` object

Required: No

**Tags**

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see [Tagging AWS Resources](https://aws.amazon.com/articles/21682/).

Type: Array of `Tag (p. 1377)` objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No
See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProjectSummary
Service: Amazon SageMaker Service

Information about a project.

Contents

CreationTime
The time that the project was created.
Type: Timestamp
Required: Yes

ProjectArn
The Amazon Resource Name (ARN) of the project.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 2048.
Pattern: arn:aws[a-zA-Z-]*:sagemaker:[a-zA-Z0-9-]*:[0-9]{12}:project:.*
Required: Yes

ProjectDescription
The description of the project.
Type: String
Length Constraints: Maximum length of 1024.
Pattern: \[^\p{L}\p{M}\p{Z}\p{S}\p{N}\p{P}]\]*
Required: No

ProjectId
The ID of the project.
Type: String
Pattern: ^[a-zA-Z0-9-]*([a-zA-Z0-9])*
Required: Yes

ProjectName
The name of the project.
Type: String
Pattern: ^[a-zA-Z0-9-]*([a-zA-Z0-9])\{0,31\}
Required: Yes
ProjectStatus

The status of the project.

Type: String

Valid Values: Pending | CreateInProgress | CreateCompleted | CreateFailed | DeleteInProgress | DeleteFailed | DeleteCompleted | UpdateInProgress | UpdateCompleted | UpdateFailed

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PropertyNameQuery

Service: Amazon SageMaker Service

Part of the SuggestionQuery type. Specifies a hint for retrieving property names that begin with the specified text.

Contents

PropertyNameHint

Text that begins a property's name.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100.

Pattern: .*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Property Name Suggestion

Service: Amazon SageMaker Service

A property name returned from a GetSearchSuggestions call that specifies a value in the Property Name Query field.

Contents

Property Name

A suggested property name based on what you entered in the search textbox in the Amazon SageMaker console.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: .+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ProvisioningParameter

Service: Amazon SageMaker Service

A key value pair used when you provision a project as a service catalog product. For information, see What is AWS Service Catalog.

Contents

Key

The key that identifies a provisioning parameter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1000.

Pattern: .*

Required: No

Value

The value of the provisioning parameter.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: .*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
PublicWorkforceTaskPrice
Service: Amazon SageMaker Service

Defines the amount of money paid to an Amazon Mechanical Turk worker for each task performed.

Use one of the following prices for bounding box tasks. Prices are in US dollars and should be based on the complexity of the task; the longer it takes in your initial testing, the more you should offer.

- 0.036
- 0.048
- 0.060
- 0.072
- 0.120
- 0.240
- 0.360
- 0.480
- 0.600
- 0.720
- 0.840
- 0.960
- 1.080
- 1.200

Use one of the following prices for image classification, text classification, and custom tasks. Prices are in US dollars.

- 0.012
- 0.024
- 0.036
- 0.048
- 0.060
- 0.072
- 0.120
- 0.240
- 0.360
- 0.480
- 0.600
- 0.720
- 0.840
- 0.960
- 1.080
- 1.200

Use one of the following prices for semantic segmentation tasks. Prices are in US dollars.

- 0.840
- 0.960
- 1.080
Use one of the following prices for Textract AnalyzeDocument Important Form Key Amazon Augmented AI review tasks. Prices are in US dollars.

- 1.200
- 2.400
- 2.280
- 2.160
- 2.040
- 1.920
- 1.800
- 1.680
- 1.560
- 1.440
- 1.320
- 1.200
- 1.080
- 0.960
- 0.840
- 0.720
- 0.600
- 0.480
- 0.360
- 0.240
- 0.120
- 0.072
- 0.060
- 0.048
- 0.036
- 0.024
- 0.012

Use one of the following prices for Rekognition DetectModerationLabels Amazon Augmented AI review tasks. Prices are in US dollars.

- 1.200
- 1.080
- 0.960
- 0.840
- 0.720
- 0.600
- 0.480
- 0.360
- 0.240
- 0.120
- 0.072
- 0.060
Use one of the following prices for Amazon Augmented AI custom human review tasks. Prices are in US dollars.

- 1.200
- 1.080
- 0.960
- 0.840
- 0.720
- 0.600
- 0.480
- 0.360
- 0.240
- 0.120
- 0.072
- 0.060
- 0.048
- 0.036
- 0.024
- 0.012

Contents

AmountInUsd

Defines the amount of money paid to an Amazon Mechanical Turk worker in United States dollars.

Type: USD (p. 1448) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
QualityCheckStepMetadata

Service: Amazon SageMaker Service

Container for the metadata for a Quality check step. For more information, see the topic on QualityCheck step in the Amazon SageMaker Developer Guide.

Contents

BaselineUsedForDriftCheckConstraints

The Amazon S3 URI of the baseline constraints file used for the drift check.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

BaselineUsedForDriftCheckStatistics

The Amazon S3 URI of the baseline statistics file used for the drift check.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

CalculatedBaselineConstraints

The Amazon S3 URI of the newly calculated baseline constraints file.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

CalculatedBaselineStatistics

The Amazon S3 URI of the newly calculated baseline statistics file.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

CheckJobArn

The Amazon Resource Name (ARN) of the Quality check processing job that was run by this step execution.

Type: String

Length Constraints: Maximum length of 256.

Required: No

CheckType

The type of the Quality check step.

Type: String
Length Constraints: Maximum length of 256.
Required: No

**ModelPackageName**

The model package group name.
Type: String
Length Constraints: Maximum length of 256.
Required: No

**RegisterNewBaseline**

This flag indicates if a newly calculated baseline can be accessed through step properties `BaselineUsedForDriftCheckConstraints` and `BaselineUsedForDriftCheckStatistics`. If it is set to `False`, the previous baseline of the configured check type must also be available. These can be accessed through the `BaselineUsedForDriftCheckConstraints` and `BaselineUsedForDriftCheckStatistics` properties.

Type: Boolean
Required: No

**SkipCheck**

This flag indicates if the drift check against the previous baseline will be skipped or not. If it is set to `False`, the previous baseline of the configured check type must be available.

Type: Boolean
Required: No

**ViolationReport**

The Amazon S3 URI of violation report if violations are detected.

Type: String
Length Constraints: Maximum length of 1024.
Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
QueryFilters

Service: Amazon SageMaker Service

A set of filters to narrow the set of lineage entities connected to the `StartArn(s)` returned by the `QueryLineage` API action.

Contents

**CreatedAfter**

Filter the lineage entities connected to the `StartArn(s)` after the create date.

Type: Timestamp

Required: No

**CreatedBefore**

Filter the lineage entities connected to the `StartArn(s)` by created date.

Type: Timestamp

Required: No

**LineageTypes**

Filter the lineage entities connected to the `StartArn(s)` by the type of the lineage entity.

Type: Array of strings

Array Members: Maximum number of 4 items.

Valid Values: TrialComponent | Artifact | Context | Action

Required: No

**ModifiedAfter**

Filter the lineage entities connected to the `StartArn(s)` after the last modified date.

Type: Timestamp

Required: No

**ModifiedBefore**

Filter the lineage entities connected to the `StartArn(s)` before the last modified date.

Type: Timestamp

Required: No

**Properties**

Filter the lineage entities connected to the `StartArn(s)` by a set if property key value pairs. If multiple pairs are provided, an entity is included in the results if it matches any of the provided pairs.

Type: String to string map

Map Entries: Maximum number of 5 items.

Key Length Constraints: Maximum length of 256.

Value Length Constraints: Maximum length of 256.
Required: No

**Types**

Filter the lineage entities connected to the `StartArn` by type. For example: `DataSet`, `Model`, `Endpoint`, or `ModelDeployment`.

Type: Array of strings

Array Members: Maximum number of 5 items.

Length Constraints: Maximum length of 40.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**RecommendationJobCompiledOutputConfig**

Service: Amazon SageMaker Service

Provides information about the output configuration for the compiled model.

**Contents**

**S3OutputUri**

Identifies the Amazon S3 bucket where you want SageMaker to store the compiled model artifacts.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: `^(https|s3)://([^/]+)/?([^/]+)$`

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**RecommendationJobInputConfig**

Service: Amazon SageMaker Service

The input configuration of the recommendation job.

**Contents**

**EndpointConfigurations**

Specifies the endpoint configuration to use for a job.

Type: Array of `EndpointInputConfiguration (p. 1019)` objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Required: No

**JobDurationInSeconds**

Specifies the maximum duration of the job, in seconds.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

**ModelPackageVersionArn**

The Amazon Resource Name (ARN) of a versioned model package.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:model-package/.*`

Required: Yes

**ResourceLimit**

Defines the resource limit of the job.

Type: `RecommendationJobResourceLimit (p. 1328)` object

Required: No

**TrafficPattern**

Specifies the traffic pattern of the job.

Type: `TrafficPattern (p. 1382)` object

Required: No

**VolumeKmsKeyId**

The Amazon Resource Name (ARN) of an AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt data on the storage volume attached to the ML compute instance that hosts the endpoint. This key will be passed to SageMaker Hosting for endpoint creation.

The SageMaker execution role must have `kms:CreateGrant` permission in order to encrypt data on the storage volume of the endpoints created for inference recommendation. The inference
recommendation job will fail asynchronously during endpoint configuration creation if the role passed does not have `kms:CreateGrant` permission.

The `KmsKeyId` can be any of the following formats:

- // KMS Key ID
  
  "1234abcd-12ab-34cd-56ef-1234567890ab"

- // Amazon Resource Name (ARN) of a KMS Key
  
  "arn:aws:kms:<region>:<account>:key/<key-id-12ab-34cd-56ef-1234567890ab>"

- // KMS Key Alias
  
  "alias/ExampleAlias"

- // Amazon Resource Name (ARN) of a KMS Key Alias
  
  "arn:aws:kms:<region>:<account>:alias/<ExampleAlias>"

For more information about key identifiers, see `Key identifiers (KeyID)` in the AWS Key Management Service (AWS KMS) documentation.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RecommendationJobOutputConfig

Service: Amazon SageMaker Service

Provides information about the output configuration for the compiled model.

Contents

CompiledOutputConfig

Provides information about the output configuration for the compiled model.

Type: RecommendationJobCompiledOutputConfig (p. 1324) object

Required: No

KmsKeyId

The Amazon Resource Name (ARN) of a AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt your output artifacts with Amazon S3 server-side encryption. The SageMaker execution role must have kms:GenerateDataKey permission.

The KmsKeyId can be any of the following formats:

- // KMS Key ID
  "1234abcd-12ab-34cd-56ef-1234567890ab"
- // Amazon Resource Name (ARN) of a KMS Key
  "arn:aws:kms:<region>:<account>:key/<key-id-12ab-34cd-56ef-1234567890ab>"
- // KMS Key Alias
  "alias/ExampleAlias"
- // Amazon Resource Name (ARN) of a KMS Key Alias
  "arn:aws:kms:<region>:<account>:alias/<ExampleAlias>"

For more information about key identifiers, see Key identifiers (KeyID) in the AWS Key Management Service (AWS KMS) documentation.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**RecommendationJobResourceLimit**

Service: Amazon SageMaker Service

Specifies the maximum number of jobs that can run in parallel and the maximum number of jobs that can run.

**Contents**

**MaxNumberOfTests**

Defines the maximum number of load tests.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

**MaxParallelOfTests**

Defines the maximum number of parallel load tests.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RecommendationJobStoppingConditions

Service: Amazon SageMaker Service

Specifies conditions for stopping a job. When a job reaches a stopping condition limit, SageMaker ends the job.

Contents

MaxInvocations

The maximum number of requests per minute expected for the endpoint.

Type: Integer

Required: No

ModelLatencyThresholds

The interval of time taken by a model to respond as viewed from SageMaker. The interval includes the local communication time taken to send the request and to fetch the response from the container of a model and the time taken to complete the inference in the container.

Type: Array of ModelLatencyThreshold (p. 1163) objects

Array Members: Fixed number of 1 item.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RecommendationMetrics
Service: Amazon SageMaker Service
The metrics of recommendations.

Contents

CostPerHour
 Defines the cost per hour for the instance.
 Type: Float
 Required: Yes

CostPerInference
 Defines the cost per inference for the instance.
 Type: Float
 Required: Yes

MaxInvocations
 The expected maximum number of requests per minute for the instance.
 Type: Integer
 Required: Yes

ModelLatency
 The expected model latency at maximum invocation per minute for the instance.
 Type: Integer
 Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RedshiftDatasetDefinition

Service: Amazon SageMaker Service

Configuration for Redshift Dataset Definition input.

Contents

ClusterId

The Redshift cluster Identifier.

Type: String


Pattern: .*

Required: Yes

ClusterRoleArn

The IAM role attached to your Redshift cluster that Amazon SageMaker uses to generate datasets.

Type: String


Pattern: ^arn:aws[a-zA-z\-]*:iam::\d{12}:role/?[a-zA-z0-9+=,.@\-_\/]+$

Required: Yes

Database

The name of the Redshift database used in Redshift query execution.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: .*

Required: Yes

DbUser

The database user name used in Redshift query execution.

Type: String


Pattern: .*

Required: Yes

KmsKeyId

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt data from a Redshift execution.

Type: String

Length Constraints: Maximum length of 2048.
Pattern: .*
Required: No

**OutputCompression**

The compression used for Redshift query results.

Type: String

Valid Values: None | GZIP | BZIP2 | ZSTD | SNAPPY

Required: No

**OutputFormat**

The data storage format for Redshift query results.

Type: String

Valid Values: PARQUET | CSV

Required: Yes

**OutputS3Uri**

The location in Amazon S3 where the Redshift query results are stored.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^https?://([^/]+)/(.*).*$

Required: Yes

**QueryString**

The SQL query statements to be executed.

Type: String


Pattern: [\s\S]+

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RegisterModelStepMetadata
Service: Amazon SageMaker Service
Metadata for a register model job step.

Contents

Arn
The Amazon Resource Name (ARN) of the model package.
Type: String
Length Constraints: Maximum length of 256.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RenderableTask
Service: Amazon SageMaker Service

Contains input values for a task.

Contents

Input

A JSON object that contains values for the variables defined in the template. It is made available to the template under the substitution variable `task.input`. For example, if you define a variable `task.input.text` in your template, you can supply the variable in the JSON object as "text": "sample text".

Type: String


Pattern: \[\S\S\]+

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RenderingError
Service: Amazon SageMaker Service

A description of an error that occurred while rendering the template.

Contents

Code
A unique identifier for a specific class of errors.
Type: String
Required: Yes

Message
A human-readable message describing the error.
Type: String
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RepositoryAuthConfig

Service: Amazon SageMaker Service

Specifies an authentication configuration for the private docker registry where your model image is hosted. Specify a value for this property only if you specified `Vpc` as the value for the `RepositoryAccessMode` field of the `ImageConfig` object that you passed to a call to `CreateModel` and the private Docker registry where the model image is hosted requires authentication.

Contents

RepositoryCredentialsProviderArn

The Amazon Resource Name (ARN) of an AWS Lambda function that provides credentials to authenticate to the private Docker registry where your model image is hosted. For information about how to create an AWS Lambda function, see Create a Lambda function with the console in the AWS Lambda Developer Guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: .*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ResolvedAttributes
Service: Amazon SageMaker Service
The resolved attributes.

Contents

AutoMLJobObjective
Specifies a metric to minimize or maximize as the objective of a job.
Type: AutoMLJobObjective (p. 917) object
Required: No

CompletionCriteria
How long a job is allowed to run, or how many candidates a job is allowed to generate.
Type: AutoMLJobCompletionCriteria (p. 915) object
Required: No

ProblemType
The problem type.
Type: String
Valid Values: BinaryClassification | MulticlassClassification | Regression
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ResourceConfig

Service: Amazon SageMaker Service

Describes the resources, including ML compute instances and ML storage volumes, to use for model training.

Contents

InstanceCount

The number of ML compute instances to use. For distributed training, provide a value greater than 1.

Type: Integer

Valid Range: Minimum value of 1.

Required: Yes

InstanceType

The ML compute instance type.

Type: String

Valid Values: ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge | ml.m4.16xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.m5.large | ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge | ml.m5.24xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge | ml.c4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge | ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.p3dn.24xlarge | ml.p4d.24xlarge | ml.c5.xlarge | ml.c5.2xlarge | ml.c5.4xlarge | ml.c5.9xlarge | ml.c5.18xlarge | ml.c5n.xlarge | ml.c5n.2xlarge | ml.c5n.4xlarge | ml.c5n.9xlarge | ml.c5n.18xlarge | ml.g5.xlarge | ml.g5.2xlarge | ml.g5.4xlarge | ml.g5.8xlarge | ml.g5.16xlarge | ml.g5.12xlarge | ml.g5.24xlarge | ml.g5.48xlarge

Required: Yes

VolumeKmsKeyId

The AWS KMS key that SageMaker uses to encrypt data on the storage volume attached to the ML compute instance(s) that run the training job.

Note

Certain Nitro-based instances include local storage, dependent on the instance type. Local storage volumes are encrypted using a hardware module on the instance. You can't request a VolumeKmsKeyId when using an instance type with local storage.

For a list of instance types that support local instance storage, see Instance Store Volumes.

For more information about local instance storage encryption, see SSD Instance Store Volumes.

The VolumeKmsKeyId can be in any of the following formats:

- // KMS Key ID

  "1234abcd-12ab-34cd-56ef-1234567890ab"

- // Amazon Resource Name (ARN) of a KMS Key

  "arn:aws:kms:us-west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
Type: String

Length Constraints: Maximum length of 2048.

Pattern: . *

Required: No

VolumeSizeInGB

The size of the ML storage volume that you want to provision.

ML storage volumes store model artifacts and incremental states. Training algorithms might also use the ML storage volume for scratch space. If you want to store the training data in the ML storage volume, choose File as the TrainingInputMode in the algorithm specification.

You must specify sufficient ML storage for your scenario.

Note
SageMaker supports only the General Purpose SSD (gp2) ML storage volume type.

Note
Certain Nitro-based instances include local storage with a fixed total size, dependent on the instance type. When using these instances for training, SageMaker mounts the local instance storage instead of Amazon EBS gp2 storage. You can't request a VolumeSizeInGB greater than the total size of the local instance storage.

For a list of instance types that support local instance storage, including the total size per instance type, see Instance Store Volumes.

Type: Integer

Valid Range: Minimum value of 1.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ResourceLimits
Service: Amazon SageMaker Service

Specifies the maximum number of training jobs and parallel training jobs that a hyperparameter tuning job can launch.

Contents

MaxNumberOfTrainingJobs

The maximum number of training jobs that a hyperparameter tuning job can launch.

Type: Integer

Valid Range: Minimum value of 1.

Required: Yes

MaxParallelTrainingJobs

The maximum number of concurrent training jobs that a hyperparameter tuning job can launch.

Type: Integer

Valid Range: Minimum value of 1.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ResourceSpec

Service: Amazon SageMaker Service

Specifies the ARN’s of a SageMaker image and SageMaker image version, and the instance type that the version runs on.

Contents

InstanceType

The instance type that the image version runs on.

Note

JupyterServer apps only support the system value.
For KernelGateway apps, the system value is translated to ml.t3.medium. KernelGateway apps also support all other values for available instance types.

Type: String

Valid Values: system | ml.t3.micro | ml.t3.small | ml.t3.medium | ml.t3.large
| ml.t3.xlarge | ml.t3.2xlarge | ml.m5.large | ml.m5.xlarge | ml.m5.2xlarge
| ml.m5.4xlarge | ml.m5.8xlarge | ml.m5.12xlarge | ml.m5.16xlarge
| ml.m5.24xlarge | ml.m5d.large | ml.m5d.xlarge | ml.m5d.2xlarge | ml.m5d.4xlarge | ml.m5d.8xlarge | ml.m5d.12xlarge | ml.m5d.16xlarge | ml.m5d.24xlarge | ml.c5.large | ml.c5.xlarge | ml.c5.2xlarge | ml.c5.4xlarge
| ml.c5.9xlarge | ml.c5.12xlarge | ml.c5.18xlarge | ml.c5.24xlarge |
| ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.p3dn.24xlarge |
| ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge | ml.g4dn.8xlarge |
| ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.r5.large | ml.r5.xlarge
| ml.r5.2xlarge | ml.r5.4xlarge | ml.r5.8xlarge | ml.r5.12xlarge |
| ml.r5.16xlarge | ml.r5.24xlarge | ml.g5.large | ml.g5.xlarge |
| ml.g5.4xlarge | ml.g5.8xlarge | ml.g5.16xlarge | ml.g5.12xlarge |
| ml.g5.24xlarge | ml.g5.48xlarge |

Required: No

LifecycleConfigArn

The Amazon Resource Name (ARN) of the Lifecycle Configuration attached to the Resource.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[-a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:studio-lifecycle-config/.*

Required: No

SageMakerImageArn

The ARN of the SageMaker image that the image version belongs to.

Type: String

Length Constraints: Maximum length of 256.

Pattern: ^arn:aws(-[\w]+)*:sagemaker:.+::[0-9]{12}:image/[a-z0-9\-][\-]?[a-z0-9\-]*/$

Required: No
SageMakerImageVersionArn

The ARN of the image version created on the instance.

Type: String

Length Constraints: Maximum length of 256.

Pattern: \^arn:aws(-[\w]+)*:sagemaker:.+:[0-9]{12}:image-version/[a-z0-9][-.]?[a-z0-9]*/[0-9]+$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RetentionPolicy
Service: Amazon SageMaker Service

The retention policy for data stored on an Amazon Elastic File System (EFS) volume.

Contents

HomeEfsFileSystem

The default is Retain, which specifies to keep the data stored on the EFS volume.

Specify Delete to delete the data stored on the EFS volume.

Type: String

Valid Values: Retain | Delete

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
RetryStrategy

Service: Amazon SageMaker Service

The retry strategy to use when a training job fails due to an InternalServerError. RetryStrategy is specified as part of the CreateTrainingJob and CreateHyperParameterTuningJob requests. You can add the StoppingCondition parameter to the request to limit the training time for the complete job.

Contents

MaximumRetryAttempts

The number of times to retry the job. When the job is retried, its SecondaryStatus is changed to STARTING.

Type: Integer


Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RSessionAppSettings
Service: Amazon SageMaker Service

A collection of settings that apply to an RSessionGateway app.

Contents

CustomImages
A list of custom SageMaker images that are configured to run as a RSession app.

Type: Array of CustomImage (p. 965) objects

Array Members: Maximum number of 200 items.

Required: No

DefaultResourceSpec
Specifies the ARN's of a SageMaker image and SageMaker image version, and the instance type that the version runs on.

Type: ResourceSpec (p. 1341) object

Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RStudioServerProAppSettings

Service: Amazon SageMaker Service

A collection of settings that configure user interaction with the RStudioServerPro app. RStudioServerProAppSettings cannot be updated. The RStudioServerPro app must be deleted and a new one created to make any changes.

Contents

AccessStatus

Indicates whether the current user has access to the RStudioServerPro app.

Type: String

Valid Values: ENABLED | DISABLED

Required: No

UserGroup

The level of permissions that the user has within the RStudioServerPro app. This value defaults to `User`. The `Admin` value allows the user access to the RStudio Administrative Dashboard.

Type: String

Valid Values: R_STUDIO_ADMIN | R_STUDIO_USER

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**RStudioServerProDomainSettings**

Service: Amazon SageMaker Service

A collection of settings that configure the RStudioServerPro Domain-level app.

## Contents

**DefaultResourceSpec**

Specifies the ARN's of a SageMaker image and SageMaker image version, and the instance type that the version runs on.

Type: ResourceSpec (p. 1341) object

Required: No

**DomainExecutionRoleArn**

The ARN of the execution role for the RStudioServerPro Domain-level app.

Type: String


Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@-_/]+$

Required: Yes

**RStudioConnectUrl**

A URL pointing to an RStudio Connect server.

Type: String

Required: No

**RStudioPackageManagerUrl**

A URL pointing to an RStudio Package Manager server.

Type: String

Required: No

## See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
RStudioServerProDomainSettingsForUpdate

Service: Amazon SageMaker Service

A collection of settings that update the current configuration for the RStudioServerPro Domain-level app.

Contents

DefaultResourceSpec

Specifies the ARN's of a SageMaker image and SageMaker image version, and the instance type that the version runs on.

Type: ResourceSpec (p. 1341) object

Required: No

DomainExecutionRoleArn

The execution role for the RStudioServerPro Domain-level app.

Type: String


Pattern: ^arn:aws[a-zA-Z\-]*:iam::\d{12}:role/\?[a-zA-Z0-9+=,.@-_]+\+$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**S3DataSource**

Service: Amazon SageMaker Service

Describes the S3 data source.

## Contents

**AttributeNames**

A list of one or more attribute names to use that are found in a specified augmented manifest file.

Type: Array of strings

Array Members: Maximum number of 16 items.

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: .+

Required: No

**S3DataDistributionType**

If you want SageMaker to replicate the entire dataset on each ML compute instance that is launched for model training, specify `FullyReplicated`.

If you want SageMaker to replicate a subset of data on each ML compute instance that is launched for model training, specify `ShardedByS3Key`. If there are `n` ML compute instances launched for a training job, each instance gets approximately `1/n` of the number of S3 objects. In this case, model training on each machine uses only the subset of training data.

Don't choose more ML compute instances for training than available S3 objects. If you do, some nodes won't get any data and you will pay for nodes that aren't getting any training data. This applies in both File and Pipe modes. Keep this in mind when developing algorithms.

In distributed training, where you use multiple ML compute EC2 instances, you might choose `ShardedByS3Key`. If the algorithm requires copying training data to the ML storage volume (when `TrainingInputMode` is set to `File`), this copies `1/n` of the number of objects.

Type: String

Valid Values: `FullyReplicated` | `ShardedByS3Key`

Required: No

**S3DataType**

If you choose `S3Prefix`, `S3Uri` identifies a key name prefix. SageMaker uses all objects that match the specified key name prefix for model training.

If you choose `ManifestFile`, `S3Uri` identifies an object that is a manifest file containing a list of object keys that you want SageMaker to use for model training.

If you choose `AugmentedManifestFile`, `S3Uri` identifies an object that is an augmented manifest file in JSON lines format. This file contains the data you want to use for model training. `AugmentedManifestFile` can only be used if the Channel's input mode is `Pipe`.

Type: String

Valid Values: `ManifestFile` | `S3Prefix` | `AugmentedManifestFile`
S3DataSource

 Required: Yes

**S3Uri**

Depending on the value specified for the `S3DataType`, identifies either a key name prefix or a manifest. For example:

- A key name prefix might look like this: `s3://bucketname/exampleprefix`
- A manifest might look like this: `s3://bucketname/example.manifest`

A manifest is an S3 object which is a JSON file consisting of an array of elements. The first element is a prefix which is followed by one or more suffixes. SageMaker appends the suffix elements to the prefix to get a full set of `S3Uri`. Note that the prefix must be a valid non-empty `S3Uri` that precludes users from specifying a manifest whose individual `S3Uri` is sourced from different S3 buckets.

The following code example shows a valid manifest format:

```json
[ {
  "prefix": "s3://customer_bucket/some/prefix/",
  "relative/path/to/custdata-1",
  "relative/path/custdata-2",
  ...
  "relative/path/custdata-N"
}
]
```

This JSON is equivalent to the following `S3Uri` list:

- `s3://customer_bucket/some/prefix/relative/path/to/custdata-1`
- `s3://customer_bucket/some/prefix/relative/path/custdata-2`
- ...
- `s3://customer_bucket/some/prefix/relative/path/custdata-N`

The complete set of `S3Uri` in this manifest is the input data for the channel for this data source. The object that each `S3Uri` points to must be readable by the IAM role that SageMaker uses to perform tasks on your behalf.

**Type:** String

**Length Constraints:** Maximum length of 1024.

**Pattern:** `^(https|s3)://([^/]+)/?(.*)$`

**Required:** Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
S3StorageConfig

Service: Amazon SageMaker Service

The Amazon Simple Storage (Amazon S3) location and and security configuration for OfflineStore.

Contents

KmsKeyId

The AWS Key Management Service (KMS) key ID of the key used to encrypt any objects written into the OfflineStore S3 location.

The IAM roleARN that is passed as a parameter to CreateFeatureGroup must have below permissions to the KmsKeyId:

- "kms:GenerateDataKey"

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

ResolvedOutputS3Uri

The S3 path where offline records are written.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^([https|s3])://([^/]+)/?(.*)$

Required: No

S3Uri

The S3 URI, or location in Amazon S3, of OfflineStore.

S3 URIs have a format similar to the following: s3://example-bucket/prefix/.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^([https|s3])://([^/]+)/?(.*)$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ScheduleConfig
Service: Amazon SageMaker Service

Configuration details about the monitoring schedule.

Contents

ScheduleExpression

A cron expression that describes details about the monitoring schedule.

Currently the only supported cron expressions are:

- If you want to set the job to start every hour, please use the following:
  
  **Hourly:** cron(0 * ? * *)

- If you want to start the job daily:
  
  cron(0 [00-23] ? * * *)

For example, the following are valid cron expressions:

- Daily at noon UTC: cron(0 12 ? * *)
- Daily at midnight UTC: cron(0 0 ? * *)

To support running every 6, 12 hours, the following are also supported:

cron(0 [00-23]/[01-24] ? * * *)

For example, the following are valid cron expressions:

- Every 12 hours, starting at 5pm UTC: cron(0 17/12 ? * *)
- Every two hours starting at midnight: cron(0 0/2 ? * *)

**Note**

- Even though the cron expression is set to start at 5PM UTC, note that there could be a delay of 0-20 minutes from the actual requested time to run the execution.
- We recommend that if you would like a daily schedule, you do not provide this parameter. Amazon SageMaker will pick a time for running every day.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**SearchExpression**

*Service: Amazon SageMaker Service*

A multi-expression that searches for the specified resource or resources in a search. All resource objects that satisfy the expression's condition are included in the search results. You must specify at least one subexpression, filter, or nested filter. A `SearchExpression` can contain up to twenty elements.

A `SearchExpression` contains the following components:

- A list of `Filter` objects. Each filter defines a simple Boolean expression comprised of a resource property name, Boolean operator, and value.
- A list of `NestedFilter` objects. Each nested filter defines a list of Boolean expressions using a list of resource properties. A nested filter is satisfied if a single object in the list satisfies all Boolean expressions.
- A list of `SearchExpression` objects. A search expression object can be nested in a list of search expression objects.
- A Boolean operator: `And` or `Or`.

### Contents

#### Filters

A list of filter objects.

Type: Array of `Filter (p. 1046)` objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Required: No

#### NestedFilters

A list of nested filter objects.

Type: Array of `NestedFilters (p. 1222)` objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Required: No

#### Operator

A Boolean operator used to evaluate the search expression. If you want every conditional statement in all lists to be satisfied for the entire search expression to be true, specify `And`. If only a single conditional statement needs to be true for the entire search expression to be true, specify `Or`. The default value is `And`.

Type: String

Valid Values: `And` | `Or`

Required: No

#### SubExpressions

A list of search expression objects.

Type: Array of `SearchExpression (p. 1355)` objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
SearchRecord
Service: Amazon SageMaker Service

A single resource returned as part of the Search API response.

Contents

Endpoint
A hosted endpoint for real-time inference.
Type: Endpoint object
Required: No

Experiment
The properties of an experiment.
Type: Experiment object
Required: No

FeatureGroup
Amazon SageMaker Feature Store stores features in a collection called Feature Group. A Feature Group can be visualized as a table which has rows, with a unique identifier for each row where each column in the table is a feature. In principle, a Feature Group is composed of features and values per features.
Type: FeatureGroup object
Required: No

ModelPackage
A versioned model that can be deployed for SageMaker inference.
Type: ModelPackage object
Required: No

ModelPackageGroup
A group of versioned models in the model registry.
Type: ModelPackageGroup object
Required: No

Pipeline
A SageMaker Model Building Pipeline instance.
Type: Pipeline object
Required: No

PipelineExecution
An execution of a pipeline.
Type: PipelineExecution object
Required: No
Project

The properties of a project.

Type: Project (p. 1309) object

Required: No

TrainingJob

The properties of a training job.

Type: TrainingJob (p. 1385) object

Required: No

Trial

The properties of a trial.

Type: Trial (p. 1421) object

Required: No

TrialComponent

The properties of a trial component.

Type: TrialComponent (p. 1424) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
SecondaryStatusTransition
Service: Amazon SageMaker Service

An array element of DescribeTrainingJob:SecondaryStatusTransitions (p. 456). It provides additional details about a status that the training job has transitioned through. A training job can be in one of several states, for example, starting, downloading, training, or uploading. Within each state, there are a number of intermediate states. For example, within the starting state, SageMaker could be starting the training job or launching the ML instances. These transitional states are referred to as the job's secondary status.

Contents

EndTime
A timestamp that shows when the training job transitioned out of this secondary status state into another secondary status state or when the training job has ended.

Type: Timestamp
Required: No

StartTime
A timestamp that shows when the training job transitioned to the current secondary status state.

Type: Timestamp
Required: Yes

Status
Contains a secondary status information from a training job.

Status might be one of the following secondary statuses:

InProgress
• Starting - Starting the training job.
• Downloading - An optional stage for algorithms that support File training input mode. It indicates that data is being downloaded to the ML storage volumes.
• Training - Training is in progress.
• Uploading - Training is complete and the model artifacts are being uploaded to the S3 location.

Completed
• Completed - The training job has completed.

Failed
• Failed - The training job has failed. The reason for the failure is returned in the FailureReason field of DescribeTrainingJobResponse.

Stopped
• MaxRuntimeExceeded - The job stopped because it exceeded the maximum allowed runtime.
• Stopped - The training job has stopped.

Stopping
• Stopping - Stopping the training job.

We no longer support the following secondary statuses:
• LaunchingMLInstances
• PreparingTrainingStack
• DownloadingTrainingImage

Type: String

Valid Values: Starting | LaunchingMLInstances | PreparingTrainingStack | Downloading | DownloadingTrainingImage | Training | Uploading | Stopping | Stopped | MaxRuntimeExceeded | Completed | Failed | Interrupted | MaxWaitTimeExceeded | Updating | Restarting

Required: Yes

**StatusMessage**

A detailed description of the progress within a secondary status.

SageMaker provides secondary statuses and status messages that apply to each of them:

**Starting**
- Starting the training job.
- Launching requested ML instances.
- Insufficient capacity error from EC2 while launching instances, retrying!
- Launched instance was unhealthy, replacing it!
- Preparing the instances for training.

**Training**
- Downloading the training image.
- Training image download completed. Training in progress.

**Important**
Status messages are subject to change. Therefore, we recommend not including them in code that programmatically initiates actions. For examples, don’t use status messages in if statements.

To have an overview of your training job’s progress, view TrainingJobStatus and SecondaryStatus in DescribeTrainingJob (p. 448), and StatusMessage together. For example, at the start of a training job, you might see the following:

- TrainingJobStatus - InProgress
- SecondaryStatus - Training
- StatusMessage - Downloading the training image

Type: String

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ServiceCatalogProvisionedProductDetails

Service: Amazon SageMaker Service

Details of a provisioned service catalog product. For information about service catalog, see What is AWS Service Catalog.

Contents

ProvisionedProductId

The ID of the provisioned product.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: ^[a-zA-Z0-9_\-]*

Required: No

ProvisionedProductStatusMessage

The current status of the product.

- AVAILABLE - Stable state, ready to perform any operation. The most recent operation succeeded and completed.
- UNDER_CHANGE - Transitive state. Operations performed might not have valid results. Wait for an AVAILABLE status before performing operations.
- TAINTED - Stable state, ready to perform any operation. The stack has completed the requested operation but is not exactly what was requested. For example, a request to update to a new version failed and the stack rolled back to the current version.
- ERROR - An unexpected error occurred. The provisioned product exists but the stack is not running. For example, CloudFormation received a parameter value that was not valid and could not launch the stack.
- PLAN_IN_PROGRESS - Transitive state. The plan operations were performed to provision a new product, but resources have not yet been created. After reviewing the list of resources to be created, execute the plan. Wait for an AVAILABLE status before performing operations.

Type: String

Pattern: .*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ServiceCatalogProvisioningDetails

Service: Amazon SageMaker Service

Details that you specify to provision a service catalog product. For information about service catalog, see What is AWS Service Catalog.

Contents

PathId

The path identifier of the product. This value is optional if the product has a default path, and required if the product has more than one path.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: ^[a-zA-Z0-9_\-]*

Required: No

ProductId

The ID of the product to provision.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: ^[a-zA-Z0-9_\-]*

Required: Yes

ProvisioningArtifactId

The ID of the provisioning artifact.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: ^[a-zA-Z0-9_\-]*

Required: No

ProvisioningParameters

A list of key value pairs that you specify when you provision a product.

Type: Array of ProvisioningParameter (p. 1316) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for Ruby V3
ServiceCatalogProvisioningUpdateDetails

Service: Amazon SageMaker Service

Details that you specify to provision a service catalog product. For information about service catalog, see What is AWS Service Catalog.

Contents

ProvisioningArtifactId

The ID of the provisioning artifact.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: ^[a-zA-Z0-9_\-]*$

Required: No

ProvisioningParameters

A list of key value pairs that you specify when you provision a product.

Type: Array of ProvisioningParameter (p. 1316) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
SharingSettings
Service: Amazon SageMaker Service

Specifies options for sharing SageMaker Studio notebooks. These settings are specified as part of DefaultUserSettings when the CreateDomain API is called, and as part of UserSettings when the CreateUserProfile API is called. When SharingSettings is not specified, notebook sharing isn't allowed.

Contents

NotebookOutputOption

Whether to include the notebook cell output when sharing the notebook. The default is Disabled.

Type: String

Valid Values: Allowed | Disabled

Required: No

S3KmsKeyId

When NotebookOutputOption is Allowed, the AWS Key Management Service (KMS) encryption key ID used to encrypt the notebook cell output in the Amazon S3 bucket.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No

S3OutputPath

When NotebookOutputOption is Allowed, the Amazon S3 bucket used to store the shared notebook snapshots.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3):/(/[^/]+)?([^/]+)?(.*)$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
ShuffleConfig

Service: Amazon SageMaker Service

A configuration for a shuffle option for input data in a channel. If you use S3Prefix for S3DataType, the results of the S3 key prefix matches are shuffled. If you use ManifestFile, the order of the S3 object references in the ManifestFile is shuffled. If you use AugmentedManifestFile, the order of the JSON lines in the AugmentedManifestFile is shuffled. The shuffling order is determined using the Seed value.

For Pipe input mode, when ShuffleConfig is specified shuffling is done at the start of every epoch. With large datasets, this ensures that the order of the training data is different for each epoch, and it helps reduce bias and possible overfitting. In a multi-node training job when ShuffleConfig is combined with S3DataDistributionType of ShardedByS3Key, the data is shuffled across nodes so that the content sent to a particular node on the first epoch might be sent to a different node on the second epoch.

Contents

Seed

Determines the shuffling order in ShuffleConfig value.

Type: Long

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
SourceAlgorithm

Service: Amazon SageMaker Service

Specifies an algorithm that was used to create the model package. The algorithm must be either an algorithm resource in your SageMaker account or an algorithm in AWS Marketplace that you are subscribed to.

Contents

AlgorithmName

The name of an algorithm that was used to create the model package. The algorithm must be either an algorithm resource in your SageMaker account or an algorithm in AWS Marketplace that you are subscribed to.

Type: String


Pattern: \(\text{arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]\{12\}:[a-z\-]*\/?}\{a-zA-Z0-9\{0,62\}\{0,62\}\{?<!-)\}\$\)

Required: Yes

ModelDataUrl

The Amazon S3 path where the model artifacts, which result from model training, are stored. This path must point to a single gzip compressed tar archive (\(.tar.gz\) suffix).

Note

The model artifacts must be in an S3 bucket that is in the same region as the algorithm.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^\(https|s3\)://(\[^/\]+)(((^\//)?([^\./])\$\)

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
SourceAlgorithmSpecification
Service: Amazon SageMaker Service
A list of algorithms that were used to create a model package.

Contents

SourceAlgorithms
A list of the algorithms that were used to create a model package.
Type: Array of SourceAlgorithm (p. 1367) objects
Array Members: Fixed number of 1 item.
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
SourceIpConfig

Service: Amazon SageMaker Service

A list of IP address ranges (CIDRs). Used to create an allow list of IP addresses for a private workforce. Workers will only be able to login to their worker portal from an IP address within this range. By default, a workforce isn't restricted to specific IP addresses.

Contents

Cidrs

A list of one to ten Classless Inter-Domain Routing (CIDR) values.

Maximum: Ten CIDR values

Note
The following Length Constraints apply to individual CIDR values in the CIDR value list.

Type: Array of strings

Length Constraints: Minimum length of 4. Maximum length of 64.

Pattern:

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
StoppingCondition

Service: Amazon SageMaker Service

Specifies a limit to how long a model training job or model compilation job can run. It also specifies how long a managed spot training job has to complete. When the job reaches the time limit, SageMaker ends the training or compilation job. Use this API to cap model training costs.

To stop a training job, SageMaker sends the algorithm the `SIGTERM` signal, which delays job termination for 120 seconds. Algorithms can use this 120-second window to save the model artifacts, so the results of training are not lost.

The training algorithms provided by SageMaker automatically save the intermediate results of a model training job when possible. This attempt to save artifacts is only a best effort case as model might not be in a state from which it can be saved. For example, if training has just started, the model might not be ready to save. When saved, this intermediate data is a valid model artifact. You can use it to create a model with `CreateModel`.

**Note**
The Neural Topic Model (NTM) currently does not support saving intermediate model artifacts. When training NTMs, make sure that the maximum runtime is sufficient for the training job to complete.

Contents

**MaxRuntimeInSeconds**

The maximum length of time, in seconds, that a training or compilation job can run.

For compilation jobs, if the job does not complete during this time, a `Timeout` error is generated. We recommend starting with 900 seconds and increasing as necessary based on your model.

For all other jobs, if the job does not complete during this time, SageMaker ends the job. When `RetryStrategy` is specified in the job request, `MaxRuntimeInSeconds` specifies the maximum time for all of the attempts in total, not each individual attempt. The default value is 1 day. The maximum value is 28 days.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

**MaxWaitTimeInSeconds**

The maximum length of time, in seconds, that a managed Spot training job has to complete. It is the amount of time spent waiting for Spot capacity plus the amount of time the job can run. It must be equal to or greater than `MaxRuntimeInSeconds`. If the job does not complete during this time, SageMaker ends the job.

When `RetryStrategy` is specified in the job request, `MaxWaitTimeInSeconds` specifies the maximum time for all of the attempts in total, not each individual attempt.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
StudioLifecycleConfigDetails

Service: Amazon SageMaker Service

Details of the Studio Lifecycle Configuration.

Contents

CreationTime

The creation time of the Studio Lifecycle Configuration.
Type: Timestamp
Required: No

LastModifiedTime

This value is equivalent to CreationTime because Studio Lifecycle Configurations are immutable.
Type: Timestamp
Required: No

StudioLifecycleConfigAppType

The App type to which the Lifecycle Configuration is attached.
Type: String
Valid Values: JupyterServer | KernelGateway
Required: No

StudioLifecycleConfigArn

The Amazon Resource Name (ARN) of the Lifecycle Configuration.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:studio-lifecycle-config/.*
Required: No

StudioLifecycleConfigName

The name of the Studio Lifecycle Configuration.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9\-]*\-[^a-zA-Z0-9\-]{0,62}$
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
SubscribedWorkteam
Service: Amazon SageMaker Service

Describes a work team of a vendor that does the a labelling job.

Contents

ListingId

Marketplace product listing ID.
Type: String
Required: No

MarketplaceDescription

The description of the vendor from the Amazon Marketplace.
Type: String
Pattern: .*
Required: No

MarketplaceTitle

The title of the service provided by the vendor in the Amazon Marketplace.
Type: String
Pattern: .*
Required: No

SellerName

The name of the vendor in the Amazon Marketplace.
Type: String
Required: No

WorkteamArn

The Amazon Resource Name (ARN) of the vendor that you have subscribed.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[\a-z\-]*:sagemaker:[\a-z0-9\-]*:[0-9]{12}:workteam/.*
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
SuggestionQuery
Service: Amazon SageMaker Service

Specified in the GetSearchSuggestions (p. 492) request. Limits the property names that are included in the response.

Contents

PropertyNameQuery

Defines a property name hint. Only property names that begin with the specified hint are included in the response.

Type: PropertyNameQuery (p. 1314) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Tag
Service: Amazon SageMaker Service

A tag object that consists of a key and an optional value, used to manage metadata for SageMaker AWS resources.

You can add tags to notebook instances, training jobs, hyperparameter tuning jobs, batch transform jobs, models, labeling jobs, work teams, endpoint configurations, and endpoints. For more information on adding tags to SageMaker resources, see AddTags (p. 17).

For more information on adding metadata to your AWS resources with tagging, see Tagging AWS resources. For advice on best practices for managing AWS resources with tagging, see Tagging Best Practices: Implement an Effective AWS Resource Tagging Strategy.

Contents

Key

The tag key. Tag keys must be unique per resource.

Type: String


Pattern: `^\[^{\w}\p{Z}\p{N}\_.:/=+\-@}*\]$`

Required: Yes

Value

The tag value.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `^\[^{\w}\p{Z}\p{N}\_.:/=+/\-@}*\]$`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TargetPlatform
Service: Amazon SageMaker Service

Contains information about a target platform that you want your model to run on, such as OS, architecture, and accelerators. It is an alternative of TargetDevice.

Contents

Accelerator

Specifies a target platform accelerator (optional).

- **NVIDIA**: Nvidia graphics processing unit. It also requires `gpu-code`, `trt-ver`, `cuda-ver` compiler options
- **MALI**: ARM Mali graphics processor
- **INTEL_GRAPHICS**: Integrated Intel graphics

Type: String

Valid Values: INTEL_GRAPHICS | MALI | NVIDIA | NNA

Required: No

Arch

Specifies a target platform architecture.

- **X86_64**: 64-bit version of the x86 instruction set.
- **X86**: 32-bit version of the x86 instruction set.
- **ARM64**: ARMv8 64-bit CPU.
- **ARM_EABIHF**: ARMv7 32-bit, Hard Float.

Type: String

Valid Values: X86_64 | X86 | ARM64 | ARM_EABI | ARM_EABIHF

Required: Yes

Os

Specifies a target platform OS.

- **LINUX**: Linux-based operating systems.
- **ANDROID**: Android operating systems. Android API level can be specified using the `ANDROID_PLATFORM` compiler option. For example, "CompilerOptions": {'ANDROID_PLATFORM': 28}

Type: String

Valid Values: ANDROID | LINUX

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
TensorBoardAppSettings
Service: Amazon SageMaker Service

The TensorBoard app settings.

Contents

DefaultResourceSpec

The default instance type and the Amazon Resource Name (ARN) of the SageMaker image created on the instance.

Type: ResourceSpec (p. 1341) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TensorBoardOutputConfig
Service: Amazon SageMaker Service

Configuration of storage locations for the Debugger TensorBoard output data.

Contents

**LocalPath**

Path to local storage location for tensorBoard output. Defaults to /opt/ml/output/tensorboard.

Type: String

Length Constraints: Maximum length of 4096.

Pattern: .*

Required: No

**S3OutputPath**

Path to Amazon S3 storage location for TensorBoard output.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)/(.*)$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrafficPattern
Service: Amazon SageMaker Service

Defines the traffic pattern of the load test.

Contents

Phases

Defines the phases traffic specification.

Type: Array of Phase (p. 1257) objects

Array Members: Minimum number of 1 item.

Required: No

TrafficType

Defines the traffic patterns.

Type: String

Valid Values: PHASES

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrafficRoutingConfig

Service: Amazon SageMaker Service

Defines the traffic routing strategy during an endpoint deployment to shift traffic from the old fleet to the new fleet.

Contents

CanarySize

Batch size for the first step to turn on traffic on the new endpoint fleet. Value must be less than or equal to 50% of the variant's total instance count.

Type: CapacitySize (p. 935) object

Required: No

LinearStepSize

Batch size for each step to turn on traffic on the new endpoint fleet. Value must be 10-50% of the variant's total instance count.

Type: CapacitySize (p. 935) object

Required: No

Type

Traffic routing strategy type.

- ALL_AT_ONCE: Endpoint traffic shifts to the new fleet in a single step.
- CANARY: Endpoint traffic shifts to the new fleet in two steps. The first step is the canary, which is a small portion of the traffic. The second step is the remainder of the traffic.
- LINEAR: Endpoint traffic shifts to the new fleet in n steps of a configurable size.

Type: String

Valid Values: ALL_AT_ONCE | CANARY | LINEAR

Required: Yes

WaitIntervalInSeconds

The waiting time (in seconds) between incremental steps to turn on traffic on the new endpoint fleet.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 3600.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
• AWS SDK for Ruby V3
TrainingJob
Service: Amazon SageMaker Service
Contains information about a training job.

Contents

AlgorithmSpecification
Information about the algorithm used for training, and algorithm metadata.
Type: AlgorithmSpecification (p. 867) object
Required: No

AutoMLJobArn
The Amazon Resource Name (ARN) of the job.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*
Required: No

BillableTimeInSeconds
The billable time in seconds.
Type: Integer
Valid Range: Minimum value of 1.
Required: No

CheckpointConfig
Contains information about the output location for managed spot training checkpoint data.
Type: CheckpointConfig (p. 945) object
Required: No

CreationTime
A timestamp that indicates when the training job was created.
Type: Timestamp
Required: No

DebugHookConfig
Configuration information for the Debugger hook parameters, metric and tensor collections, and storage paths. To learn more about how to configure the DebugHookConfig parameter, see Use the SageMaker and Debugger Configuration API Operations to Create, Update, and Debug Your Training Job.
Type: DebugHookConfig (p. 980) object
Required: No
**DebugRuleConfigurations**

Information about the debug rule configuration.

Type: Array of DebugRuleConfiguration (p. 982) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

Required: No

**DebugRuleEvaluationStatuses**

Information about the evaluation status of the rules for the training job.

Type: Array of DebugRuleEvaluationStatus (p. 984) objects

Array Members: Minimum number of 0 items. Maximum number of 20 items.

Required: No

**EnableInterContainerTrafficEncryption**

To encrypt all communications between ML compute instances in distributed training, choose True. Encryption provides greater security for distributed training, but training might take longer. How long it takes depends on the amount of communication between compute instances, especially if you use a deep learning algorithm in distributed training.

Type: Boolean

Required: No

**EnableManagedSpotTraining**

When true, enables managed spot training using Amazon EC2 Spot instances to run training jobs instead of on-demand instances. For more information, see Managed Spot Training.

Type: Boolean

Required: No

**EnableNetworkIsolation**

If the **TrainingJob** was created with network isolation, the value is set to true. If network isolation is enabled, nodes can't communicate beyond the VPC they run in.

Type: Boolean

Required: No

**Environment**

The environment variables to set in the Docker container.

Type: String to string map

Map Entries: Maximum number of 48 items.

Key Length Constraints: Maximum length of 512.

Key Pattern: \[a-zA-Z_]\[a-zA-Z0-9_]\*

Value Length Constraints: Maximum length of 512.

Value Pattern: \[\S\s\]*

Required: No
**ExperimentConfig**

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:

- `CreateProcessingJob (p. 162)`
- `CreateTrainingJob (p. 173)`
- `CreateTransformJob (p. 182)`

Type: `ExperimentConfig (p. 1029)` object

Required: No

**FailureReason**

If the training job failed, the reason it failed.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

**FinalMetricDataList**

A list of final metric values that are set when the training job completes. Used only if the training job was configured to use metrics.

Type: Array of `MetricData (p. 1145)` objects

Array Members: Minimum number of 0 items. Maximum number of 40 items.

Required: No

**HyperParameters**

Algorithm-specific parameters.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 2500.

Value Pattern: .*

Required: No

**InputDataConfig**

An array of `Channel` objects that describes each data input channel.

Type: Array of `Channel (p. 941)` objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Required: No

**LabelingJobArn**

The Amazon Resource Name (ARN) of the labeling job.
TrainingJob

Type: String
Length Constraints: Maximum length of 2048.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:labeling-job/.*
Required: No

LastModifiedTime
A timestamp that indicates when the status of the training job was last modified.
Type: Timestamp
Required: No

ModelArtifacts
Information about the Amazon S3 location that is configured for storing model artifacts.
Type: ModelArtifacts (p. 1149) object
Required: No

OutputDataConfig
The S3 path where model artifacts that you configured when creating the job are stored. SageMaker creates subfolders for model artifacts.
Type: OutputDataConfig (p. 1244) object
Required: No

ResourceConfig
Resources, including ML compute instances and ML storage volumes, that are configured for model training.
Type: ResourceConfig (p. 1338) object
Required: No

RetryStrategy
The number of times to retry the job when the job fails due to an InternalServerError.
Type: RetryStrategy (p. 1344) object
Required: No

RoleArn
The AWS Identity and Access Management (IAM) role configured for the training job.
Type: String
Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z0-9+,.@\-_\/]+$
Required: No

SecondaryStatus
Provides detailed information about the state of the training job. For detailed information about the secondary status of the training job, see StatusMessage under SecondaryStatusTransition (p. 1359).
SageMaker provides primary statuses and secondary statuses that apply to each of them:

### InProgress
- **Starting** - Starting the training job.
- **Downloading** - An optional stage for algorithms that support *File* training input mode. It indicates that data is being downloaded to the ML storage volumes.
- **Training** - Training is in progress.
- **Uploading** - Training is complete and the model artifacts are being uploaded to the S3 location.

### Completed
- **Completed** - The training job has completed.

### Failed
- **Failed** - The training job has failed. The reason for the failure is returned in the *FailureReason* field of *DescribeTrainingJobResponse*.

### Stopped
- **MaxRuntimeExceeded** - The job stopped because it exceeded the maximum allowed runtime.
- **Stopped** - The training job has stopped.

### Stopping
- **Stopping** - Stopping the training job.

**Important**
Valid values for *SecondaryStatus* are subject to change.

We no longer support the following secondary statuses:
- **LaunchingMLInstances**
- **PreparingTrainingStack**
- **DownloadingTrainingImage**

**Type:** String

**Valid Values:**
- `Starting`
- `LaunchingMLInstances`
- `PreparingTrainingStack`
- `Downloading`
- `DownloadingTrainingImage`
- `Training`
- `Uploading`
- `Stopping`
- `Stopped`
- `MaxRuntimeExceeded`
- `Completed`
- `Failed`
- `Interrupted`
- `MaxWaitTimeExceeded`
- `Updating`
- `Restarting`

**Required:** No

### SecondaryStatusTransitions
A history of all of the secondary statuses that the training job has transitioned through.

**Type:** Array of *SecondaryStatusTransition* (p. 1359) objects

**Required:** No

### StoppingCondition
Specifies a limit to how long a model training job can run. It also specifies how long a managed Spot training job has to complete. When the job reaches the time limit, SageMaker ends the training job.

Use this API to cap model training costs.

To stop a job, SageMaker sends the algorithm the **SIGTERM** signal, which delays job termination for 120 seconds. Algorithms can use this 120-second window to save the model artifacts, so the results of training are not lost.

**Type:** *StoppingCondition* (p. 1370) object
Tags

An array of key-value pairs. You can use tags to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. For more information, see Tagging AWS Resources.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

TensorBoardOutputConfig

Configuration of storage locations for the Debugger TensorBoard output data.

Type: TensorBoardOutputConfig (p. 1381) object

Required: No

TrainingEndTime

Indicates the time when the training job ends on training instances. You are billed for the time interval between the value of TrainingStartTime and this time. For successful jobs and stopped jobs, this is the time after model artifacts are uploaded. For failed jobs, this is the time when SageMaker detects a job failure.

Type: Timestamp

Required: No

TrainingJobArn

The Amazon Resource Name (ARN) of the training job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:training-job/.*

Required: No

TrainingJobName

The name of the training job.

Type: String


Pattern: ^[a-zA-Z0-9-](\*[a-zA-Z0-9])\{0,62}$

Required: No

TrainingJobStatus

The status of the training job.

Training job statuses are:

- InProgress - The training is in progress.
- Completed - The training job has completed.
- Failed - The training job has failed. To see the reason for the failure, see the FailureReason field in the response to a DescribeTrainingJobResponse call.
• Stopping - The training job is stopping.
  • Stopped - The training job has stopped.

For more detailed information, see SecondaryStatus.

Type: String

Valid Values: InProgress | Completed | Failed | Stopping | Stopped

Required: No

TrainingStartTime

Indicates the time when the training job starts on training instances. You are billed for the time interval between this time and the value of TrainingEndTime. The start time in CloudWatch Logs might be later than this time. The difference is due to the time it takes to download the training data and to the size of the training container.

Type: Timestamp

Required: No

TrainingTimeInSeconds

The training time in seconds.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

TuningJobArn

The Amazon Resource Name (ARN) of the associated hyperparameter tuning job if the training job was launched by a hyperparameter tuning job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:hyper-parameter-tuning-job/.*

Required: No

VpcConfig

A VpcConfig (p. 1456) object that specifies the VPC that this training job has access to. For more information, see Protect Training Jobs by Using an Amazon Virtual Private Cloud.

Type: VpcConfig (p. 1456) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
**TrainingJobDefinition**

Service: Amazon SageMaker Service

Defines the input needed to run a training job using the algorithm.

**Contents**

**HyperParameters**

The hyperparameters used for the training job.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: .*

Value Length Constraints: Maximum length of 2500.

Value Pattern: .*

Required: No

**InputDataConfig**

An array of Channel objects, each of which specifies an input source.

Type: Array of Channel (p. 941) objects

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Required: Yes

**OutputDataConfig**

the path to the S3 bucket where you want to store model artifacts. SageMaker creates subfolders for the artifacts.

Type: OutputDataConfig (p. 1244) object

Required: Yes

**ResourceConfig**

The resources, including the ML compute instances and ML storage volumes, to use for model training.

Type: ResourceConfig (p. 1338) object

Required: Yes

**StoppingCondition**

Specifies a limit to how long a model training job can run. It also specifies how long a managed Spot training job has to complete. When the job reaches the time limit, SageMaker ends the training job. Use this API to cap model training costs.

To stop a job, SageMaker sends the algorithm the SIGTERM signal, which delays job termination for 120 seconds. Algorithms can use this 120-second window to save the model artifacts.

Type: StoppingCondition (p. 1370) object
TrainingInputMode

The training input mode that the algorithm supports. For more information about input modes, see Algorithms.

Pipe mode

If an algorithm supports Pipe mode, Amazon SageMaker streams data directly from Amazon S3 to the container.

File mode

If an algorithm supports File mode, SageMaker downloads the training data from S3 to the provisioned ML storage volume, and mounts the directory to the Docker volume for the training container.

You must provision the ML storage volume with sufficient capacity to accommodate the data downloaded from S3. In addition to the training data, the ML storage volume also stores the output model. The algorithm container uses the ML storage volume to also store intermediate information, if any.

For distributed algorithms, training data is distributed uniformly. Your training duration is predictable if the input data objects sizes are approximately the same. SageMaker does not split the files any further for model training. If the object sizes are skewed, training won't be optimal as the data distribution is also skewed when one host in a training cluster is overloaded, thus becoming a bottleneck in training.

FastFile mode

If an algorithm supports FastFile mode, SageMaker streams data directly from S3 to the container with no code changes, and provides file system access to the data. Users can author their training script to interact with these files as if they were stored on disk.

FastFile mode works best when the data is read sequentially. Augmented manifest files aren't supported. The startup time is lower when there are fewer files in the S3 bucket provided.

Type: String

Valid Values: Pipe | File | FastFile

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**TrainingJobStatusCounters**

*Service: Amazon SageMaker Service*

The numbers of training jobs launched by a hyperparameter tuning job, categorized by status.

**Contents**

**Completed**

The number of completed training jobs launched by the hyperparameter tuning job.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**InProgress**

The number of in-progress training jobs launched by a hyperparameter tuning job.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**NonRetryableError**

The number of training jobs that failed and can't be retried. A failed training job can't be retried if it failed because a client error occurred.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**RetryableError**

The number of training jobs that failed, but can be retried. A failed training job can be retried only if it failed because an internal service error occurred.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**Stopped**

The number of training jobs launched by a hyperparameter tuning job that were manually stopped.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrainingJobStepMetadata
Service: Amazon SageMaker Service

Metadata for a training job step.

Contents

Arn

The Amazon Resource Name (ARN) of the training job that was run by this step execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:training-job/.*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrainingJobSummary
Service: Amazon SageMaker Service
Provides summary information about a training job.

Contents

CreationTime
A timestamp that shows when the training job was created.
Type: Timestamp
Required: Yes

LastModifiedTime
Timestamp when the training job was last modified.
Type: Timestamp
Required: No

TrainingEndTime
A timestamp that shows when the training job ended. This field is set only if the training job has one of the terminal statuses (Completed, Failed, or Stopped).
Type: Timestamp
Required: No

TrainingJobArn
The Amazon Resource Name (ARN) of the training job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-zA-Z\-]*sagemaker:[a-zA-Z0-9\-]*:[0-9]{12}:training-job/.*
Required: Yes

TrainingJobName
The name of the training job that you want a summary for.
Type: String
Pattern: ^[a-zA-Z0-9\-]*(-*[a-zA-Z0-9])\{0,62\}
Required: Yes

TrainingJobStatus
The status of the training job.
Type: String
Valid Values: InProgress | Completed | Failed | Stopping | Stopped
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrainingSpecification
Service: Amazon SageMaker Service

 Defines how the algorithm is used for a training job.

Contents

MetricDefinitions

A list of MetricDefinition objects, which are used for parsing metrics generated by the algorithm.

Type: Array of MetricDefinition (p. 1147) objects

Array Members: Minimum number of 0 items. Maximum number of 40 items.

Required: No

SupportedHyperParameters

A list of the HyperParameterSpecification objects, that define the supported hyperparameters. This is required if the algorithm supports automatic model tuning.

Type: Array of HyperParameterSpecification (p. 1080) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: No

SupportedTrainingInstanceTypes

A list of the instance types that this algorithm can use for training.

Type: Array of strings

Valid Values: ml.m4.xlarge | ml.m4.2xlarge | ml.m4.4xlarge | ml.m4.10xlarge
| ml.m4.16xlarge | ml.g4dn.xlarge | ml.g4dn.2xlarge | ml.g4dn.4xlarge
| ml.g4dn.8xlarge | ml.g4dn.12xlarge | ml.g4dn.16xlarge | ml.m5.large
| ml.m5.xlarge | ml.m5.2xlarge | ml.m5.4xlarge | ml.m5.12xlarge
| ml.m5.24xlarge | ml.c4.xlarge | ml.c4.2xlarge | ml.c4.4xlarge
| ml.c4.8xlarge | ml.p2.xlarge | ml.p2.8xlarge | ml.p2.16xlarge
| ml.p3.2xlarge | ml.p3.8xlarge | ml.p3.16xlarge | ml.p3dn.24xlarge
| ml.p4d.24xlarge | ml.c5.xlarge | ml.c5.2xlarge | ml.c5.4xlarge
| ml.c5.9xlarge | ml.c5.18xlarge | ml.c5n.xlarge | ml.c5n.2xlarge
| ml.c5n.4xlarge | ml.c5n.9xlarge | ml.c5n.18xlarge | ml.g5.xlarge
| ml.g5.2xlarge | ml.g5.4xlarge | ml.g5.8xlarge | ml.g5.16xlarge
| ml.g5.12xlarge | ml.g5.24xlarge | ml.g5.48xlarge

Required: Yes

SupportedTuningJobObjectiveMetrics

A list of the metrics that the algorithm emits that can be used as the objective metric in a hyperparameter tuning job.

Type: Array of HyperParameterTuningJobObjective (p. 1091) objects

Required: No

SupportsDistributedTraining

Indicates whether the algorithm supports distributed training. If set to false, buyers can't request more than one instance during training.
TrainingSpecification

Type: Boolean
Required: No

TrainingChannels

A list of ChannelSpecification objects, which specify the input sources to be used by the algorithm.

Type: Array of ChannelSpecification (p. 943) objects
Array Members: Minimum number of 1 item. Maximum number of 8 items.
Required: Yes

TrainingImage

The Amazon ECR registry path of the Docker image that contains the training algorithm.

Type: String
Length Constraints: Maximum length of 255.
Pattern: [\S]+
Required: Yes

TrainingImageDigest

An MD5 hash of the training algorithm that identifies the Docker image used for training.

Type: String
Length Constraints: Maximum length of 72.
Pattern: ^[Ss][Hh][Aa]256:[0-9a-fA-F]{64}$
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformDataSource
Service: Amazon SageMaker Service

Describes the location of the channel data.

Contents

S3DataSource

The S3 location of the data source that is associated with a channel.

Type: TransformS3DataSource (p. 1419) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformInput
Service: Amazon SageMaker Service

Describes the input source of a transform job and the way the transform job consumes it.

Contents

CompressionType

If your transform data is compressed, specify the compression type. Amazon SageMaker automatically decompresses the data for the transform job accordingly. The default value is None.

Type: String

Valid Values: None | Gzip

Required: No

ContentType

The multipurpose internet mail extension (MIME) type of the data. Amazon SageMaker uses the MIME type with each http call to transfer data to the transform job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

DataSource

Describes the location of the channel data, which is, the S3 location of the input data that the model can consume.

Type: TransformDataSource (p. 1402) object

Required: Yes

SplitType

The method to use to split the transform job's data files into smaller batches. Splitting is necessary when the total size of each object is too large to fit in a single request. You can also use data splitting to improve performance by processing multiple concurrent mini-batches. The default value for SplitType is None, which indicates that input data files are not split, and request payloads contain the entire contents of an input object. Set the value of this parameter to Line to split records on a newline character boundary. SplitType also supports a number of record-oriented binary data formats. Currently, the supported record formats are:

- RecordIO
- TFRecord

When splitting is enabled, the size of a mini-batch depends on the values of the BatchStrategy and MaxPayloadInMB parameters. When the value of BatchStrategy is MultiRecord, Amazon SageMaker sends the maximum number of records in each request, up to the MaxPayloadInMB limit. If the value of BatchStrategy is SingleRecord, Amazon SageMaker sends individual records in each request.

Note

Some data formats represent a record as a binary payload wrapped with extra padding bytes. When splitting is applied to a binary data format, padding is removed if the value
of BatchStrategy is set to SingleRecord. Padding is not removed if the value of BatchStrategy is set to MultiRecord.

For more information about RecordIO, see Create a Dataset Using RecordIO in the MXNet documentation. For more information about TFRecord, see Consuming TFRecord data in the TensorFlow documentation.

Type: String

Valid Values: None | Line | RecordIO | TFRecord

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformJob

Service: Amazon SageMaker Service

A batch transform job. For information about SageMaker batch transform, see Use Batch Transform.

Contents

AutoMLJobArn

The Amazon Resource Name (ARN) of the AutoML job that created the transform job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:automl-job/.*

Required: No

BatchStrategy

Specifies the number of records to include in a mini-batch for an HTTP inference request. A record is a single unit of input data that inference can be made on. For example, a single line in a CSV file is a record.

Type: String

Valid Values: MultiRecord | SingleRecord

Required: No

CreationTime

A timestamp that shows when the transform Job was created.

Type: Timestamp

Required: No

DataProcessing

The data structure used to specify the data to be used for inference in a batch transform job and to associate the data that is relevant to the prediction results in the output. The input filter provided allows you to exclude input data that is not needed for inference in a batch transform job. The output filter provided allows you to include input data relevant to interpreting the predictions in the output from the job. For more information, see Associate Prediction Results with their Corresponding Input Records.

Type: DataProcessing (p. 971) object

Required: No

Environment

The environment variables to set in the Docker container. We support up to 16 key and values entries in the map.

Type: String to string map

Map Entries: Maximum number of 16 items.

Key Length Constraints: Maximum length of 1024.
Key Pattern: [a-zA-Z\-_]\[a-zA-Z0-9\-_]{0,1023}

Value Length Constraints: Maximum length of 10240.

Value Pattern: [\S\s]*

Required: No

**ExperimentConfig**

Associates a SageMaker job as a trial component with an experiment and trial. Specified when you call the following APIs:
- CreateProcessingJob (p. 162)
- CreateTrainingJob (p. 173)
- CreateTransformJob (p. 182)

Type: ExperimentConfig (p. 1029) object

Required: No

**FailureReason**

If the transform job failed, the reason it failed.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

**LabelingJobArn**

The Amazon Resource Name (ARN) of the labeling job that created the transform job.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:labeling-job/.*

Required: No

**MaxConcurrentTransforms**

The maximum number of parallel requests that can be sent to each instance in a transform job. If MaxConcurrentTransforms is set to 0 or left unset, SageMaker checks the optional execution-parameters to determine the settings for your chosen algorithm. If the execution-parameters endpoint is not enabled, the default value is 1. For built-in algorithms, you don't need to set a value for MaxConcurrentTransforms.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**MaxPayloadInMB**

The maximum allowed size of the payload, in MB. A payload is the data portion of a record (without metadata). The value in MaxPayloadInMB must be greater than, or equal to, the size of a single record. To estimate the size of a record in MB, divide the size of your dataset by the number of records. To ensure that the records fit within the maximum payload size, we recommend using a slightly larger value. The default value is 6 MB. For cases where the payload might be arbitrarily
large and is transmitted using HTTP chunked encoding, set the value to 0. This feature works only in supported algorithms. Currently, SageMaker built-in algorithms do not support HTTP chunked encoding.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

**ModelClientConfig**

Configures the timeout and maximum number of retries for processing a transform job invocation.

Type: ModelClientConfig (p. 1153) object

Required: No

**ModelName**

The name of the model associated with the transform job.

Type: String

Length Constraints: Maximum length of 63.

Pattern: ^[a-zA-Z0-9\-\[\*\-]*[a-zA-Z0-9\-\] )* $

Required: No

**Tags**

A list of tags associated with the transform job.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**TransformEndTime**

Indicates when the transform job has been completed, or has stopped or failed. You are billed for the time interval between this time and the value of **TransformStartTime**.

Type: Timestamp

Required: No

**TransformInput**

Describes the input source of a transform job and the way the transform job consumes it.

Type: TransformInput (p. 1403) object

Required: No

**TransformJobArn**

The Amazon Resource Name (ARN) of the transform job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-\[\*\-]*:sagemaker\[a-z0-9\[\-\[\*\[0-9\[\{12\]:transform-job/.*
Required: No

**TransformJobName**

The name of the transform job.

Type: String


Pattern: ^[a-zA-Z0-9][-\[a-zA-Z0-9]]\{0,62}\}

Required: No

**TransformJobStatus**

The status of the transform job.

Transform job statuses are:
- **InProgress** - The job is in progress.
- **Completed** - The job has completed.
- **Failed** - The transform job has failed. To see the reason for the failure, see the FailureReason field in the response to a DescribeTransformJob call.
- **Stopping** - The transform job is stopping.
- **Stopped** - The transform job has stopped.

Type: String

Valid Values: InProgress | Completed | Failed | Stopping | Stopped

Required: No

**TransformOutput**

Describes the results of a transform job.

Type: TransformOutput (p. 1415) object

Required: No

**TransformResources**

Describes the resources, including ML instance types and ML instance count, to use for transform job.

Type: TransformResources (p. 1417) object

Required: No

**TransformStartTime**

Indicates when the transform job starts on ML instances. You are billed for the time interval between this time and the value of TransformEndTime.

Type: Timestamp

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformJobDefinition
Service: Amazon SageMaker Service

Defines the input needed to run a transform job using the inference specification specified in the algorithm.

**Contents**

**BatchStrategy**

A string that determines the number of records included in a single mini-batch.

- **SingleRecord** means only one record is used per mini-batch. **MultiRecord** means a mini-batch is set to contain as many records that can fit within the MaxPayloadInMB limit.

  Type: String

  Valid Values: MultiRecord | SingleRecord

  Required: No

**Environment**

The environment variables to set in the Docker container. We support up to 16 key and values entries in the map.

  Type: String to string map

  Map Entries: Maximum number of 16 items.

  Key Length Constraints: Maximum length of 1024.

  Key Pattern: [a-zA-Z_][a-zA-Z0-9_]{0,1023}

  Value Length Constraints: Maximum length of 10240.

  Value Pattern: \[\S\s\]*

  Required: No

**MaxConcurrentTransforms**

The maximum number of parallel requests that can be sent to each instance in a transform job. The default value is 1.

  Type: Integer

  Valid Range: Minimum value of 0.

  Required: No

**MaxPayloadInMB**

The maximum payload size allowed, in MB. A payload is the data portion of a record (without metadata).

  Type: Integer

  Valid Range: Minimum value of 0.

  Required: No
TransformInput

A description of the input source and the way the transform job consumes it.

Type: TransformInput (p. 1403) object

Required: Yes

TransformOutput

Identifies the Amazon S3 location where you want Amazon SageMaker to save the results from the transform job.

Type: TransformOutput (p. 1415) object

Required: Yes

TransformResources

Identifies the ML compute instances for the transform job.

Type: TransformResources (p. 1417) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformJobStepMetadata
Service: Amazon SageMaker Service

Metadata for a transform job step.

Contents

Arn

The Amazon Resource Name (ARN) of the transform job that was run by this step execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:transform-job/.*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformJobSummary
Service: Amazon SageMaker Service

Provides a summary of a transform job. Multiple TransformJobSummary objects are returned as a list after in response to a ListTransformJobs (p. 658) call.

Contents
CreationTime
A timestamp that shows when the transform Job was created.
Type: Timestamp
Required: Yes

FailureReason
If the transform job failed, the reason it failed.
Type: String
Length Constraints: Maximum length of 1024.
Required: No

LastModifiedTime
Indicates when the transform job was last modified.
Type: Timestamp
Required: No

TransformEndTime
Indicates when the transform job ends on compute instances. For successful jobs and stopped jobs, this is the exact time recorded after the results are uploaded. For failed jobs, this is when Amazon SageMaker detected that the job failed.
Type: Timestamp
Required: No

TransformJobArn
The Amazon Resource Name (ARN) of the transform job.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:transform-job/.*
Required: Yes

TransformJobName
The name of the transform job.
Type: String
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}$
Required: Yes

TransformJobStatus

The status of the transform job.

Type: String

Valid Values: InProgress | Completed | Failed | Stopping | Stopped

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformOutput

Service: Amazon SageMaker Service

Describes the results of a transform job.

Contents

Accept

The MIME type used to specify the output data. Amazon SageMaker uses the MIME type with each http call to transfer data from the transform job.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

AssembleWith

 Defines how to assemble the results of the transform job as a single S3 object. Choose a format that is most convenient to you. To concatenate the results in binary format, specify None. To add a newline character at the end of every transformed record, specify Line.

Type: String

Valid Values: None | Line

Required: No

KmsKeyId

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt the model artifacts at rest using Amazon S3 server-side encryption. The KmsKeyId can be any of the following formats:

- Key ID: 1234abcd-12ab-34cd-56ef-1234567890ab
- Key ARN: arn:aws:kms:us-west-2:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab
- Alias name: alias/ExampleAlias

If you don't provide a KMS key ID, Amazon SageMaker uses the default KMS key for Amazon S3 for your role's account. For more information, see KMS-Managed Encryption Keys in the Amazon Simple Storage Service Developer Guide.

The KMS key policy must grant permission to the IAM role that you specify in your CreateModel request. For more information, see Using Key Policies in AWS KMS in the AWS Key Management Service Developer Guide.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: No
S3OutputPath

The Amazon S3 path where you want Amazon SageMaker to store the results of the transform job. For example, s3://bucket-name/key-name-prefix.

For every S3 object used as input for the transform job, batch transform stores the transformed data with an .out suffix in a corresponding subfolder in the location in the output prefix. For example, for the input data stored at s3://bucket-name/input-name-prefix/dataset01/data.csv, batch transform stores the transformed data at s3://bucket-name/output-name-prefix/input-name-prefix/data.csv.out. Batch transform doesn't upload partially processed objects. For an input S3 object that contains multiple records, it creates an .out file only if the transform job succeeds on the entire file. When the input contains multiple S3 objects, the batch transform job processes the listed S3 objects and uploads only the output for successfully processed objects. If any object fails in the transform job batch transform marks the job as failed to prompt investigation.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: ^(https|s3)://([^/]+)?(.*$)

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformResources

Service: Amazon SageMaker Service

Describes the resources, including ML instance types and ML instance count, to use for transform job.

Contents

InstanceCount

The number of ML compute instances to use in the transform job. For distributed transform jobs, specify a value greater than 1. The default value is 1.

Type: Integer

Valid Range: Minimum value of 1.

Required: Yes

InstanceType

The ML compute instance type for the transform job. If you are using built-in algorithms to transform moderately sized datasets, we recommend using ml.m4.xlarge or ml.m5.large instance types.

Type: String

Valid Values:
- ml.m4.xlarge
- ml.m4.2xlarge
- ml.m4.4xlarge
- ml.m4.10xlarge
- ml.m4.16xlarge
- ml.m4.2xlarge
- ml.m4.4xlarge
- ml.m4.8xlarge
- ml.m4.16xlarge
- ml.m4.24xlarge
- ml.m4.large
- ml.m5.xlarge
- ml.m5.2xlarge
- ml.m5.4xlarge
- ml.m5.8xlarge
- ml.m5.16xlarge
- ml.m5.24xlarge
- ml.m5.large
- ml.m5.2xlarge
- ml.m5.4xlarge
- ml.m5.8xlarge
- ml.m5.16xlarge
- ml.m5.24xlarge
- ml.m5.large

Required: Yes

VolumeKmsKeyId

The AWS Key Management Service (AWS KMS) key that Amazon SageMaker uses to encrypt model data on the storage volume attached to the ML compute instance(s) that run the batch transform job.

Note

Certain Nitro-based instances include local storage, dependent on the instance type. Local storage volumes are encrypted using a hardware module on the instance. You can't request a VolumeKmsKeyId when using an instance type with local storage.

For a list of instance types that support local instance storage, see Instance Store Volumes.

For more information about local instance storage encryption, see SSD Instance Store Volumes.

The VolumeKmsKeyId can be any of the following formats:
- Key ID: 1234abcd-12ab-34cd-56ef-1234567890ab
- Key ARN: arn:aws:kms:us-west-2:1234abcd-12ab-34cd-56ef-1234567890ab
- Alias name: alias/ExampleAlias
- Alias name ARN: arn:aws:kms:us-west-2:1234abcd-12ab-34cd-56ef-1234567890ab

Type: String
Length Constraints: Maximum length of 2048.

Pattern: . *

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TransformS3DataSource
Service: Amazon SageMaker Service

Describes the S3 data source.

Contents

S3DataType

If you choose **S3Prefix**, **S3Uri** identifies a key name prefix. Amazon SageMaker uses all objects with the specified key name prefix for batch transform.

If you choose **ManifestFile**, **S3Uri** identifies an object that is a manifest file containing a list of object keys that you want Amazon SageMaker to use for batch transform.

The following values are compatible: **ManifestFile, S3Prefix**

The following value is not compatible: **AugmentedManifestFile**

Type: String

Valid Values: ManifestFile | S3Prefix | AugmentedManifestFile

Required: Yes

S3Uri

Depending on the value specified for the **S3DataType**, identifies either a key name prefix or a manifest. For example:

- A key name prefix might look like this: `s3://bucketname/exampleprefix`.
- A manifest might look like this: `s3://bucketname/example.manifest`

The manifest is an S3 object which is a JSON file with the following format:

```json
[  
  {"prefix": "s3://customer_bucket/some/prefix/"},
  "relative/path/to/custdata-1",
  "relative/path/custdata-2",
  ...
  "relative/path/custdata-N"
]
```

The preceding JSON matches the following **S3Uris**:

- `s3://customer_bucket/some/prefix/relative/path/to/custdata-1`
- `s3://customer_bucket/some/prefix/relative/path/custdata-2`
- ...
- `s3://customer_bucket/some/prefix/relative/path/custdata-N`

The complete set of **S3Uris** in this manifest constitutes the input data for the channel for this datasource. The object that each **S3Uri** points to must be readable by the IAM role that Amazon SageMaker uses to perform tasks on your behalf.

Type: String
Length Constraints: Maximum length of 1024.

Pattern: ^((https|s3)://([^/]+)/*.*$^$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Trial
Service: Amazon SageMaker Service

The properties of a trial as returned by the Search (p. 691) API.

Contents

CreatedBy
  Who created the trial.
  
  Type: UserContext (p. 1449) object
  
  Required: No

CreationTime
  When the trial was created.
  
  Type: Timestamp
  
  Required: No

DisplayName
  The name of the trial as displayed. If DisplayName isn't specified, TrialName is displayed.
  
  Type: String
  
  Length Constraints: Minimum length of 1. Maximum length of 120.
  
  Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}
  
  Required: No

ExperimentName
  The name of the experiment the trial is part of.
  
  Type: String
  
  Length Constraints: Minimum length of 1. Maximum length of 120.
  
  Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}
  
  Required: No

LastModifiedBy
  Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
  
  Type: UserContext (p. 1449) object
  
  Required: No

LastModifiedTime
  Who last modified the trial.
  
  Type: Timestamp
  
  Required: No
**MetadataProperties**

Metadata properties of the tracking entity, trial, or trial component.

Type: MetadataProperties (p. 1143) object

Required: No

**Source**

The source of the trial.

Type: TrialSource (p. 1439) object

Required: No

**Tags**

The list of tags that are associated with the trial. You can use Search (p. 691) API to search on the tags.

Type: Array of Tag (p. 1377) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

**TrialArn**

The Amazon Resource Name (ARN) of the trial.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial/.*`

Required: No

**TrialComponentSummaries**

A list of the components associated with the trial. For each component, a summary of the component's properties is included.

Type: Array of TrialComponentSimpleSummary (p. 1432) objects

Required: No

**TrialName**

The name of the trial.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9\-]{0,119}$`

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:
• AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
TrialComponent
Service: Amazon SageMaker Service

The properties of a trial component as returned by the Search (p. 691) API.

Contents

CreatedBy
Who created the trial component.
Type: UserContext (p. 1449) object
Required: No

CreationTime
When the component was created.
Type: Timestamp
Required: No

DisplayName
The name of the component as displayed. If DisplayName isn't specified, TrialComponentName is displayed.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$
Required: No

EndTime
When the component ended.
Type: Timestamp
Required: No

InputArtifacts
The input artifacts of the component.
Type: String to TrialComponentArtifact (p. 1428) object map
Map Entries: Maximum number of 30 items.
Key Length Constraints: Maximum length of 64.
Key Pattern: .*
Required: No

LastModifiedBy
Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.
Type: UserContext (p. 1449) object
TrialComponent

Required: No

**LastModifiedTime**

When the component was last modified.

Type: Timestamp

Required: No

**LineageGroupArn**

The Amazon Resource Name (ARN) of the lineage group resource.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:lineage-group/.*`

Required: No

**MetadataProperties**

Metadata properties of the tracking entity, trial, or trial component.

Type: `MetadataProperties (p. 1143)` object

Required: No

**Metrics**

The metrics for the component.

Type: Array of `TrialComponentMetricSummary (p. 1429)` objects

Required: No

**OutputArtifacts**

The output artifacts of the component.

Type: `String to TrialComponentArtifact (p. 1428)` object map

Map Entries: Maximum number of 30 items.

Key Length Constraints: Maximum length of 64.

Key Pattern: `. *`

Required: No

**Parameters**

The hyperparameters of the component.

Type: `String to TrialComponentParameterValue (p. 1431)` object map

Map Entries: Maximum number of 150 items.

Key Length Constraints: Maximum length of 256.

Key Pattern: `. *`

Required: No
Parents
An array of the parents of the component. A parent is a trial the component is associated with and the experiment the trial is part of. A component might not have any parents.
Type: Array of Parent (p. 1251) objects
Required: No

Source
The Amazon Resource Name (ARN) and job type of the source of the component.
Type: TrialComponentSource (p. 1434) object
Required: No

SourceDetail
Details of the source of the component.
Type: TrialComponentSourceDetail (p. 1435) object
Required: No

StartTime
When the component started.
Type: Timestamp
Required: No

Status
The status of the trial component.
Type: TrialComponentStatus (p. 1436) object
Required: No

Tags
The list of tags that are associated with the component. You can use Search (p. 691) API to search on the tags.
Type: Array of Tag (p. 1377) objects
Array Members: Minimum number of 0 items. Maximum number of 50 items.
Required: No

TrialComponentArn
The Amazon Resource Name (ARN) of the trial component.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial-component/.*
Required: No

TrialComponentName
The name of the trial component.
Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrialComponentArtifact

Service: Amazon SageMaker Service

Represents an input or output artifact of a trial component. You specify TrialComponentArtifact as part of the InputArtifacts and OutputArtifacts parameters in the CreateTrialComponent request.

Examples of input artifacts are datasets, algorithms, hyperparameters, source code, and instance types. Examples of output artifacts are metrics, snapshots, logs, and images.

Contents

MediaType

The media type of the artifact, which indicates the type of data in the artifact file. The media type consists of a type and a subtype concatenated with a slash (/) character, for example, text/csv, image/jpeg, and s3/uri. The type specifies the category of the media. The subtype specifies the kind of data.

Type: String

Length Constraints: Maximum length of 64.

Pattern: ^[-\w]+/[-\w]+$

Required: No

Value

The location of the artifact.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: .*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrialComponentMetricSummary

Service: Amazon SageMaker Service

A summary of the metrics of a trial component.

Contents

Avg

The average value of the metric.

Type: Double

Required: No

Count

The number of samples used to generate the metric.

Type: Integer

Required: No

Last

The most recent value of the metric.

Type: Double

Required: No

Max

The maximum value of the metric.

Type: Double

Required: No

MetricName

The name of the metric.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: . +

Required: No

Min

The minimum value of the metric.

Type: Double

Required: No

SourceArn

The Amazon Resource Name (ARN) of the source.

Type: String
Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-\]:[0-9]{12}:.*`

Required: No

**StdDev**

The standard deviation of the metric.

Type: Double

Required: No

**TimeStamp**

When the metric was last updated.

Type: Timestamp

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**TrialComponentParameterValue**

Service: Amazon SageMaker Service

The value of a hyperparameter. Only one of `NumberValue` or `StringValue` can be specified.

This object is specified in the `CreateTrialComponent (p. 191)` request.

**Contents**

**NumberValue**

The numeric value of a numeric hyperparameter. If you specify a value for this parameter, you can't specify the `StringValue` parameter.

Type: Double

Required: No

**StringValue**

The string value of a categorical hyperparameter. If you specify a value for this parameter, you can't specify the `NumberValue` parameter.

Type: String

Length Constraints: Maximum length of 256.

Pattern: .*

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrialComponentSimpleSummary

Service: Amazon SageMaker Service

A short summary of a trial component.

Contents

CreatedBy

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Type: UserContext (p. 1449) object

Required: No

CreationTime

When the component was created.

Type: Timestamp

Required: No

TrialComponentArn

The Amazon Resource Name (ARN) of the trial component.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial-component/.*

Required: No

TrialComponentName

The name of the trial component.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9-]*(-*[a-zA-Z0-9])\{0,119\}

Required: No

TrialComponentSource

The Amazon Resource Name (ARN) and job type of the source of a trial component.

Type: TrialComponentSource (p. 1434) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
• AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
**TrialComponentSource**

Service: Amazon SageMaker Service

The Amazon Resource Name (ARN) and job type of the source of a trial component.

**Contents**

**SourceArn**

The source ARN.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:.*`

Required: Yes

**SourceType**

The source job type.

Type: String

Length Constraints: Maximum length of 128.

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrialComponentSourceDetail
Service: Amazon SageMaker Service

Detailed information about the source of a trial component. Either ProcessingJob or TrainingJob is returned.

Contents

ProcessingJob

Information about a processing job that's the source of a trial component.

Type: ProcessingJob (p. 1279) object

Required: No

SourceArn

The Amazon Resource Name (ARN) of the source.

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:.*

Required: No

TrainingJob

Information about a training job that's the source of a trial component.

Type: TrainingJob (p. 1385) object

Required: No

TransformJob

Information about a transform job that's the source of a trial component.

Type: TransformJob (p. 1405) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**TrialComponentStatus**

Service: Amazon SageMaker Service

The status of the trial component.

**Contents**

**Message**

If the component failed, a message describing why.

Type: String

Length Constraints: Maximum length of 1024.

Pattern: .*

Required: No

**PrimaryStatus**

The status of the trial component.

Type: String

Valid Values: InProgress | Completed | Failed | Stopping | Stopped

Required: No

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**TrialComponentSummary**

Service: Amazon SageMaker Service

A summary of the properties of a trial component. To get all the properties, call the *DescribeTrialComponent* (p. 467) API and provide the *TrialComponentName*.

**Contents**

**CreatedBy**

Who created the trial component.

Type: *UserContext* (p. 1449) object

Required: No

**CreationTime**

When the component was created.

Type: Timestamp

Required: No

**DisplayName**

The name of the component as displayed. If *DisplayName* isn't specified, *TrialComponentName* is displayed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: `^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,119}$`

Required: No

**EndTime**

When the component ended.

Type: Timestamp

Required: No

**LastModifiedBy**

Who last modified the component.

Type: *UserContext* (p. 1449) object

Required: No

**LastModifiedTime**

When the component was last modified.

Type: Timestamp

Required: No

**StartTime**

When the component started.
TrialComponentSummary

Type: Timestamp
Required: No

Status
The status of the component. States include:
• InProgress
• Completed
• Failed

Type: TrialComponentStatus (p. 1436) object
Required: No

TrialComponentArn
The ARN of the trial component.
Type: String
Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial-component/.*
Required: No

TrialComponentName
The name of the trial component.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.

Pattern: ^[a-zA-Z0-9\-\(\)*\+\-,\.;\[\]\]\{0,119\}$
Required: No

TrialComponentSource
The Amazon Resource Name (ARN) and job type of the source of a trial component.
Type: TrialComponentSource (p. 1434) object
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**TrialSource**
Service: Amazon SageMaker Service
The source of the trial.

**Contents**

**SourceArn**
The Amazon Resource Name (ARN) of the source.
Type: String
Length Constraints: Maximum length of 256.
Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:.*`
Required: Yes

**SourceType**
The source job type.
Type: String
Length Constraints: Maximum length of 128.
Required: No

**See Also**
For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TrialSummary
Service: Amazon SageMaker Service

A summary of the properties of a trial. To get the complete set of properties, call the DescribeTrial (p. 464) API and provide the TrialName.

Contents

CreationTime
When the trial was created.
Type: Timestamp
Required: No

DisplayName
The name of the trial as displayed. If DisplayName isn't specified, TrialName is displayed.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9]){0,119}$
Required: No

LastModifiedTime
When the trial was last modified.
Type: Timestamp
Required: No

TrialArn
The Amazon Resource Name (ARN) of the trial.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:experiment-trial/.*
Required: No

TrialName
The name of the trial.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 120.
Pattern: ^[a-zA-Z0-9](\-*[a-zA-Z0-9]){0,119}$
Required: No

TrialSource
The source of the trial.
Type: TrialSource (p. 1439) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**TuningJobCompletionCriteria**
Service: Amazon SageMaker Service

The job completion criteria.

**Contents**

**TargetObjectiveMetricValue**

The value of the objective metric.

Type: Float

Required: Yes

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
TuningJobStepMetaData
Service: Amazon SageMaker Service

Metadata for a tuning step.

Contents

Arn

The Amazon Resource Name (ARN) of the tuning job that was run by this step execution.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:hyper-parameter-tuning-job/.*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
UiConfig

Service: Amazon SageMaker Service

Provided configuration information for the worker UI for a labeling job. Provide either HumanTaskUiArn or UiTemplateS3Uri.

For named entity recognition, 3D point cloud and video frame labeling jobs, use HumanTaskUiArn.

For all other Ground Truth built-in task types and custom task types, use UiTemplateS3Uri to specify the location of a worker task template in Amazon S3.

Contents

HumanTaskUiArn

The ARN of the worker task template used to render the worker UI and tools for labeling job tasks.

Use this parameter when you are creating a labeling job for named entity recognition, 3D point cloud and video frame labeling jobs. Use your labeling job task type to select one of the following ARNs and use it with this parameter when you create a labeling job. Replace aws-region with the AWS Region you are creating your labeling job in. For example, replace aws-region with us-west-1 if you create a labeling job in US West (N. California).

Named Entity Recognition

Use the following HumanTaskUiArn for named entity recognition labeling jobs:


3D Point Cloud HumanTaskUiArns

Use this HumanTaskUiArn for 3D point cloud object detection and 3D point cloud object detection adjustment labeling jobs.

• arn:aws:sagemaker:aws-region:394669845002:human-task-ui/PointCloudObjectDetection

Use this HumanTaskUiArn for 3D point cloud object tracking and 3D point cloud object tracking adjustment labeling jobs.

• arn:aws:sagemaker:aws-region:394669845002:human-task-ui/PointCloudObjectTracking

Use this HumanTaskUiArn for 3D point cloud semantic segmentation and 3D point cloud semantic segmentation adjustment labeling jobs.

• arn:aws:sagemaker:aws-region:394669845002:human-task-ui/PointCloudSemanticSegmentation

Video Frame HumanTaskUiArns

Use this HumanTaskUiArn for video frame object detection and video frame object detection adjustment labeling jobs.

• arn:aws:sagemaker:region:394669845002:human-task-ui/VideoObjectDetection

Use this HumanTaskUiArn for video frame object tracking and video frame object tracking adjustment labeling jobs.

• arn:aws:sagemaker:aws-region:394669845002:human-task-ui/VideoObjectTracking
UiConfig

Type: String
Length Constraints: Maximum length of 1024.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:human-task-ui/.*
Required: No

UiTemplateS3Uri

The Amazon S3 bucket location of the UI template, or worker task template. This is the template used to render the worker UI and tools for labeling job tasks. For more information about the contents of a UI template, see Creating Your Custom Labeling Task Template.

Type: String
Length Constraints: Maximum length of 1024.
Pattern: ^(https|s3)://([^/]+)/(.*)$
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
UiTemplate
Service: Amazon SageMaker Service

The Liquid template for the worker user interface.

Contents

Content

The content of the Liquid template for the worker user interface.

Type: String
Pattern: \[\S\s\]+
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
UiTemplateInfo

Service: Amazon SageMaker Service

Container for user interface template information.

Contents

ContentSha256

The SHA-256 digest of the contents of the template.

Type: String


Required: No

Url

The URL for the user interface template.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
USD
Service: Amazon SageMaker Service

Represents an amount of money in United States dollars.

Contents

Cents
The fractional portion, in cents, of the amount.
Type: Integer
Valid Range: Minimum value of 0. Maximum value of 99.
Required: No

Dollars
The whole number of dollars in the amount.
Type: Integer
Valid Range: Minimum value of 0. Maximum value of 2.
Required: No

TenthFractionsOfACent
Fractions of a cent, in tenths.
Type: Integer
Valid Range: Minimum value of 0. Maximum value of 9.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
UserContext
Service: Amazon SageMaker Service

Information about the user who created or modified an experiment, trial, trial component, lineage group, or project.

Contents

DomainId
The domain associated with the user.
Type: String
Required: No

UserProfileArn
The Amazon Resource Name (ARN) of the user's profile.
Type: String
Required: No

UserProfileName
The name of the user's profile.
Type: String
Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
UserProfileDetails
Service: Amazon SageMaker Service
The user profile details.

Contents

CreationTime
The creation time.
Type: Timestamp
Required: No

DomainId
The domain ID.
Type: String
Length Constraints: Maximum length of 63.
Required: No

LastModifiedTime
The last modified time.
Type: Timestamp
Required: No

Status
The status.
Type: String
Valid Values: Deleting | Failed | InService | Pending | Updating | Update_Failed | Delete_Failed
Required: No

UserProfileName
The user profile name.
Type: String
Length Constraints: Maximum length of 63.
Pattern: ^[a-zA-Z0-9](-*[a-zA-Z0-9]){0,62}
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

• AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
UserSettings

Service: Amazon SageMaker Service

A collection of settings that apply to users of Amazon SageMaker Studio. These settings are specified when the CreateUserProfile API is called, and as DefaultUserSettings when the CreateDomain API is called.

SecurityGroups is aggregated when specified in both calls. For all other settings in UserSettings, the values specified in CreateUserProfile take precedence over those specified in CreateDomain.

Contents

ExecutionRole

The execution role for the user.
Type: String
Pattern: ^arn:aws[a-z\-]*:iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_\/]++$
Required: No

JupyterServerAppSettings

The Jupyter server’s app settings.
Type: JupyterServerAppSettings (p. 1115) object
Required: No

KernelGatewayAppSettings

The kernel gateway app settings.
Type: KernelGatewayAppSettings (p. 1116) object
Required: No

RSessionAppSettings

A collection of settings that configure the RSessionGateway app.
Type: RSessionAppSettings (p. 1345) object
Required: No

RStudioServerProAppSettings

A collection of settings that configure user interaction with the RStudioServerPro app.
Type: RStudioServerProAppSettings (p. 1346) object
Required: No

SecurityGroups

The security groups for the Amazon Virtual Private Cloud (VPC) that Studio uses for communication.
Optional when the CreateDomain.AppNetworkAccessType parameter is set to PublicInternetOnly.
Required when the CreateDomain.AppNetworkAccessType parameter is set to VpcOnly.
UserSettings

Amazon SageMaker adds a security group to allow NFS traffic from SageMaker Studio. Therefore, the number of security groups that you can specify is one less than the maximum number shown.

Type: Array of strings

Array Members: Maximum number of 5 items.

Length Constraints: Maximum length of 32.

Pattern: [-0-9a-zA-Z]+

Required: No

SharingSettings

Specifies options for sharing SageMaker Studio notebooks.

Type: SharingSettings (p. 1365) object

Required: No

TensorBoardAppSettings

The TensorBoard app settings.

Type: TensorBoardAppSettings (p. 1380) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
VariantProperty

Service: Amazon SageMaker Service

Specifies a production variant property type for an Endpoint.

If you are updating an endpoint with the UpdateEndpoint:RetainAllVariantProperties (p. 768) option set to true, the VariantProperty objects listed in UpdateEndpoint:ExcludeRetainedVariantProperties (p. 768) override the existing variant properties of the endpoint.

Contents

VariantPropertyType

The type of variant property. The supported values are:

- DesiredInstanceCount: Overrides the existing variant instance counts using the ProductionVariant:InitialInstanceCount (p. 1293) values in the CreateEndpointConfig:ProductionVariants (p. 75).
- DesiredWeight: Overrides the existing variant weights using the ProductionVariant:InitialVariantWeight (p. 1293) values in the CreateEndpointConfig:ProductionVariants (p. 75).
- DataCaptureConfig: (Not currently supported.)

Type: String

Valid Values: DesiredInstanceCount | DesiredWeight | DataCaptureConfig

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Vertex
Service: Amazon SageMaker Service
A lineage entity connected to the starting entity(ies).

Contents

Arn
The Amazon Resource Name (ARN) of the lineage entity resource.
Type: String
Length Constraints: Maximum length of 256.
Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:(experiment|experiment-trial-component|artifact|action|context)/.*
Required: No

LineageType
The type of resource of the lineage entity.
Type: String
Valid Values: TrialComponent | Artifact | Context | Action
Required: No

Type
The type of the lineage entity resource. For example: DataSet, Model, Endpoint, etc...
Type: String
Length Constraints: Maximum length of 40.
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
VpcConfig
Service: Amazon SageMaker Service

Specifies a VPC that your training jobs and hosted models have access to. Control access to and from your training and model containers by configuring the VPC. For more information, see Protect Endpoints by Using an Amazon Virtual Private Cloud and Protect Training Jobs by Using an Amazon Virtual Private Cloud.

Contents
SecurityGroupIds

The VPC security group IDs, in the form sg-xxxxxxx. Specify the security groups for the VPC that is specified in the Subnets field.

Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 5 items.
Length Constraints: Maximum length of 32.
Pattern: [-0-9a-zA-Z]+
Required: Yes

Subnets

The ID of the subnets in the VPC to which you want to connect your training job or model. For information about the availability of specific instance types, see Supported Instance Types and Availability Zones.

Type: Array of strings
Array Members: Minimum number of 1 item. Maximum number of 16 items.
Length Constraints: Maximum length of 32.
Pattern: [-0-9a-zA-Z]+
Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
**Workforce**

Service: Amazon SageMaker Service

A single private workforce, which is automatically created when you create your first private work team. You can create one private work force in each AWS Region. By default, any workforce-related API operation used in a specific region will apply to the workforce created in that region. To learn how to create a private workforce, see Create a Private Workforce.

**Contents**

**CognitoConfig**

The configuration of an Amazon Cognito workforce. A single Cognito workforce is created using and corresponds to a single Amazon Cognito user pool.

Type: CognitoConfig (p. 950) object

Required: No

**CreateDate**

The date that the workforce is created.

Type: Timestamp

Required: No

**LastUpdatedDate**

The most recent date that UpdateWorkforce (p. 811) was used to successfully add one or more IP address ranges (CIDRs) to a private workforce's allow list.

Type: Timestamp

Required: No

**OidcConfig**

The configuration of an OIDC Identity Provider (IdP) private workforce.

Type: OidcConfigForResponse (p. 1235) object

Required: No

**SourceIpConfig**

A list of one to ten IP address ranges (CIDRs) to be added to the workforce allow list. By default, a workforce isn't restricted to specific IP addresses.

Type: SourceIpConfig (p. 1369) object

Required: No

**SubDomain**

The subdomain for your OIDC Identity Provider.

Type: String

Required: No

**WorkforceArn**

The Amazon Resource Name (ARN) of the private workforce.
Workforce

Type: String

Length Constraints: Maximum length of 256.

Pattern: arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workforce/.*

Required: Yes

WorkforceName

The name of the private workforce.

Type: String


Pattern: ^[a-zA-Z0-9][a-zA-Z0-9\-]{0,62}$

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Workteam
Service: Amazon SageMaker Service
Provides details about a labeling work team.

Contents

CreateDate
The date and time that the work team was created (timestamp).
Type: Timestamp
Required: No
Description
A description of the work team.
Type: String
Pattern: .+
Required: Yes
LastUpdatedDate
The date and time that the work team was last updated (timestamp).
Type: Timestamp
Required: No
MemberDefinitions
A list of MemberDefinition objects that contains objects that identify the workers that make up the work team.
Workforces can be created using Amazon Cognito or your own OIDC Identity Provider (IdP). For private workforces created using Amazon Cognito use CognitoMemberDefinition. For workforces created using your own OIDC identity provider (IdP) use OidcMemberDefinition.
Type: Array of MemberDefinition (p. 1142) objects
Array Members: Minimum number of 1 item. Maximum number of 10 items.
Required: Yes
NotificationConfiguration
Configures SNS notifications of available or expiring work items for work teams.
Type: NotificationConfiguration (p. 1229) object
Required: No
ProductListingIds
The Amazon Marketplace identifier for a vendor's work team.
Type: Array of strings
Required: No

**SubDomain**

The URI of the labeling job's user interface. Workers open this URI to start labeling your data objects.

Type: String

Required: No

**WorkforceArn**

The Amazon Resource Name (ARN) of the workforce.

Type: String

Length Constraints: Maximum length of 256.

Pattern: \[arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workforce/.*\]

Required: No

**WorkteamArn**

The Amazon Resource Name (ARN) that identifies the work team.

Type: String

Length Constraints: Maximum length of 256.

Pattern: \[arn:aws[a-z\-]*:sagemaker:[a-z0-9\-]*:[0-9]{12}:workteam/.*\]

Required: Yes

**WorkteamName**

The name of the work team.

Type: String


Pattern: ^\[a-zA-Z0-9\-\]{0,62}\}

Required: Yes

---

**See Also**

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3

---

**Amazon SageMaker Runtime**

The following data types are supported by Amazon SageMaker Runtime:
Amazon Sagemaker Edge Manager

The following data types are supported by Amazon Sagemaker Edge Manager:

- EdgeMetric (p. 1462)
- Model (p. 1463)
EdgeMetric
Service: Amazon Sagemaker Edge Manager

Information required for edge device metrics.

Contents

**Dimension**
- The dimension of metrics published.
  Type: String
  Length Constraints: Minimum length of 1. Maximum length of 1000.
  Pattern: `^[a-zA-Z0-9]-*[^a-zA-Z0-9]$/`*
  Required: No

**MetricName**
- Returns the name of the metric.
  Type: String
  Pattern: `^[a-zA-Z0-9]-*[^a-zA-Z0-9]$/`
  Required: No

**Timestamp**
- Timestamp of when the metric was requested.
  Type: Timestamp
  Required: No

**Value**
- Returns the value of the metric.
  Type: Double
  Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:
- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Model
Service: Amazon SageMaker Edge Manager

Information about a model deployed on an edge device that is registered with SageMaker Edge Manager.

Contents

LatestInference
The timestamp of the last inference that was made.
Type: Timestamp
Required: No

LatestSampleTime
The timestamp of the last data sample taken.
Type: Timestamp
Required: No

ModelMetrics
Information required for model metrics.
Type: Array of EdgeMetric (p. 1462) objects
Required: No

ModelName
The name of the model.
Type: String
Pattern: ^[a-zA-Z0-9-]*[a-zA-Z0-9]+$
Required: No

ModelVersion
The version of the model.
Type: String
Length Constraints: Minimum length of 1. Maximum length of 64.
Pattern: [a-zA-Z0-9\-_\.]+$
Required: No

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
Amazon SageMaker Feature Store Runtime

The following data types are supported by Amazon SageMaker Feature Store Runtime:

- BatchGetRecordError (p. 1465)
- BatchGetRecordIdentifier (p. 1467)
- BatchGetRecordResultDetail (p. 1468)
- FeatureValue (p. 1469)
BatchGetRecordError
Service: Amazon SageMaker Feature Store Runtime

The error that has occurred when attempting to retrieve a batch of Records.

Contents

ErrorCode
The error code of an error that has occurred when attempting to retrieve a batch of Records. For more information on errors, see Errors.

Type: String
Length Constraints: Maximum length of 358400.
Pattern: . *
Required: Yes

ErrorMessage
The error message of an error that has occurred when attempting to retrieve a record in the batch.

Type: String
Length Constraints: Maximum length of 2048.
Required: Yes

FeatureGroupName
The name of the feature group that the record belongs to.

Type: String
Length Constraints: Maximum length of 358400.
Pattern: . *
Required: Yes

RecordIdentifierValueAsString
The value for the RecordIdentifier in string format of a Record from a FeatureGroup that is causing an error when attempting to be retrieved.

Type: String
Length Constraints: Maximum length of 358400.
Pattern: . *
Required: Yes

See Also
For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
• AWS SDK for Java V2
• AWS SDK for Ruby V3
BatchGetRecordIdentifier

Service: Amazon SageMaker Feature Store Runtime

The identifier that identifies the batch of Records you are retrieving in a batch.

Contents

FeatureGroupName

A FeatureGroupName containing Records you are retrieving in a batch.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9\-]*[a-zA-Z0-9]\

Required: Yes

FeatureNames

List of names of Features to be retrieved. If not specified, the latest value for all the Features are returned.

Type: Array of strings

Array Members: Minimum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9\-\_]\([\-\_]\*[a-zA-Z0-9]\)*

Required: No

RecordIdentifiersValueAsString

The value for a list of record identifiers in string format.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Maximum length of 358400.

Pattern: .*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
BatchGetRecordResultDetail

Service: Amazon SageMaker Feature Store Runtime

The output of Records that have been retrieved in a batch.

Contents

FeatureGroupName

The FeatureGroupName containing Records you retrieved in a batch.

Type: String

Length Constraints: Maximum length of 358400.

Pattern: . *

Required: Yes

Record

The Record retrieved.

Type: Array of FeatureValue (p. 1469) objects

Array Members: Minimum number of 1 item.

Required: Yes

RecordIdentifierValueAsString

The value of the record identifier in string format.

Type: String

Length Constraints: Maximum length of 358400.

Pattern: . *

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
FeatureValue

Service: Amazon SageMaker Feature Store Runtime

The value associated with a feature.

Contents

FeatureName

The name of a feature that a feature value corresponds to.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9][-_*][a-zA-Z0-9]*

Required: Yes

ValueAsString

The value associated with a feature, in string format. Note that features types can be String, Integral, or Fractional. This value represents all three types as a string.

Type: String

Length Constraints: Maximum length of 358400.

Pattern: .*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Go
- AWS SDK for Java V2
- AWS SDK for Ruby V3
Common Parameters

The following list contains the parameters that all actions use for signing Signature Version 4 requests with a query string. Any action-specific parameters are listed in the topic for that action. For more information about Signature Version 4, see Signature Version 4 Signing Process in the Amazon Web Services General Reference.

**Action**
- The action to be performed.
  - Type: string
  - Required: Yes

**Version**
- The API version that the request is written for, expressed in the format YYYY-MM-DD.
  - Type: string
  - Required: Yes

**X-Amz-Algorithm**
- The hash algorithm that you used to create the request signature.
  - Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.
  - Type: string
  - Valid Values: AWS4-HMAC-SHA256
  - Required: Conditional

**X-Amz-Credential**
- The credential scope value, which is a string that includes your access key, the date, the region you are targeting, the service you are requesting, and a termination string ("aws4_request"). The value is expressed in the following format: \access_key/YYYYMMDD\/region\/service/aws4_request. For more information, see Task 2: Create a String to Sign for Signature Version 4 in the Amazon Web Services General Reference.
  - Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.
  - Type: string
  - Required: Conditional

**X-Amz-Date**
- The date that is used to create the signature. The format must be ISO 8601 basic format (’YYYYMMDD’T’HHMMSS’Z’). For example, the following date time is a valid X-Amz-Date value: 20120325T120000Z.
  - Condition: X-Amz-Date is optional for all requests; it can be used to override the date used for signing requests. If the Date header is specified in the ISO 8601 basic format, X-Amz-Date is
not required. When X-Amz-Date is used, it always overrides the value of the Date header. For more information, see Handling Dates in Signature Version 4 in the Amazon Web Services General Reference.

Type: string
Required: Conditional

**X-Amz-Security-Token**

The temporary security token that was obtained through a call to AWS Security Token Service (AWS STS). For a list of services that support temporary security credentials from AWS Security Token Service, go to AWS Services That Work with IAM in the IAM User Guide.

Condition: If you're using temporary security credentials from the AWS Security Token Service, you must include the security token.

Type: string
Required: Conditional

**X-Amz-Signature**

Specifies the hex-encoded signature that was calculated from the string to sign and the derived signing key.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string
Required: Conditional

**X-Amz-SignedHeaders**

Specifies all the HTTP headers that were included as part of the canonical request. For more information about specifying signed headers, see Task 1: Create a Canonical Request For Signature Version 4 in the Amazon Web Services General Reference.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string
Required: Conditional
Common Errors

This section lists the errors common to the API actions of all AWS services. For errors specific to an API action for this service, see the topic for that API action.

**AccessDeniedException**
You do not have sufficient access to perform this action.
HTTP Status Code: 400

**IncompleteSignature**
The request signature does not conform to AWS standards.
HTTP Status Code: 400

**InternalFailure**
The request processing has failed because of an unknown error, exception or failure.
HTTP Status Code: 500

**InvalidAction**
The action or operation requested is invalid. Verify that the action is typed correctly.
HTTP Status Code: 400

**InvalidClientTokenId**
The X.509 certificate or AWS access key ID provided does not exist in our records.
HTTP Status Code: 403

**InvalidParameterCombination**
Parameters that must not be used together were used together.
HTTP Status Code: 400

**InvalidParameterValue**
An invalid or out-of-range value was supplied for the input parameter.
HTTP Status Code: 400

**InvalidQueryParameter**
The AWS query string is malformed or does not adhere to AWS standards.
HTTP Status Code: 400

**MalformedQueryString**
The query string contains a syntax error.
HTTP Status Code: 404

**MissingAction**
The request is missing an action or a required parameter.
HTTP Status Code: 400
MissingAuthenticationToken

The request must contain either a valid (registered) AWS access key ID or X.509 certificate.

HTTP Status Code: 403

MissingParameter

A required parameter for the specified action is not supplied.

HTTP Status Code: 400

NotAuthorized

You do not have permission to perform this action.

HTTP Status Code: 400

OptInRequired

The AWS access key ID needs a subscription for the service.

HTTP Status Code: 403

RequestExpired

The request reached the service more than 15 minutes after the date stamp on the request or more than 15 minutes after the request expiration date (such as for pre-signed URLs), or the date stamp on the request is more than 15 minutes in the future.

HTTP Status Code: 400

ServiceUnavailable

The request has failed due to a temporary failure of the server.

HTTP Status Code: 503

ThrottlingException

The request was denied due to request throttling.

HTTP Status Code: 400

ValidationError

The input fails to satisfy the constraints specified by an AWS service.

HTTP Status Code: 400