

AWS ホワイトペーパー

Amazon Forecast による時系列予測の原理



Amazon Forecast による時系列予測の原理: AWS ホワイトペーパー

Copyright © 2023 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon の商標およびトレードドレスは、Amazon のものではない製品またはサービスと関連付けてはならず、また、お客様に混乱を招くような形や Amazon の信用を傷つけたり失わせたりする形で使用することはできません。Amazon が所有しない商標はすべてそれぞれの所有者に所属します。所有者は必ずしも Amazon と提携していたり、関連しているわけではありません。また、Amazon 後援を受けているとはかぎりません。

Table of Contents

要約と概要	i
概要	1
優れた設計をしていますか?	2
予測について	3
予測システム	3
予測の問題はどのような分野で発生するのですか。	3
予測の問題の解決を目指す前の考慮事項	4
ケーススタディ: e コマース企業の小売需要予測の問題	6
ステップ 1: データを収集して集計する	9
例	10
ステップ 2: データを準備する	12
欠損データの処理方法	12
例 1	12
例 2	14
特徴量化および関連時系列の概念	15
例 3	15
ステップ 3: 予測子を作成する	17
ステップ 4: 予測子を評価する	19
バックテスト	19
予測分位数と精度指標	21
重み付き分位損失 (wQL)	21
重み付き絶対誤差率 (WAPE)	22
二乗平均平方根誤差 (RMSE)	22
WAPE と RMSE に関する問題	23
ステップ 5: 意思決定のための予測を生成して使用する	25
確率的予測	25
視覚化	26
予測ワークフローと API の概要	28
一般的なシナリオで Amazon Forecast を使用する	29
予測を本番環境に導入する	30
まとめ	31
寄稿者	32
その他の資料	33
付録 A: よくある質問	34

付録 B: 参考文献	38
ドキュメント履歴	39
注意	40
AWS の用語集	41

Amazon Forecast による時系列予測の原理

公開日: 2021 年 9 月 1 日 ([ドキュメント履歴](#))

企業は現在、単純なスプレッドシートから複雑な財務計画ソフトウェアにおよぶすべてのシステムを使用して、製品の需要、リソースのニーズ、財務パフォーマンスなどの今後のビジネス成果の正確な予測を実現することに取り組んでいます。このペーパーでは、予測とその用語、課題、ユースケースを紹介します。このドキュメントでは、ケーススタディを使用して予測のコンセプトと予測手順を説明し、さらに [Amazon Forecast](#) が実際の予測の問題において数多くの現実的な課題に対処するうえでどのように役立つかについて説明します。

概要

予測は将来を予測する科学です。履歴データを使用することで、企業は傾向を理解し、どのようなことがどのタイミングで起こるかを判断し、この情報を製品需要から在庫計画、人員配置まで、将来に向けたあらゆる計画に組み込むことができます。

予測がもたらす結果を考えると、精度が重要となります。予測値が高すぎると、製品や人財に過剰投資し、その投資が無駄になる可能性があります。予測値が低すぎると、過小投資となり、原材料や在庫が不足し、顧客体験の低下につながる可能性があります。

企業は現在、単純なスプレッドシートから複雑な財務計画ソフトウェアにおよぶすべてのシステムを使用して、予想を生成していますが、以下の 2 つの理由により、精度の高い予測を入手することは困難です。

- まず、従来の予測では、大量の履歴データの組み込み作業が困難を極め、結果、ノイズにより過去データの重要なシグナルを見逃してしまうことがあります。
- 次に、重要なコンテキスト (価格、休日やイベント、品切れ、マーケティングプロモーションなど) を提供できる関連性があるデータが独立して管理されており、従来の予測に組み込まれることはほとんどありません。完全な履歴と広範囲のコンテキストがなければ、ほとんどの予測は将来を正確に予測することができません。

[Amazon Forecast](#) は、このような問題を解決するフルマネージドサービスです。Amazon Forecast は、お客様の予測シナリオに最適なアルゴリズムを提供しています。必要に応じてモダンな機械学習 (ML) と深層学習を活用して、精度の高い予測を実現します。Amazon Forecast は使いやすく、機械学習の経験は必要ありません。このサービスでは、必要なインフラストラクチャを自動的にプロビジョニングし、データを処理し、AWS 上でホストし、カスタムまたはプライベート機械学習モデル

を構築するため、すぐに予測を作成できます。さらに、急速に進化を遂げている機械学習技術の進歩についても、Amazon Forecast には組み込まれているため、お客様は最小限の追加作業で精度の向上を継続して確認できます。

優れた設計をしていますか？

[AWS Well-Architected フレームワーク](#)は、クラウドでシステムを構築する際に行う決定の長所と短所を理解するのに役立ちます。フレームワークの6つの柱により、信頼性が高く、安全かつ効率的で、費用対効果が高く、持続可能なシステムを設計および運用するためのアーキテクチャのベストプラクティスを学ぶことができます。[AWS Management Console](#)で無料で提供されている[AWS Well-Architected Tool](#)を使用し、各柱の一連の質問に回答することで、これらのベストプラクティスに照らしてワークロードを確認できます。

[Machine Learning Lens](#)では、AWS クラウドの機械学習ワークロードの設計方法、デプロイ方法、アーキテクチャ方法に焦点を当てています。このレンズは、Well-Architected フレームワークに説明されるベストプラクティスに追加されるものです。

クラウドアーキテクチャに関するより専門的なガイダンスとベストプラクティス (リファレンスアーキテクチャのデプロイ、図、ホワイトペーパー) については、[AWSアーキテクチャセンター](#)を参照してください。

予測について

このドキュメントにおいて、予測とは時系列の将来の値を予測することを意味します。問題への入力または出力は時系列の性質のもので、

予測システム

予測システムには、次のようなさまざまなユーザーが関与します。

- エンドユーザーは、特定の製品の予測のクエリを実行し、購入すべきユニットの数を決定します。決定を行うのは、個人の場合も、自動化システムの場合もあります。
- ビジネスアナリストまたはビジネスインテリジェンスは、エンドユーザーをサポートし、集計レポートを実行して整理します。
- データサイエンティストは、需要パターンや因果関係を繰り返し分析し、新しい機能を追加してモデルを段階的に改善したり、予測モデルを改善したりします。
- エンジニアは、データ収集のインフラストラクチャをセットアップして、システムへの入力データの可用性を確保します。

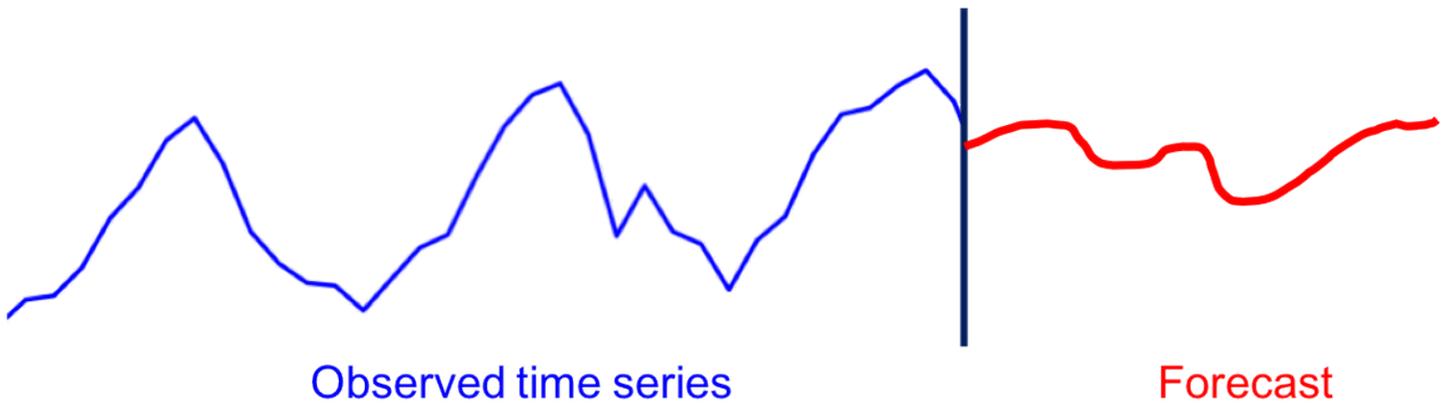
Amazon Forecast を使用すると、ソフトウェアエンジニアの労力が軽減し、データサイエンス能力が限られている企業でも最先端の予測技術を活用できます。データサイエンス能力のある企業の場合は、Amazon Forecast で提供されているさまざまな診断機能を利用して、予測の問題に適切に対処できます。

予測の問題はどのような分野で発生するのですか。

予測の問題は、自然に時系列データが生成される数多くの分野で発生します。これには、小売販売、医療分析、キャパシティプランニング、センサーネットワークモニタリング、財務分析、社会的活動のマイニング、データベースシステムがあります。例えば、データ駆動型意思決定を行っている企業のほとんどでは、業務プロセスの自動化と最適化において予測が重要な役割を果たしています。製品の需要と供給の予測は、最適な在庫管理、スタッフのスケジューリング、トポロジープランニングに利用できます。さらに一般的には、サプライチェーン最適化のほとんどの側面において、予測は重要なテクノロジーとなっています。

あるパターン (この例では季節性) を示す観測された時系列に基づいて、特定の期間にわたって予測を作成した場合の予測の問題の概要が、次の図で説明されています。横軸は、過去 (左) から将来

(右) にわたる時系列を表しています。縦軸は測定単位を表しています。過去 (青色) から黒色の縦線までの現在の測定値を元に作成されているのが、将来 (赤色) の予測タスクです。



予測タスクの概要

予測の問題の解決を目指す前の考慮事項

予測の問題を解決する前に理解しておくべき最も重要な質問事項は次のとおりです。

- 予測の問題を解決する必要がありますか。
- 予測の問題を解決する理由はなんですか。

時系列データは至る所で発生しており、予測の問題は場所を問わずすぐに見つかります。ただし、予測の問題を解決する必要があるのか、それともビジネス上の効率的な意思決定を犠牲にすることなく予測の問題を完全に回避できるのか、ということが重要な問題となります。科学的観点では、予測は機械学習で最も難しい問題の1つであり、このような問いを投げかけることが重要です。

例えば、オンライン小売業者のお勧め商品を考えてみましょう。このようなお勧め商品の問題は、お客様と SKU (stock keeping unit) のペアごとに、この特定のお客様が購入するであろう特定の商品のユニット数を予測するという予測の問題として表現することができます。このように問題を定式化することには多くの利点があります。時間の要素が明確に考慮されるため、お客様の購入パターンに応じて商品をお勧めできることが、利点の1つです。

ただし、お勧め商品の問題が予測の問題として定式化されることはほとんどありません。このような予測の問題を解決することは、(お客様と SKU のレベルでの情報がわずかであることや、問題の規模などにより) お勧めに関する問題を直接解決するよりもはるかに難しいためです。したがって、予測アプリケーションについて検討する場合、予測の下流での使用状況と、別の方法でこの問題に対処できるかどうかを考慮することが重要となります。

このような場合、[Amazon Personalize](#) が役に立ちます。Amazon Personalize は、デベロッパーがアプリケーションを使用しているお客様に応じて個別のレコメンデーションを簡単に作成できる機械学習サービスです。

予測の問題を解決する必要があるという結論に至った場合、次に尋ねるべき質問は、予測の問題を解決する理由です。ビジネス環境では、通常ほとんどの場合、予測は目的を達成するための手段として利用されています。例えば、小売業界での需要予測の場合、その予測を使用して在庫管理の意思決定を行うことができます。予測の問題は通常、意思決定の問題に対する入力情報であり、これを最適化問題としてモデル化する場合もあります。

このような意思決定上の問題の例には、購入するユニット数や、既存の在庫を処理する最善の方法などがあります。ビジネス予測に関するその他の問題としては、サーバーのキャパシティ予測や、製造業における原材料や部品の需要予測などがあります。このような予測は、上記のように意思決定の問題やシナリオシミュレーションなど、その他のプロセスへの入力情報として使用できます。シナリオシミュレーションはその後、明示的なモデルを使用することなくプランニングに使用できます。予測自体は目的ではないと言っても、例外があります。例えば、財務予測では、予測は手持ち資金の準備に直接使用されたり、投資家に提示されたりする場合があります。

予測の目的を理解するには、次の質問を検討してください。

- 予測の対象とするのは、今後どれくらいの期間ですか。
- 予測はどのくらいの頻度で生成する必要がありますか。
- 予測には、深く掘り下げるべき特定の側面はありますか。

ケーススタディ: e コマース企業の小売需要予測の問題

予測のコンセプトの詳細を説明するために、オンラインで商品を販売している、ある e コマース企業のケースを考えてみましょう。サプライチェーンについての意思決定 (在庫管理など) を最適化すると、適切な配送拠点に正確な数の商品の在庫を維持することができるため、この企業の中心的な競争力という点で重要です。つまりこれは、より短い配送時間と競争力のある価格で幅広い選択肢を用意することを意味します。これが、顧客満足度の向上につながります。このサプライチェーンソフトウェアシステムへの重要な入力データとなるのが、カタログ内のすべての商品の需要予測または潜在的な売上予測です。この予測により、下流での重要な意思決定が実現します。中でも重要な予測には以下があります。

- マクロレベルのプランニング (戦略的予測): 企業全体として、総売上高や収益の観点から予測される成長率はどれくらいですか。この企業が地理的に (さらに) 力を入れるべき場所はどこですか。人事配置はどのようにすべきですか。
- 需要 (または在庫) 予測: 各商品の拠点別の販売数見込みはどれくらいですか。
- プロモーション活動 (戦術的予測): どのようなプロモーションを展開すべきですか。清算すべき商品はありますか。

ケーススタディの後半では、運用上の予測の問題 (Januschowski & Kolassa、2019 年) の一部である 2 つ目の問題に焦点を当てます。このドキュメントでは、データ、モデル (予測子)、推論 (予測)、プロダクト化などの重要な点について説明しています。

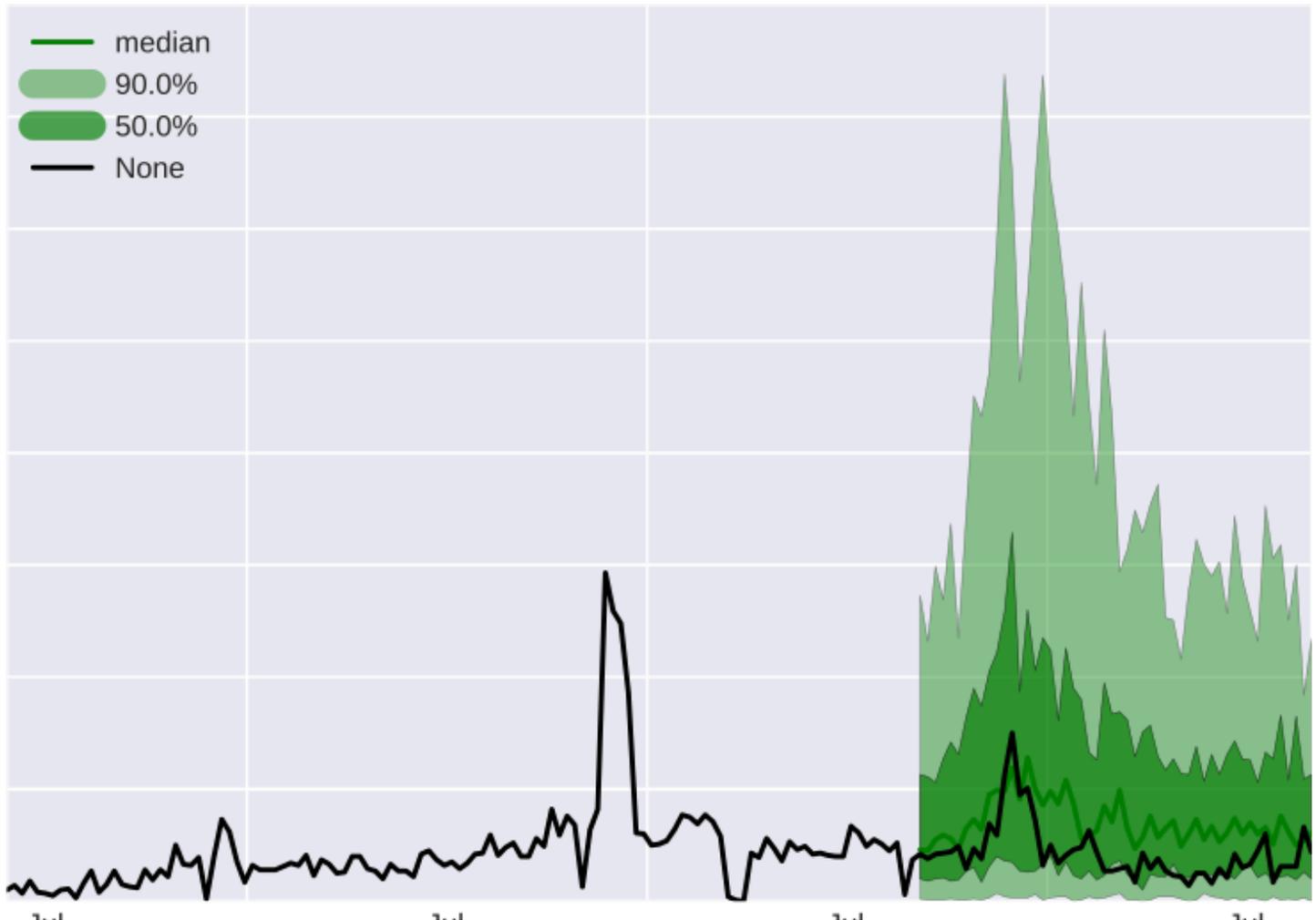
このケーススタディの場合、予測の問題が目的を達成するための手段であることに留意することが重要です。予測はこの企業にとって非常に重要であるとはいえ、下流のサプライチェーンにおける意思決定はさらに重要です。このケーススタディでは、意思決定はオペレーションズリサーチの数学的最適化モデルを基とする自動購入システムが行っています。これらのシステムでは、この企業の見込まれるコストを最小限に抑えることが目標です。

ここでのキーワードは「見込み」です。つまり、予測は、特定の結果の確率に応じて適切な重み付けを行い、将来の 1 つの可能性ではなく、考えられるすべての将来の可能性を対象とする必要があります。そのためには、単にポイント予測を行うのではなく、予測値の分布を得ることが下流の意思決定を実現する鍵となります。次の図は、確率的予測 (密度予測とも呼ばれる) を説明しています。この確率的予測から単一点における予測 (最も可能性の高い将来) を簡単に導き出すことはできますが、単一点の予測から確率的予測への移行は、より難しくなることに注意してください。

確率的予測があれば、そこからさまざまな統計情報を取得し、意思決定に役立つ結果を調整できます。この e コマース企業には、在庫切れを避けるべき主要な商品が数多くある可能性があります。

この場合、高分位数 (例えば、90 パーセンタイル) を使用してください。これは、商品の在庫がある時間の 90% に相当します。代替品を見つけやすい製品 (鉛筆など) など、その他の商品については、低パーセンタイルを使用することが、より適切である場合があります。

Amazon Forecast では、確率的予測からさまざまな分位数を簡単に取得できます。



確率的予測の説明図

上の図の黒い線は実際の値、濃い緑の線は予測分布の中央値、濃い緑色の影付きの領域は値の 50% があると予想される予測区間、薄緑色の領域は実際の値の 90% があると予想される予測区間です。

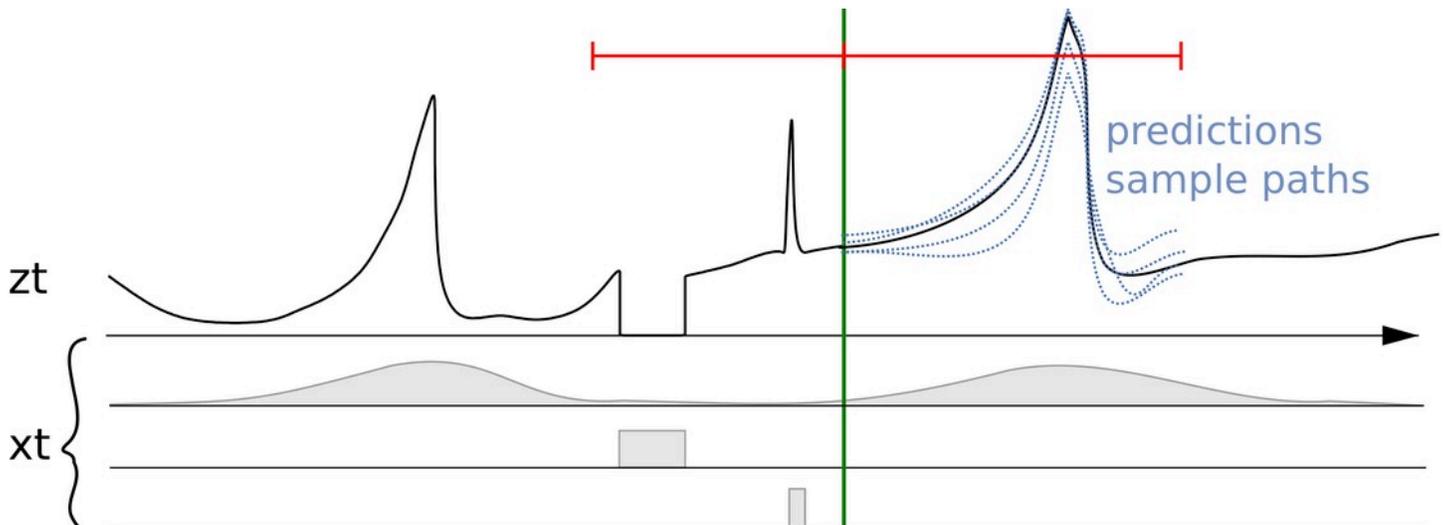
次のセクションでは、この企業における予測の問題を解決するために必要な手順について説明します。これには以下が含まれます。

- [データを収集して集計する \(ステップ 1\)](#)
- [データを準備する \(ステップ 2\)](#)
- [予測子を作成する \(ステップ 3\)](#)

- [予測子を評価する \(ステップ 4\)](#)
- [予測生成を自動化する \(ステップ 5\)](#)

ステップ 1: データを収集して集計する

次の図は、予測の問題に関するメンタルモデルを説明しています。可能な限り予測の精度を向上するために、できるだけ多くの関連情報を使用して、時系列 z_t を将来にわたり予測することが目標です。そのため、最初の最も重要なステップは、できるだけ多くの正確なデータを収集することです。



時系列 z_t と関連する特徴または共変量 (x_t)、および複数の予測

上の図では、縦線の右側に複数の予測が表示されています。これらの予測は、確率的予測分布の例です (逆の言い方をすれば、確率的予測を示すために使用できます)。

小売業が記録すべき重要な情報は次のとおりです。

- 取引売上データ — SKU (Stock Keeping Unit)、場所、タイムスタンプ、販売数など。
- SKU 商品詳細データ — 商品のメタデータ。例としては、色、部門、サイズなどがあります。
- 価格データ — タイムスタンプが付随した各商品の時系列の価格。
- プロモーション情報データ — 商品コレクション (カテゴリ) またはタイムスタンプが付随した個別の商品についてのさまざまなタイプのプロモーション。
- 在庫情報データ — SKU が在庫にあるか購入可能ある場合と SKU が在庫切れである場合を比較する情報が時間単位ごとに表示されます。
- 位置情報データ — 特定の時点での商品または販売の拠点は、`location_idstore_id` 文字列または実際の位置情報として表示できます。位置情報には、国コードと 5 桁の郵便番号、または `latitude_longitude` を使用できます。拠点は、取引販売の「ディメンション」として扱われます。

[Amazon Forecast](#) では、予測対象の数量の履歴データをターゲット時系列 (TTS) と呼びます。小売業の場合、この TTS が取引の売上データです。すべての販売取引と完全に同じタイミングで入手できるその他の履歴データは、関連時系列 (RTS) と呼ばれます。小売業の場合、RTS には価格、プロモーション、在庫変数などがあります。

この問題では売上ではなく需要の予測に焦点をあてているにもかかわらず、この企業で記録されているのは売上のみであるため、在庫情報が重要であることに注意してください。SKU が在庫切れになると、販売数は潜在的な需要を下回るため、このような在庫切れのイベントが発生した日時を把握して記録することが重要です。

考慮すべきその他のデータセットには、ウェブページ訪問数、検索用語の詳細、ソーシャルメディア、気象情報などがあります。多くの場合、このデータをモデルで使用するには、データ履歴と将来のデータがあることが重要です。これは、多くの予測モデルやバックテスト ([ステップ 4: 予測を評価する](#) セクションで説明) の要件となっています。

予測の問題によっては、未加工データの頻度が予測の問題の頻度と自然に一致することがあります。この例としては、分単位での予測の場合に、分単位でサンプリングされるサーバーボリュームのリクエストがあります。

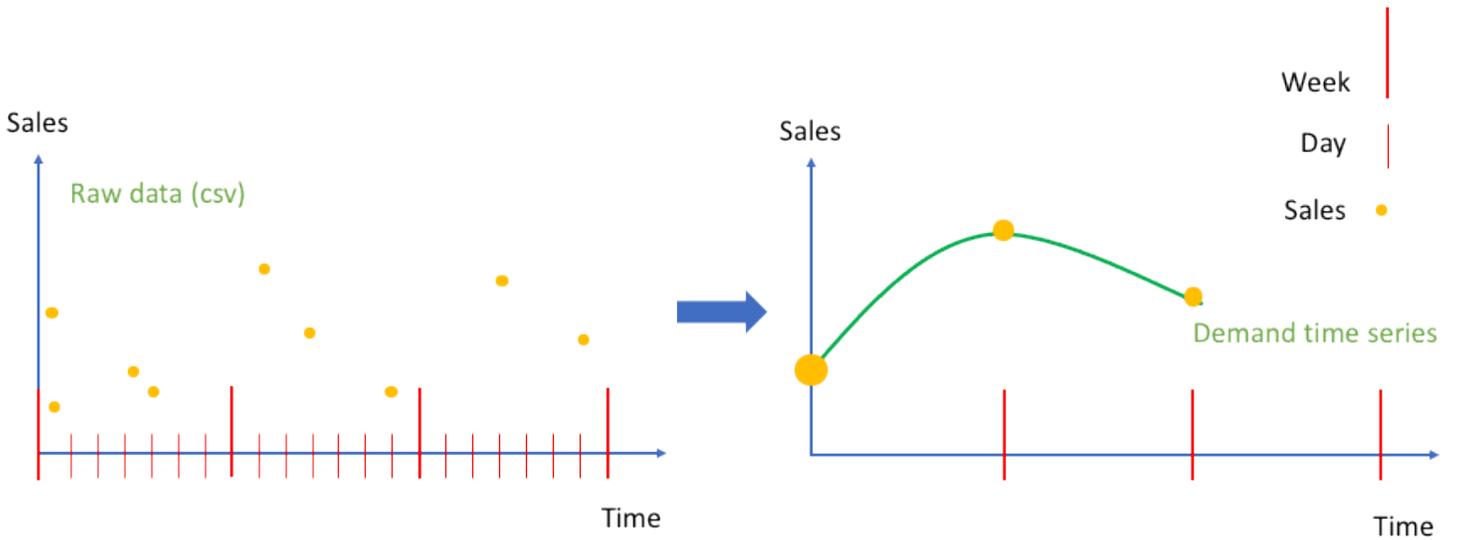
多くの場合、データはより細かい頻度で記録されるか、時間範囲内の任意のタイムスタンプで記録されますが、予測の問題の場合の詳細度はより粗くなります。これは小売業のケーススタディによく見られます。販売データは通常、取引データとして記録されます。取引データの形式は、例えば、販売が行われた日時をきめ細かく記録したタイムスタンプで構成されています。予測のユースケースでは、このような詳細度は必要ない場合があります。このデータを時間単位または日単位の売上に集計する方が適切な場合があります。このケースでは、集計のレベルは、例えば在庫管理やリソースプランニングなど、下流の問題に対応するものとしします。

例

次の図の左のグラフは、カンマ区切り値 (CSV) ファイルとして Amazon Forecast に入力できる未加工の顧客売上データの例を示しています。この例では、売上データはより細かい日次時間のグリッドで定義されています。予測の問題の場合、より粗い時間グリッドである週次需要を将来的に予測します。Amazon Forecast では、`create_predictor` API コールで特定の週の日次値を集計します。

集計された結果、未加工データは、週ごとに固定した頻度形式での時系列コレクションに変換されます。右のグラフはこの集計を、デフォルトの合計集計方法を使用したターゲット時系列で表示しています。その他の集計方法には、平均化、最大化、最小化、単一点 (最初の時点など) の選択などがあります。データが企業のユースケースに最も適したものとなるように、集

計の詳細度と方法を選択する必要があります。この例では、集計値は週次集計に対応しています。その他の集計方法は、`create_predictor` API の `FeaturizationConfig` パラメータの `FeaturizationMethodParameters` キーを使用してユーザーが設定することができます。



イベントである未加工の売上データ (左) を等間隔の時系列に集計した図 (右)

ステップ 2: データを準備する

未加工データを入手したら、欠損データなどの複雑な問題を処理し、意図した解釈を最も的確に捉える予測モデルのためにデータを準備する必要があります。

欠損データの処理方法

実際に予測を実施するうえでよくある問題として、未加工データに欠損値が存在することが挙げられます。時系列で値が欠落しているということは、指定された頻度のすべての時点に対応する値がないため、それ以上処理できないことを意味します。値が欠損していると判定される理由は複数考えられます。

欠損値は、トランザクションが存在しない、または測定エラー (特定のデータを監視していたサービスが正しく動作していなかった、測定が正しく行われなかったなど) があつたなどの理由で発生します。小売のケーススタディの場合、後者の主な例として、需要予測における在庫切れの状況があります。つまり、需要がその日の売上に等しくないということです。

クラウドコンピューティングのシナリオでも、サービスが上限に達した場合 (特定の [AWS リージョン](#) の [Amazon EC2](#) インスタンスがすべてビジー状態の場合など)、同様の影響が生じる可能性があります。欠損値のもう 1 つの例は、製品やサービスがまだ発売されていない場合や、生産が終了した場合です。

欠損値を特徴量処理コンポーネントで挿入し、パディングによって時系列の長さを揃えることもできます。欠損値が多ければ、モデルの精度に大きな影響を与える可能性があります。

例 1

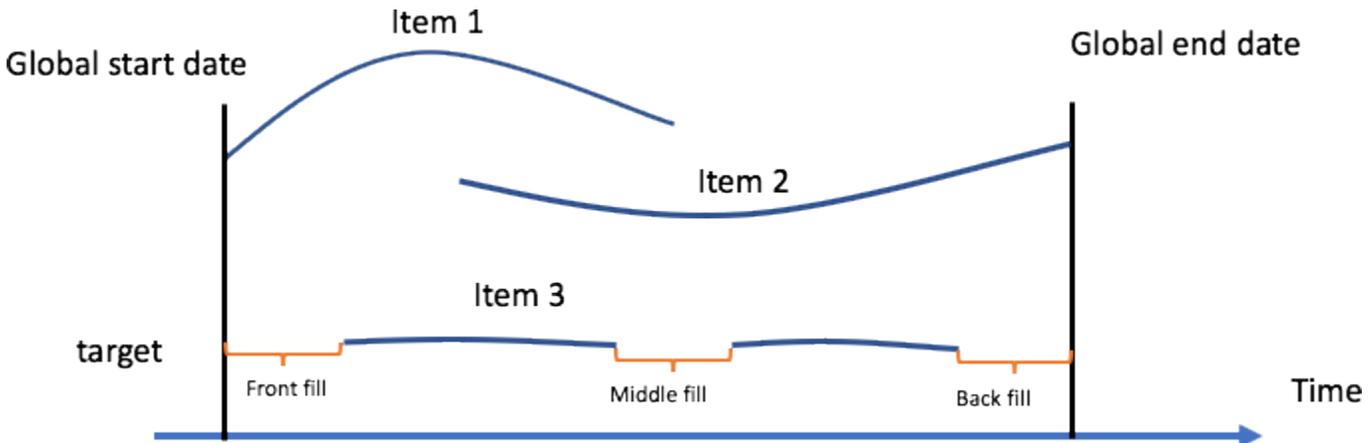
補完は、データセット内の欠落しているエントリに標準化された値を追加するプロセスです。次の図では、3 つの項目からなるデータセットの項目 2 について、Amazon Forecast で欠損値を処理するためのさまざまな方法 (前方補完、中間補完、後方補完、予測補完) を示しています。

Amazon Forecast は、ターゲット時系列と関連時系列の両方の補完に対応しています。グローバル開始日は、データセット内のすべての項目の開始日のうち、最初の開始日として定義されます。下記の例では、グローバル開始日は項目 1 で発生します。同様に、グローバル終了日は、すべての項目の時系列のうち最後の終了日として定義され、項目 2 で発生します。

前方補完では、特定の時系列の開始日からグローバル開始日までのすべての値が入力されます。本書の公開時点では、Amazon Forecast ではいかなる前方補完も有効にならず、すべての時系列を異な

る時点から開始できます。中間補完は時系列の中間 (アイテムの開始日と終了日の間など) に補完された値を表し、後方補完はその時系列の最終日付からグローバル終了日までの値を補完します。

ターゲット時系列の場合、中間と後方の補完メソッドのデフォルトの補完ロジックはゼロです。予測補完 (関連時系列にのみ適用) は、項目のグローバル終了日と、お客様が指定した予測期間との間の欠損値を補完します。[Prophet](#) と [DeepAR+](#) で関連時系列データセットを使用する場合は将来の値が必要で、[CNN-QR](#) の場合は任意です。



Amazon Forecast での欠損値の処理方法

前図では、グローバル開始日はすべての項目の開始日のうち最初の開始日を示し、グローバル終了日はすべての項目の終了日のうち最後の終了日を示しています。予測期間は、Forecast がターゲット値の予測を提供する期間です。

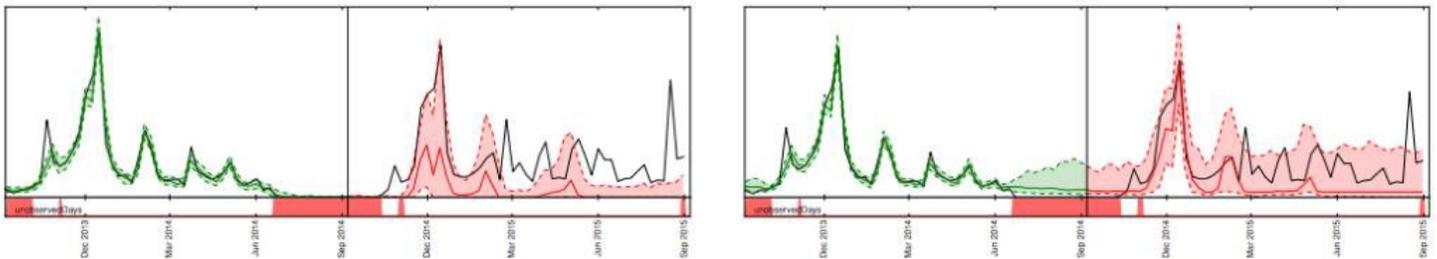
これは、小売業界の調査でよく見られるシナリオで、販売可能な商品の取引データで販売数がゼロになっているというものです。これらの値は真のゼロとして扱われ、指標評価コンポーネントで使用されます。Amazon Forecast では、ユーザーが実際に欠落している値を特定し、アルゴリズムが処理する非数 (NaN) としてエンコードできます。本書では次に、この 2 つのケースが異なる理由と、それぞれがどのような場合に役立つかを説明します。

小売のケーススタディでは、販売可能な商品の販売個数がゼロであったという情報は、その商品が存在しない期間 (発売前や廃止後など) または存在している期間 (一部在庫切れだった場合や、販売データが記録されていなかった期間があった場合など) に、販売不可能な商品の販売個数がゼロであったという情報とは異なります。前者の場合は、デフォルトのゼロ補完が適用されます。後者の場合、対応するターゲット値は通常ゼロですが、値が欠落しているとマークされることで付加的な情報が伝達されます。欠落しているデータがあったという情報は保存し、この情報を破棄しないことがベストプラクティスです。情報を保持することが重要である理由については、次の例で説明します。

Amazon Forecast では、値、平均、中央値、最小値、最大値という追加の補完ロジックをサポートしています。関連時系列 (価格やキャンペーンなど) では、正しい欠損値のロジックが属性タイプや

ユースケースによって異なるため、中間、後方、予測の補完メソッドにデフォルトのロジックは指定されていません。関連時系列でサポートされている入力ロジックには、ゼロ、値、平均、中央値、最小値、最大値があります。

欠損値の補完を実行するには、[CreatePredictor](#) オペレーションを呼び出すときに実装する補完のタイプを指定します。補完ロジックは [FeaturizationMethod](#) オブジェクトで指定されます。たとえば、ターゲットの時系列で入手できない商品の売上がゼロではない値をエンコードするには、補完タイプを NaN に設定して、値を本当に欠落しているものとしてマークします。ゼロ補完とは異なり、NaN でエンコードされた値は欠落しているものとして扱われ、指標評価コンポーネントでは使用されません。



同じ項目の予測に対する 0 補完と NaN による補完の効果の違い

前図の左のグラフでは、黒い縦線の左側の値が 0 で補完されているため、アンダーバイアス気味の予測 (黒い縦線の右側) になっています。右のグラフでは、これらの値は NaN とマークされており、適切な予測につながります。

例 2

上の図は、[ARIMA](#) や [ETS](#) などの線形状態空間モデルで欠損値を正しく処理することの重要性を示しています。一部在庫切れになっている項目の需要予測をグラフ化しており、トレーニング領域は左のグラフに緑で、予測範囲は右のパネルで赤で、真のターゲットは黒で表示されています。中央値、p10、p90 の予測は、それぞれ赤い線と影付きの領域に示されています。最下部には、在庫切れの商品 (データの 80%) が赤でマークされています。左のグラフでは、在庫切れの領域は無視され、0 で補完されています。

その結果、予測モデルでは、予測されるゼロが多数あると仮定され、予測値が低くなりすぎてしまいます。右のグラフでは、在庫切れ領域は真の欠測として扱われ、在庫切れ領域では需要が不確実になります。在庫切れ商品の欠損値が NaN と適切にマークされているため、このグラフの予測範囲にアンダーバイアスは見られません。Amazon Forecast はこのようなデータのギャップを埋めるため、入力データをすべて明示的に変更しなくても、欠落しているデータを簡単に正しく処理できます。

特徴量化および関連時系列の概念

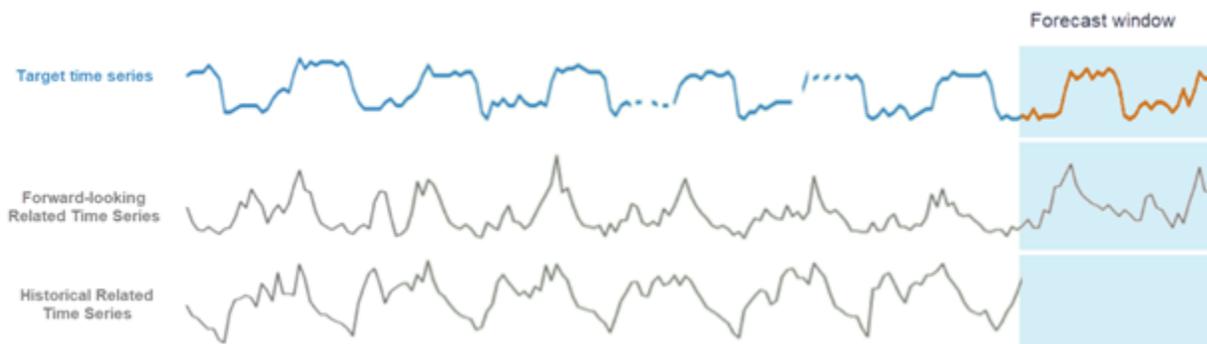
Amazon Forecast では、サポートされている特定の予測モデルの精度を向上させるのに役立つ関連データをユーザーが入力できます。このデータには、関連時系列と静的な項目メタデータの 2 種類があります。

Note

メタデータと関連データは、機械学習では特徴量と呼ばれ、統計では共変量と呼ばれます。

関連時系列とは、ターゲット値と何らかの相関関係があり、直感的に説明できるため、ターゲット値の予測に統計的に役立つはずですが (例については、「[Amazon Forecast: predicting time-series at scale](#)」(Amazon Forecast: 大規模な時系列予測) をご覧ください)。ターゲット時系列とは異なり、関連時系列とは、ターゲット時系列に影響を与える可能性のある過去の既知の値であり、将来既知の値を持つ可能性があります。

Amazon Forecast では、過去の時系列と将来の時系列の 2 種類の関連時系列を追加できます。過去に関連時系列には、予測期間までのデータポイントが含まれ、将来の予測期間内のデータポイントは含まれません。将来の関連時系列には、予測期間までのデータポイントと予測期間内のデータポイントが含まれます。



Amazon Forecast で関連時系列を使用する際のさまざまなアプローチ

例 3

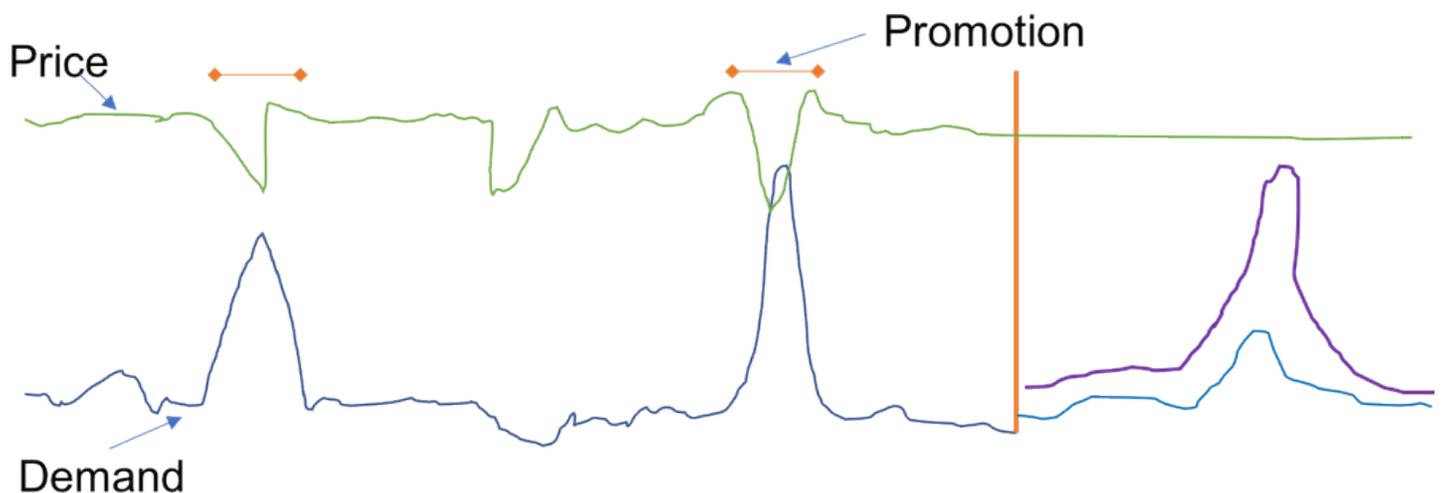
次の図は、関連時系列を使用して、人気のある本の将来の需要を予測する方法の例を示しています。青い線は、ターゲット時系列の需要を表しています。価格は緑の線で表されます。縦線は予測開始日を表し、2 つの分位数での予測が縦線の右側に表示されています。

この例では、予測粒度についてターゲット時系列と一致し、予測開始日から予測期間によって増分された予測開始日 (予測終了日) までの範囲で、将来のすべて (またはほとんど) の時点で既知である、将来の関連時系列を使用しています。

また、次の図では、価格の低下と商品の売上の上昇に相関関係が見られることから、価格が使用に適した特徴量であることもわかります。関連時系列は、商品の SKU、タイムスタンプ、関連時系列値 (この場合は価格) を含む別の CSV ファイルを使用して Amazon Forecast に読み込むことができます。

Amazon Forecast は、ターゲット時系列の平均や合計などの集計方法をサポートしていますが、関連時系列ではサポートしていません。たとえば、日次価格を週次価格に合計してもほとんど意味がありません。日次キャンペーンでも同様です。

Amazon Forecast では、組み込みの特徴量化済みデータセットを含めることで、[天気](#)や[休日](#)の情報をモデルに自動的に組み込むことができます (「[SupplementaryFeature](#)」を参照)。気象情報や休日は、小売需要に大きな影響を与える可能性があります。



特定の商品の売上 (青色、赤色の縦線の左側)

Amazon Forecast に入力できるその他の有用な機能として、項目メタデータ (カテゴリ変数とも呼ばれる) があります (例については「[Amazon Forecast: predicting time-series at scale](#)」(Amazon Forecast: 大規模な時系列予測) を参照)。カテゴリ変数と関連時系列の主な違いは、カテゴリ変数は静的であり、時間の経過とともに変化しないことです。一般的な小売の例としては、商品の色、書籍のカテゴリ、テレビがスマートテレビかどうかを示す二値の指標などがあります。この情報は、類似の SKU の売上が類似していると仮定して、深層学習アルゴリズムによって在庫管理単位 (SKU) 間の類似点を学習することで取得できます。このメタデータには時間依存性がないため、項目メタデータ CSV ファイルの各行は、商品 SKU と対応する分類ラベルまたは説明のみで構成されます。

ステップ 3: 予測子を作成する

予測変数は 2 つの方法で作成できます。1 つは [AutoML](#) を実行する方法で、もう 1 つは Amazon Forecast に組み込まれている 6 つのアルゴリズムのうちの 1 つを手動で選択する方法です。AutoML を実行すると、Amazon Forecast は 6 つの組み込みアルゴリズム (本書の作成時) を自動的にテストし、P10、P50 (中央値)、P90 の予測値から算出した平均分位損失が最も小さいアルゴリズムを選択します。

Amazon Forecast には 4 つのローカルモデルがあります。

- 自己回帰和分移動平均 ([ARIMA](#))
- 指数平滑法 ([ETS](#))
- ノンパラメトリック時系列 ([NPTS](#))
- [Prophet](#)

ローカルモデルとは、1 つのモデルを個々の時系列 (または特定の項目とディメンションの組み合わせ) に合わせ、これらのモデルを使用して時系列を将来に当てはめる予測メソッドです。

ARIMA と ETS は、R 予測パッケージの一般的なローカルモデルのスケラブルバージョンです。Amazon で開発されたローカルメソッドである NPTS は、他のローカルモデルと比較すると大きな違いがあります。前回の値や適切な季節での値を繰り返してポイント予測を提供する単純な季節予測機能とは異なり、NPTS は確率的な予測を行います。NPTS は固定時間インデックスを使用します。ここで、前のインデックス ($T - 1$) または過去のシーズン ($T - \tau$) がタイムステップ T の予測です。アルゴリズムは、セット $\{0, \dots, T - 1\}$ の時間インデックス (t) をランダムにサンプリングして、現在のタイムステップ T のサンプルを生成します。NPTS は、ゼロの多い断続的な (スパースとも呼ばれる) 時系列データに特に効果的です。Forecast には、ベイズ構造時系列モデルである Prophet の Python 実装も含まれています。

Amazon Forecast には、次の 2 つのグローバル深層学習アルゴリズムが用意されています。

- [DeepAR+](#)
- [CNN-QR](#)

グローバルモデルは、データセット内の時系列の集合全体を対象に 1 つのモデルをトレーニングします。これは、一連の横断的な単位にわたって類似した時系列データがある場合に特に便利です。た

たとえば、さまざまな商品に対する需要、サーバーの負荷、ウェブページへのリクエストを時系列でグループ化することができます。

一般に、時系列データの数が増えるほど、CNN-QR と DeepAR+ の有効性が高まります。これはローカルモデルには必ずしも当てはまりません。また、深層学習モデルを使用して、過去の売上データがほとんどない、またはまったくない新しい SKU の予測を生成することもできます。これは [コールドスタート予測](#) として知られています。

	Neural Networks		Flexible Local Algorithms	Baseline Algorithms		
	CNN-QR	DeepAR+	Prophet	NPTS	ARIMA	ETS
Computationally intensive training process	High	High	Medium	Low	Low	Low
Accepts historical related time series*	✔	✘	✘	✘	✘	✘
Accepts forward-looking related time series*	✔	✔	✔	✘	✘	✘
Accepts item metadata (product color, brand, etc)	✔	✔	✘	✘	✘	✘
Suitable for sparse datasets	✔	✔	✘	✔	✘	✘
Performs Hyperparameter Optimization (HPO)	✔	✔	✘	✘	✘	✘
Allows overriding default hyperparameter values	✔	✔	✘	✔	✘	✘
Suitable for What-if analysis	✔	✔	✔	✘	✘	✘
Suitable for Cold Start scenarios (forecasting with little to no historical data)	✔	✔	✘	✘	✘	✘

Amazon Forecast で利用できるアルゴリズムの比較

関連時系列の詳細については、「[Related Time Series](#)」(関連時系列) を参照してください。

ステップ 4: 予測子を評価する

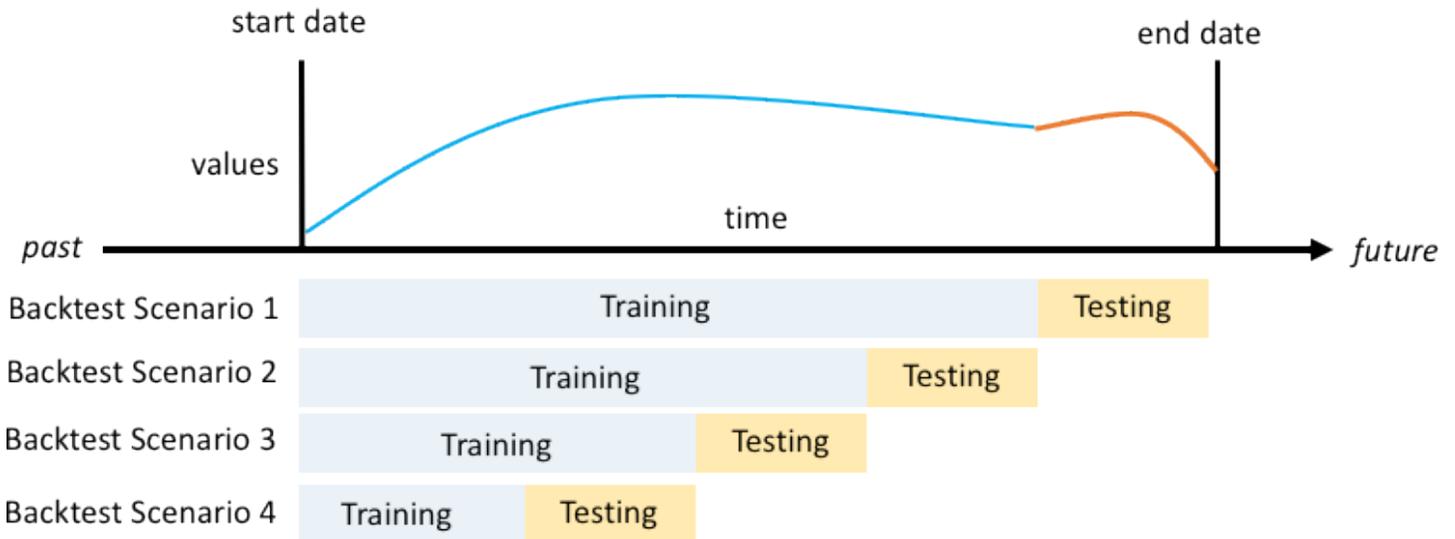
機械学習の一般的なワークフローでは、モデル一式またはモデルの組み合わせをトレーニングセットでトレーニングし、ホールドアウトデータセットでその精度を評価します。このセクションでは、履歴データを分割する方法と、時系列予測でモデルを評価するために使用する指標について説明します。予測では、バックテスト手法が予測精度を評価するための主要なツールとなっています。

バックテスト

適切な評価とバックテストのフレームワークは、機械学習アプリケーションを成功に導く上で最も重要な要素の 1 つです。モデルのバックテストが成功すれば、モデルの将来の予測能力に自信が持てるようになります。さらに、ハイパーパラメーター最適化 (HPO) によるモデルのチューニング、モデルの組み合わせの学習、メタ学習と AutoML の有効化可能です。

時系列予測が、評価やバックテストの方法論の観点から、応用機械学習の他の分野とは異なるのは、その特性である時間に起因しています。通常、ML タスクでは、バックテストの予測エラーを評価するために、データセットを項目ごとに分割します。たとえば、画像関連のタスクでの交差検証では、何割かの画像でトレーニングを行い、それ以外の部分をテストと検証に使用します。予測では、トレーニングセットからテストセットや検証セットに情報が漏れないように、また本稼働環境でのケースをできる限り厳密にシミュレートするために、主に時間別に (また主にではないが項目別に) 分割する必要があります。

1 つの時点を選択するのではなく、複数の時点を選択する必要があるため、時間による分割は慎重に行う必要があります。そうしないと、分割点によって定義される予測開始日に精度が影響されすぎてしまいます。複数の時点で一連の分割を行い、平均結果を出力するローリング予測評価では、より確実で信頼性の高いバックテスト結果が得られます。次の図は、4 つの異なるバックテスト分割を示しています。



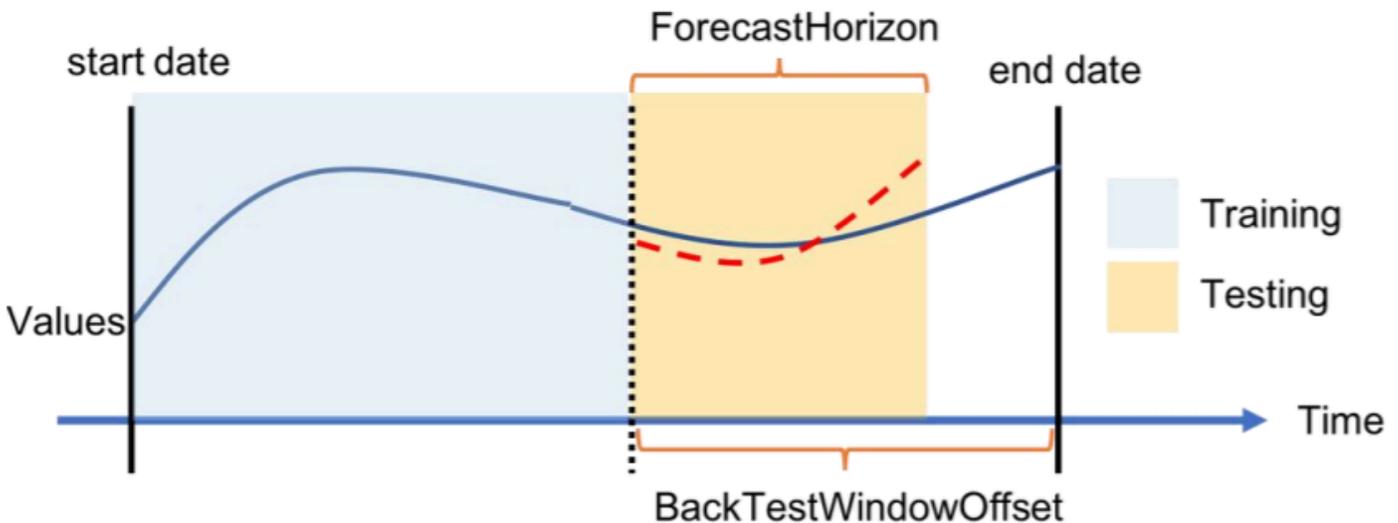
トレーニングセットのサイズは大きくなるが、テストのサイズは一定である 4 つの異なるバックテストシナリオの図

上の図では、すべてのバックテストシナリオに、予測値を実際の値と比較して評価できるように、全体を通して利用可能なデータがあります。

複数のバックテスト時間枠が必要な理由は、現実世界のほとんどの時系列が通常は非定常であるためです。このケーススタディの e コマースビジネスは北米を拠点としており、その製品需要の多くは第 4 四半期に集中し、特に感謝祭前後とクリスマス前にピークを迎えます。第 4 四半期のショッピングシーズンでは、時系列の変動が一年の他の時期よりも大きくなります。バックテストの時間枠を複数用意することで、よりバランスの取れた設定で予測モデルを評価できます。

各バックテストシナリオについて、次の図は Amazon Forecast の用語の基本要素を示しています。Amazon Forecast は、データをトレーニングデータセットとテストデータセットに自動的に分割します。Amazon Forecast は、`create_predictor` API でパラメータとして指定されている `BackTestWindowOffset` パラメータまたは `ForecastHorizon` のデフォルト値を使用して、入力データをどのように分割するかを決定します。

次の図では、`BackTestWindowOffset` パラメータと `ForecastHorizon` パラメータが等しくない場合の、より一般的な前者のケースを示しています。この `BackTestWindowOffset` パラメータは、仮定の予測開始日を定義します。次の図では、縦の破線で示されています。これを使うと、この日にモデルが導入されたら、どのような予測になるかというような仮説的な疑問に答えることができます。 `ForecastHorizon` は、仮定予測開始日から予測するタイムステップ数を定義します。



Amazon Forecast での単一バックテストシナリオとその設定の図

Amazon Forecast は、バックテスト中に生成された予測値と精度指標をエクスポートできます。エクスポートされたデータを使用して、特定の時点と分位数における特定の項目を評価できます。

予測分位数と精度指標

予測分位数では、予測値の上限と下限を指定できます。たとえば、0.1 (P10)、0.5 (P50)、0.9 (P90) の予測タイプを使用すると、P50 予測に対して 80% 信頼区間と呼ばれる値の範囲が得られます。P10、P50、P90 で予測を生成すると、真の値が 80% の確率でこの範囲に入ることが期待できます。

このホワイトペーパーでは、[ステップ 5 で分位数についてさらに説明しています](#)。

Amazon Forecast では、バックテスト中に重み付き分位損失 (wQL)、二乗平均平方根誤差 (RMSE)、重み付き絶対誤差率 (WAPE) の精度指標を使用して予測子を評価します。

重み付き分位損失 (wQL)

重み付き分位損失 (wQL) エラー指標は、指定された分位点におけるモデルの予測の精度を測定します。これは、過小予測と過大予測のコストが異なる場合に特に有用です。wQL 関数の重み (τ) を設定すると、過小予測と過大予測に対して異なるペナルティが自動的に組み込まれます。

$$wQL[\tau] = 2 \frac{\sum_{i,t} [\tau \max(y_{i,t} - q_{i,t}^{(\tau)}, 0) + (1 - \tau) \max(q_{i,t}^{(\tau)} - y_{i,t}, 0)]}{\sum_{i,t} |y_{i,t}|}$$

wQL 関数

各パラメータの意味は次のとおりです。

- τ — セット $\{0.01, 0.02, \dots, 0.99\}$ 内の分位数
- $q_{i,t}(\tau)$ — モデルが予測する τ 分位数。
- $y_{i,t}$ — ポイント (i,t) での観測値

重み付き絶対誤差率 (WAPE)

重み付き絶対誤差率 (WAPE) は、モデルの精度を測定するために一般的に使用される指標です。予測値と観測値の全体的な偏差を測定します。

$$\text{WAPE} = \frac{\sum_{i,t} |y_{i,t} - \hat{y}_{i,t}|}{\sum_{i,t} |y_{i,t}|}$$

WAPE

各パラメータの意味は次のとおりです。

- $y_{i,t}$ - ポイント (i,t) での観測値
- $\hat{y}_{i,t}$ - ポイント (i,t) での予測値

Forecast では、平均予測値を予測値 $\hat{y}_{i,t}$ として使用します。

二乗平均平方根誤差 (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{nT} \sum_{i,t} (\hat{y}_{i,t} - y_{i,t})^2}$$

二乗平均平方根誤差 (RMSE) は、モデルの精度を測定するために一般的に使用される指標です。WAPE と同様に、予測値と観測値の全体的な偏差を測定します。

各パラメータの意味は次のとおりです。

- $y_{i,t}$ - ポイント (i,t) での観測値

- $\hat{y}_{i,t}$ - ポイント (i,t) での予測値
- nT - テストセット内のデータポイントの数

Forecast では、平均予測値を予測値 $\hat{y}_{i,t}$ として使用します。予測指標を計算する場合、nT はバックテスト時間枠内のデータポイントの数です。

WAPE と RMSE に関する問題

ほとんどの場合、内部で、または他の予測ツールから生成できるポイント予測は、p50 分位数または平均予測値と一致する必要があります。WAPE と RMSE のどちらについても、Amazon Forecast は平均予測を使用して予測値 (yhat) を表します。

wQL [tau] 方程式の tau = 0.5 の場合、両方の重みが等しくなり、wQL [0.5] はポイント予測で一般的に使用されている重み付き絶対誤差率 (WAPE) になります。

$$wQL[0.5] = 2 \frac{\sum_{i,t} 0.5 [\max(y_{i,t} - q_{i,t}^{(0.5)}, 0) + \max(q_{i,t}^{(0.5)} - y_{i,t}, 0)]}{\sum_{i,t} |y_{i,t}|} = \frac{\sum_{i,t} |y_{i,t} - q_{i,t}^{(0.5)}|}{\sum_{i,t} |y_{i,t}|}$$

ここでは、yhat = q(0.5) が計算予測です。wQL の式では、倍率 2 を使用して 0.5 係数を相殺し、正確な WAPE[中央値] 式を求めます。

上記の WAPE の定義は、平均絶対誤差率 ([MAPE](#)) の一般的な解釈とは異なることに注意してください。違いは分母にあります。WAPE を上記のように定義すると、0 による除算の問題が回避されます。この問題は、このケーススタディの e コマースビジネスのように、特定の SKU の 1 日の販売数が 0 になることが多いという現実のシナリオでよく発生する問題です。

tau が 0.5 に等しくない場合の重み付き分位損失指標とは異なり、各分位数に内在する偏りは、重みが等しい WAPE のような計算では捉えられません。WAPE のその他の欠点としては、対称性がない、数値が小さいと誤差率が過度に大きくなる、単なる点単位の指標であることなどが挙げられます。

RMSE は WAPE の誤差項の二乗で、他の機械学習アプリケーションでよく使用される誤差指標です。誤差の変動が大きいと RMSE は過度に高くなるため、RMSE 指標は個々の誤差の大きさが一定のモデルに適しています。二乗誤差のため、不適切な予測値がいくつかあるだけで、他の点では良好な予測でも、RMSE が高くなる可能性があります。また、二乗項があるため、誤差項が小さいほど、WAPE よりも RMSE での方が重みが小さくなります。

精度指標を使用すると、予測を定量的に評価できます。特に大規模な比較 (メソッド A はメソッド B より総合的に優れているのか) の場合、こうした精度指標は非常に重要です。ただし、多くの場合、これを個々の SKU のビジュアルで補うことが重要です。

ステップ 5: 意思決定のための予測を生成して使用する

特定のユースケースに求められる精度を満足するモデルができたなら (バックテストで確認)、最後のステップはモデルをデプロイして予測を生成することです。Amazon Forecast にモデルをデプロイするには、Create_Forecast API を実行する必要があります。このアクションは、履歴データセット全体でトレーニングして作成されたモデルをホストします (データをトレインセットとテストセットに分割する Create_Predictor とは異なります)。予測対象期間で生成されたモデル予測は、次の 2 つの方法で利用できます。

- 特定の商品の予測を (商品または商品とディメンションの組み合わせを指定して) [AWS CLI](#) で Query_Forecast API を使用するか、直接 [AWS Management Console](#) 経由でクエリできます。
- Create_Forecast_Export_Job API を使用して、すべての分位数にわたる商品とディメンションのすべての組み合わせの予測を生成できます。この API は、生成した CSV ファイルを [Amazon Simple Storage Service](#) (Amazon S3) の選択した場所に安全に保存します。その後、この CSV ファイルのデータを使用して、意思決定に使用する下流システムに取り込むことができます。例えば、既存のサプライチェーンシステムでは、Amazon Forecast の出力を直接取り込んで、特定の SKU の製造に関する意思決定に役立てることができます。

確率的予測

Amazon Forecast では、さまざまな分位数での予測を生成できます。これは、実測値が予測値を下回る場合と上回る場合でそれぞれコストが異なる場合に特に役立ちます。予測子のトレーニング段階と同様に、p1 と p99 の間の分位数について確率的予測を生成できます。

Amazon Forecast は、デフォルトで予測子のトレーニング中に使用した分位数と同じ分位数で予測を生成します。予測子のトレーニング中に分位数が指定されていない場合、予測はデフォルトで p10、p50、p90 で生成されます。

p10 予測とは、正しい値が予測値よりも低くなる確率が 10% であると予想されるものであり、wQL [0.1] 指標を使用して精度を評価できます。つまり、P10 の予測は 90% の確率で実際の値を下回ることを示しており、在庫予測で利用する場合、90% の確率で商品が売切れとなります。P10 の予測は、保管スペースがあまりない場合や、投資する資本のコストが高い場合に役立つ場合があります。

Note

分位数予測の正式な定義は、 $\Pr(\text{実際の値} \leq \text{分位点 } q \text{ での予測}) = q$ です。厳密には、分位数は百分位数 / 100 です。統計学者は「P90 分位数レベル」と言う傾向がありますが、これ

は、「分位数 0.9」よりも言いやすいからです。例えば、P90 の分位レベルの予測では、実際の値は 90% の確率で予測値を下回ると予想されます。具体的には、時間 = t1 で分位数レベル = 0.9 のとき、予測値 = 30 で、これはつまり、時間 = t1 での実際の値は、1,000 回シミュレーションを行った場合、900 回のシミュレーションでは 30 未満になり、100 回のシミュレーションでは 30 回を超えると予想されます。

一方、P90 予測は、90% の確率で実際の値を上回るため、商品販売しないことによる機会コストが極端に高い場合や、投資資本のコストが低い場合に役立ちます。食料品店の場合、P90 予測は、牛乳やトイレットペーパーなどに使われる場合があります。このような食料品店の場合、売切れを回避することが重要であり、常に棚に商品が数点残っていてもかまいません。

p50 予測 (予測中央値とも呼ばれる) では、50% の確率で正しい値は予測値よりも低いと予想され、wQL [0.5] 指標を使用して精度を評価できます。在庫過剰がさほどの懸念事ではなく、特定の商品に対する需要が中程度の場合は、p50 分位数予測が役立ちます。

視覚化

Amazon Forecast を使用すると、AWS Management Console で予測をネイティブにプロットできます。さらに、Python のデータサイエンススタックをすべて活用することもできます ([Amazon Forecast Examples](#) (Amazon Forecast の例) を参照してください)。Amazon Forecast を使用すると、ExportForecastJob API を介して予測を CSV ファイルとしてエクスポートできます。これにより、ユーザーは好みの分析ツールで予測を視覚化できます。



Amazon Forecast コンソールでさまざまな分位数での視覚化が可能

予測ワークフローと API の概要

次の表は、予測ワークフローの各ステップと対応する Amazon Forecast API との照合表です。

表1: 予測ステップと Amazon Forecast API

ステップ	API	API の機能
ステップ 1: データを収集して集計する ステップ 2: データを準備する	Create_Dataset_Group、Create_Dataset、Create_Dataset_Import_Job	<ol style="list-style-type: none"> 1. 問題の高レベルのドメイン (小売、メトリクスなど) を決定します。 2. さまざまなデータセット (ターゲット、関連、商品のメタデータ) のスキーマを定義します。 3. Amazon S3 から Amazon Forecast にデータをインポートします。
ステップ 3: 予測子を作成する ステップ 4: 予測子を評価する	Create_Predictor	<ol style="list-style-type: none"> 1. ETL を実行します。 2. データをトレーニングセットとテストセットに分割して、モデルのトレーニングを行います。 3. オプションで、バックテスト結果を Create_predictor_backtest_export_job API を使ってバックテスト結果を CSV にエクスポートして、商品レベルのメトリクスを計算することもできます。
ステップ 5: 意思決定のための予測を生成して使用する	Create_Forecast	<ol style="list-style-type: none"> 1. モデルのトレーニングを実施してホストします。

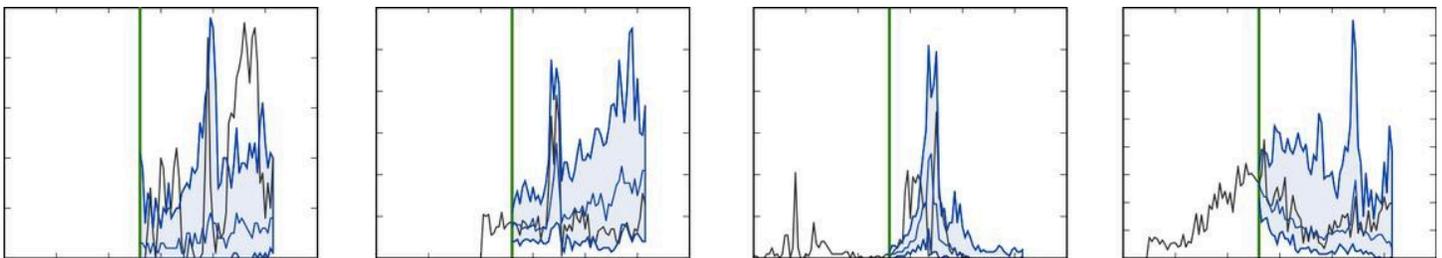
ステップ	API	API の機能
		2. 対象とする特定の分位数 (平均を含む 1 から 99 までの任意の整数) について、予測期間にわたる予測を生成します。
	Query_Forecast . Create_Forecast_Export_Job	Create_Forecast が作成した予測を使用できるようになります

一般的なシナリオで Amazon Forecast を使用する

外部変数 (価格やプロモーションなど) の変化に基づいてさまざまな予測を生成し、What-If 分析を行うこともできます。例えば、e コマースのケーススタディの例では、計画中のプロモーションに基づいてさまざまな予測を作成できます。ある商品の需要を 10% 割引で予測し、次に 20% 割引で需要を予測して、需要を満たすうえで在庫が必要な商品の数量を把握できます。これを実現するには、対象とするシナリオに基づいて、固有のデータセットグループを設定し、それぞれの関連する時系列を更新します。

さらに、これまでの履歴がない商品の予測を生成することもできます (これは、コールドスタートの問題と呼ばれることもあります)。このアプローチでは、DeepAR+ または CNN-QR とメタデータ (商品メタデータデータセットなど) を使用して予測子を作成し、新しい商品の予測を生成する必要があります。

次の図は、実際の運用上の予測の問題で起こる 4 つの異なる SKU の例を説明しています。



実際の運用上の予測の問題で起こる 4 つの異なる SKU の例

上の図では、黒色の実績値と比較して、観測実績値は縦線の左側、青色の予測値は縦線の右に表示されています。縦線の左側にある各 SKU の履歴は、緑の線の右側にある SKU の展開を示すものではないことに注意してください。

予測を本番環境に導入する

エンドツーエンドの Amazon Forecast ワークフロー完了後、`Create_Forecast` と `Create_Predictor` API の主な違いと、それぞれをいつ使用すべきかを特定することが非常に重要となります。

前者は主に概念実証の際にモデルの精度または指標を評価するために使用され、後者は本番稼働環境での予測を生成するために使用されます。

本番環境導入後、予測を生成するたびに毎回 `Create_Predictor` を実行する必要はありません。実行する必要があるのは、データの変更や事前設定した頻度 (隔週や月次など) の一環でモデルを再トレーニングする必要がある場合のみです。データセットが新しいデータで更新されるにつれ、`Create_Forecast` のみを実行すると、新しい予測期間の予測が生成されます。

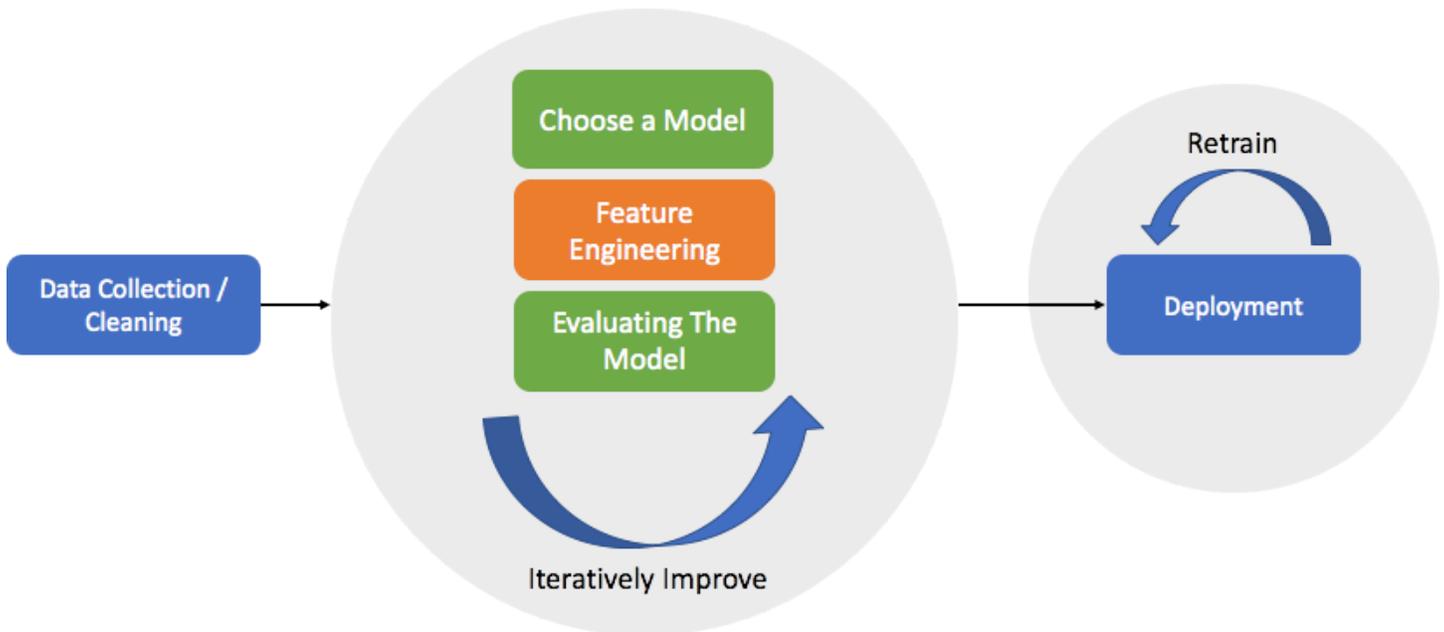
本番環境では、データセットのインポートと予測オペレーションを自動化して、新しい予測を定期的に生成する必要もあります。現在、これは、[Amazon CloudWatch Events](#) ログ、[AWS Step Functions](#)、[AWS Lambda](#) 機能を組み合わせた cron ジョブをセットアップすると実現できます。[cron](#) ジョブをセットアップすると、インポートおよび再トレーニングや予測生成のための Amazon Forecast API 呼び出しが自動化できます。最後に、サービスで規定された[システム制限](#)を超えないようにリソースを管理し、定期的に削除処理を行うことが重要です。[Amazon Redshift](#) などを含め、スケジュールジョブのセットアップの詳細については、[このブログ投稿](#)を参照してください。

まとめ

[Januschowski と Kolassa \(2019 年\)](#) は予測の問題を、企業が下す必要のある意思決定に沿って、戦略的、戦術的、運用上に分類しています。各意思決定レベルには、対応する予測タスクがあります。

運用上および戦術上の予測の問題は、大量のデータを含むという特徴があり、通常、高度な自動化が必要です。このような問題には、さまざまな予測方法を採用することができます。戦略的予測の問題には通常、局所的予測手法が、運用上の予測の問題には深層学習ベースの手法が適していますが、その中間の問題については、多少の実験が必要な場合があります。このホワイトペーパーでは、運用上の予測の問題について説明していますが、Amazon Forecast が提供するモデルについては一定の意見に縛られず、戦略的、運用上、戦術的な予測の問題に役立つモデルが幅広く提供されています。

運用上の予測の問題解決プロセスは、データの収集と準備からモデルの構築とデプロイまでの基本的なステップに分割できます。通常、これを順番に処理する線形のプロセスではなく、反復的なプロセスととらえることが、非常に役に立ちます。例えば、モデルやユースケースの理解が深まったところで、再度データ収集フェーズに戻るものが理にかなう場合も考えられます。モデルの開発自体も非常に反復的です。



予測モデルを本番環境にデプロイするための簡略化した開発プロセス。

寄稿者

本書の寄稿者は次のとおりです

- AI Vertical Services、シニア機械学習サイエンティスト、Yuyang Wang
- AI Vertical Services、アプライドサイエンティスト、Danielle Robinson
- 機械学習アプライドサイエンス、マネージャー、Tim Januschowski
- AI Vertical Services、シニアプロダクトマネージャー、Namita Das
- シニア AI/ML スペシャリストソリューションアーキテクト、Christy Bergman
- AI/ML ドキュメンテーション、テクニカルライター、Kris Tonthat

その他の資料

時系列予測と深層学習の方法の詳細については、以下を参照してください。

- [Amazon Forecast のドキュメント](#)
- [ブログ「Amazon Forecast が一般公開されました」](#)
- [Amazon SageMaker でのご利用開始: より正確な時系列予測のための DeepAR アルゴリズム](#)
- [Amazon SageMaker DeepAR が欠損値、カテゴリおよび時系列機能、一般化頻度をサポートするようになりました](#)
- [Amazon Forecast は畳み込みニューラルネットワーク \(CNN\) を使用して、予測モデルを最大 2 倍高速に、最大 30% 高い精度でトレーニング可能に](#)
- [Amazon Forecast で個別項目での精度測定が可能に](#)
- [Amazon Forecast を使用して、予測モデルの精度を測定し、ビジネス目標を最適化する](#)
- [Amazon Forecast Weather Index のリリース – 予測モデルの精度を高めるために、現地の天気を自動的に含める](#)
- [時系列予測モデルに関する科学論文](#)
- [Amazon Forecast samples GitHub ページ](#) (Amazon Forecast サンプルの GitHub ページ)
- [AWS アーキテクチャセンター](#)

付録 A: よくある質問

Q: Amazon Forecast の使用を開始するにはどうすればよいですか。

1. まず、AWS アカウント が必要です。
2. 次に、[AWS Management Console](#) で Forecast サービスを開いてデータセットグループを作成し、.csv ファイルを対象とする時系列データセットにインポートします (必須)。作業開始に最低限必要となるデータは、予測対象分の履歴データです (各世帯のタイムスタンプごとの電力量など)。
3. 最後に [CreatePredictor](#) を実行してモデルを作成し、[CreateForecast](#) を実行して予測結果を生成します。詳細については、「[Getting Started \(はじめに\)](#)」のドキュメントページを参照してください。

また、「[GitHub Introduction and Best Practice \(GitHub 入門とベストプラクティス\) ガイド](#)」も参照してください。

Q: Amazon Forecast は私に適したツールですか。

機械学習の問題のすべてが予測の問題に関するものであるわけではありません。最初に尋ねるべきは、「現在直面しているビジネス上の問題に時系列データが含まれているか」という質問です。例えば、将来の特定の日時における特定の値のみが必要なのでしょうか。時系列予測は通常、不正行為の検出やユーザーに対して映画作品をお勧めするなど、(特定の日時を問わない) 静的な問題には適していません。静的な問題には、より迅速に問題を解決できるソリューションがあります。

時系列データに加えて、データ自体の「密度」が高く、長期間の履歴がある必要があります。次の表にこれらの情報がまとめられています。

表 2 — 条件と Amazon Forecast アルゴリズムクラス

条件	Amazon Forecast アルゴリズムクラス
同様のパターンを持つ最大 500 万の時系列を含む大規模なデータセット + 季節的影響 + 関連データ。各時系列には長い期間の履歴が必要です。年間イベントをキャプチャする場合、理想的には 2 年以上の履歴、各時系列 300 以	Amazon Forecast 独自の深層学習である DeepAR+、CNN-QR

条件	Amazon Forecast アルゴリズムクラス
上、できれば最低 1,000 個のデータポイントが必要です。	
ほとんどの時系列に 300 個以上のデータポイントがあり、季節的影響と関連データを含む 1 ~ 100 の時系列で構成される小規模のデータセット。	Prophet
ほとんどの時系列に 300 個以上のデータポイントがあり、季節的影響を含む 1 ~ 数十の時系列で構成される小規模のデータセット。	ETS、ARIMA
時系列の大半に 300 個以上のデータポイントがあり、1 ~ 数十の時系列で構成される断続的な (多数の 0 を含む断続的な) データセット。	Amazon Forecast 独自の NPTS
ほとんどの時系列に 300 個未満のデータポイントがあり、1 ~ 数十の時系列で構成される小規模の (通常または断続的な) データセット。	このデータは Amazon Forecast には小規模すぎます。代わりに Excel の ETS や、従来の統計モデルである ARIMA や Prophet をお試しください。

初めてデータ作業を行う際には、Predictor の AutoML モードでトレーニングを行うことが、ベストプラクティスです。AutoML を使用すると、すべてのアルゴリズムが自動的に実行され (Deep Learning アルゴリズムはハイパーパラメータ調整 (HPO) がオンの状態で実行)、対象データに最適なアルゴリズムを判断できます。

Q: データの欠損についてはどう対応しますか。妥当な予測を生成するうえでデータの欠落が多すぎるのはどのような場合ですか。

データの記録に問題がある場合や、データの集計レベルが低すぎたり高すぎたりする場合があります。原則として、トレーニングデータの 1/3 を超えて予測の対象期間を設定することはできません。

欠損データの量以外にも、欠損データの代入も考慮すべきです。すべての 0 を NULL に変換して、欠損値を自動的に入力するという煩雑な作業は Amazon Forecast に任せることができます。Amazon Forecast を使用すると、欠損値の原因が新製品の導入 (コールドスタート) によるものか、製造中止の製品によるものかが自動的に検出されます。値、中央値、最小値、最大値、ゼロ、平

均、nan (ターゲット時系列のみ) など、いくつかの欠損値ロジックを使用できます。[NULL 入力構文の詳細については、ドキュメント](#)を参照してください。

- 「frontfill」 — (TTS のみ) 新商品やコールドスタートする商品と、その商品の履歴が発生する前の null 処理方法を指します
- 「middlefill」 — 時系列値の途中の null を指します
- 「backfill」 — 販売終了商品と、商品の販売が停止した後の null の処理方法を指します。
- 「futurefill」 — (RTS のみ) トレーニングデータ終了後に発生する null を指します

Q: 入力履歴データには負の値はないのに、需要予測には負の値が表示されています。どのような理由でこのような予測になりますか。これを回避するにはどう対処できますか。

NPTS (負の数値を含まないデータを使用したトレーニング) モデルと DeepAR (負の 2 項確率関数を使用) モデル以外のすべてのモデルを使用すると、負の数値が生成されない保証はありません。上記のモデルのいずれかに変更するか、予測値の負の値を切り捨てることが解決策です。

Q: 精度指標が分位数で異なるのはなぜですか。モデルが同じならば、同じエラーになるはずではないのですか。

加重が分位数に依存する方法の詳細については、[重み付き分位損失 \(wQL\)](#) を参照してください。

例えば、すべての予測が p10、p50、p90 の 3 つの異なる分位数にある場合、この 3 つの予測自体は確率変数です。精度は、各分位レベルの実績値と予測値間で個別に算出されます。以下のような重み付き分位損失の「wQL」の表が表示される場合があります。この wQL 値は相互に決定的な関係にあるわけではありません。(リコール損失はエラーを意味するため、非順序です。ただし、分位数の予測は順序付けられています)。したがって、例えば、p90 の wQL が p50 の wQL 以上であるべき根拠はありません。

表 3 — 予測分位数の例

	A	B	C
1	P10 wQL	P50 wQL	P90 wQL
2	0.18647	0.50879	0.30428

Q: 予測の精度を向上するにはどうすればよいですか。

予測の精度は、適切なデータが適切な量と質で入手できるかどうか依存します。精度が不十分な場合は、問題の予測可能性の程度 (またはデータのランダム性、ノイズの有無、定常性のレベル) を把握することが、適切である場合があります。考慮すべきその他の要素には、さまざまなモデルの評価、ハイパーパラメータの設定、関連する時系列と商品のメタデータデータセットを使用した追加機能の組み込みなどがあります。具体的な提案については、[GitHub に掲載されているこのベストプラクティスドキュメントを参照してください](#)。

Q: 取り組んでいるユースケースに非常に適したアルゴリズムがあっても、Amazon Forecast では提供されていません。どうすればよいですか。

Amazon Forecast チームが該当のユースケースについてのサポートを提供します。Amazon Forecast<amazonforecast-poc@amazon.com> サービスチームにメールでお問い合わせください。

付録 B: 参考文献

[Januschowski, Tim, Kolassa, Stephan 「A Classification of Business Forecasting Problems」 Foresight: 『The International Journal of Applied Forecasting』 2019 年](#)

[Salinas, David, Flunkert, Valentin, Gasthaus, Jan, Januschowski, Tim 「DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks」 『International Journal of Forecasting』 2019 年](#)

[Gasthaus, Jan, Benidis, Konstantinos, Wang, Yuyang, Rangapuram, Syama Sundar, Salinas, David, Flunkert, Valentin, Januschowski, Tim 「Probabilistic Forecasting with Spline Quantile Function RNNs」 第 22 回人工知能と統計に関する国際会議2019 年](#)

[Januschowski, Tim, Gasthaus, Jan, Wang, Yuyang, Salinas, David, Flunkert, Valentin, Bohlke-Schneider, Michael, Callot, Laurent 「Criteria for classifying forecasting methods」 『International Journal of Forecasting』 2019 年 \(サインインが必要\)](#)

[Januschowski, Tim, Gasthaus, Jan, Wang, Yuyang, Rangapuram, Syama, Callot, Laurent 「Deep Learning for Forecasting」 Foresight: 『The International Journal of Applied Forecasting』 2018 年](#)

[Januschowski, Tim, Gasthaus, Jan, Wang, Yuyang, Rangapuram, Syama Sundar, Callot, Laurent 「Deep Learning for Forecasting: Current Trends and Challenges」 Foresight: 『The International Journal of Applied Forecasting』 2018 年](#)

[Bose, Joos-Hendrik, Flunkert, Valentin, Gasthaus, Jan, Januschowski, Tim, Lange, Dustin, Salinas, David, Schelter, Sebastian, Seeger, Matthias, Wang, Yuyang 「Probabilistic Demand Forecasting at Scale」 VLDB Endowment の議事録。2017 年](#)

ドキュメント履歴

このホワイトペーパーの更新に関する通知を受け取るには、RSS フィードを購読してください。

変更	説明	日付
ホワイトペーパーの更新	更新	2021 年 9 月 1 日
初版公開	ホワイトペーパーの初版発行。	2020 年 2 月 4 日

Note

RSS の最新情報にサブスクライブするには、ご使用のブラウザで RSS プラグインを有効にする必要があります。

注意

お客様は、この文書に記載されている情報を独自に評価する責任を負うものとし、本書は、(a) 情報提供のみを目的とし、(b) AWS の現行製品と慣行について説明しており、これらは予告なしに変更されることがあり、(c) AWS およびその関連会社、サプライヤーまたはライセンサーからの契約上の義務や保証をもたらすものではありません。AWS の製品やサービスは、明示または暗示を問わず、一切の保証、表明、条件なしに「現状のまま」提供されます。お客様に対する AWS の責任は、AWS 契約により規定されます。本書は、AWS とお客様の間で締結されるいかなる契約の一部でもなく、その内容を修正するものでもありません。

© 2021 Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS の用語集

最新の AWS の用語については、AWS 全般のリファレンスの「[AWS の用語集](#)」を参照してください。