

AWS 白皮書

Amazon Forecast 的時間序列預測原則



Amazon Forecast 的時間序列預測原則: AWS 白皮書

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

Amazon 的商標和商業外觀不得用於任何非 Amazon 的產品或服務，或以任何可能造成客戶混淆、以任何貶低或損毀 Amazon 的方式使用 Amazon 的商標和商業外觀。所有其他非 Amazon 擁有的商標均為其各自擁有者的財產，這些擁有者可能隸屬於 Amazon，或與 Amazon 有合作關係，亦或受到 Amazon 贊助。

Table of Contents

摘要和概觀	i
概觀	1
採用 Well-Architected 了嗎？	1
關於預測	3
預測系統	3
預測問題發生於何處？	3
嘗試解決預測問題前的考量事項	4
案例研究：電子商務企業的零售需求預測問題	5
步驟 1：收集和彙總資料	7
範例	8
步驟 2：準備資料	10
如何處理遺漏的資料	10
範例 1	10
範例 2	12
特徵化和相關時間序列的概念	12
範例 3	13
步驟 3：建立預測器	15
步驟 4：評估預測器	17
回測	17
預測分位數和準確性指標	18
加權分位數損失 (wQL)	19
加權絕對百分比誤差 (WAPE)	19
均方根誤差 (RMSE)	20
WAPE 和 RMSE 的相關問題	20
步驟 5：產生並使用預測進行決策	22
機率預測	22
視覺效果	23
預測工作流程和 API 摘要	24
在常見案例中使用 Amazon Forecast	25
將預測導入生產中	25
結論	26
作者群	27
深入閱讀	28
附錄 A：常見問答集	29

附錄 B：參考資料	32
文件歷史記錄	33
聲明	34
AWS 詞彙表	35

Amazon Forecast 的時間序列預測原則

發佈日期：2021 年 9 月 1 日 ([文件歷史記錄](#))

如今，公司使用簡單的試算表到複雜的財務規劃軟體等各種工具，以期能準確預測未來的業務成果，例如產品需求、資源需求和財務績效。本白皮書介紹預測及其術語、挑戰和使用案例。本文件使用案例研究來補強預測概念、預測步驟，並述及 [Amazon Forecast](#) 為何有助於解決實際預測問題的許多實質挑戰。

概觀

預測是研判未來趨勢的科學。利用歷史資料，企業可以了解趨勢、判斷未來可能發生的事情及其時間，繼而將這些資訊擬定在未來的計劃中，從產品需求到庫存規劃與人員配置等，一切都做好準備。

從預測的結果來看，準確性至關重要。如果預測過高，客戶可能會過度投資於產品和員工，從而導致無效投資。如果預測太低，客戶可能會投資不足，導致原料和庫存短缺，進而產生糟糕的客戶體驗。

現今的企業為了產生預測均廣泛使用各種工具，從簡單的試算表到複雜的需求/財務規劃軟體都在其列，但仍難以掌握高準確性，原因有兩個：

- 首先，傳統預測難以納入大量的歷史資料，過去有許多重要訊息被吞沒在雜訊之中。
- 其次，傳統預測很少納入相關但獨立的資料，而這類資料可能提供了重要情境背景 (例如價格、假日/活動、缺貨、市場促銷等等)。若沒有完整的歷史和較豐富的情境背景，大多數預測都無法準確預測未來。

[Amazon Forecast](#) 是可克服這些問題的全受管服務。Amazon Forecast 為現有的預測情境提供了最佳演算法。它會適時採用現代機器學習 (ML) 和深度學習，以提供高準確度的預測。Amazon Forecast 易於使用，且不需要機器學習經驗。此服務會自動提供必要的基礎架構、處理資料，並建置由 AWS 託管的自訂/私有 ML 模型，供預測之用。此外，隨著機器學習技術不斷快速發展，Amazon Forecast 也整合了這些技術，讓客戶只需進行最低限度的額外工作，就能享有準確性的提升。

採用 Well-Architected 了嗎？

[AWS Well-Architected Framework](#) 可協助您了解在雲端建立系統時所做決策的利弊。該架構的六根支柱讓您能夠學習設計和操作可靠、安全、高效、經濟高效且永續的系統的架構最佳實務。使用 [AWS Well-Architected Tool](#) (在 [AWS Management Console](#) 中免費提供)，您可以透過回答每根支柱的一組問題來針對這些最佳實務審查您的工作負載。

在[機器學習聚焦](#)中，我們專注於如何在 AWS 雲端 中設計、部署和架構機器學習工作負載。本聚焦會新增至 Well-Architected 架構中說明的最佳實務。

如需更多關於雲端架構的專家指導和最佳實務 (參考架構部署、圖表和白皮書)，請參閱[AWS架構中心](#)。

關於預測

在本文件中，預測是指預測時間序列的未來值：問題的輸入或輸出具有時間序列的屬性。

預測系統

預測系統包含各式各樣的使用者：

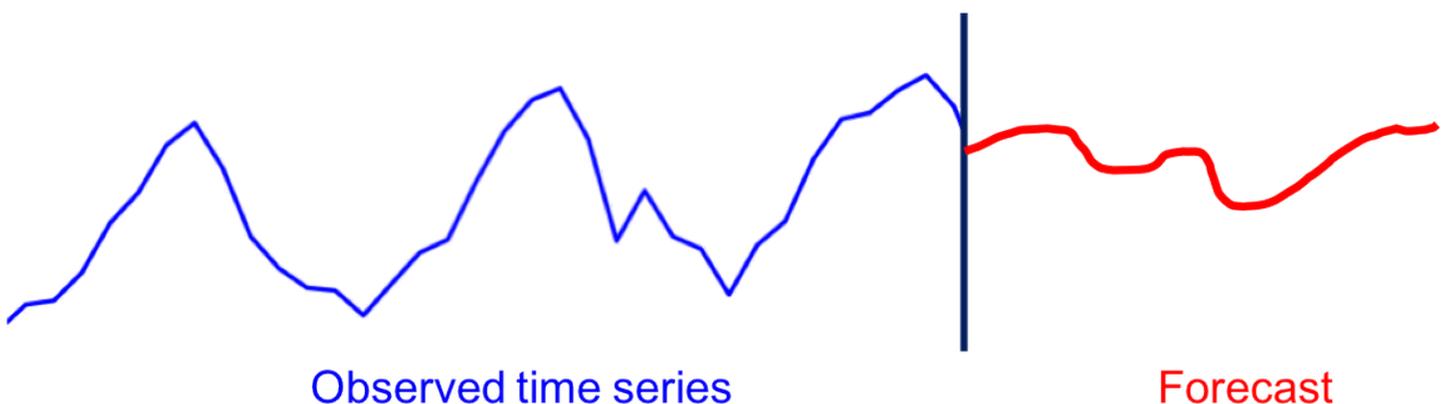
- 最終使用者，會查詢特定產品的預測，並決定要購買多少單位；有可能是個人或自動化系統。
- 業務分析師/商業智慧，負責支援最終使用者，執行和組織彙總報告。
- 資料科學家，會反覆分析需求模式、因果關係，並添加新特徵，以逐漸改良模型或改進預測模型。
- 工程師，負責設定資料收集的基礎架構，並確保系統的輸入資料可供使用。

Amazon Forecast 可減輕軟體工程師的工作負擔，並且讓資料科學能力有限的企業得以運用最先進的預測技術。對於具有資料科學能力的企業，Amazon Forecast 包含許多診斷功能，因此能夠妥善解決預測問題。

預測問題發生於何處？

在許多會自然產生時間序列資料的領域中，都會發生預測問題。其中包括零售銷售、醫學分析、容量規劃、感應器網路監控、財務分析、社交活動探勘和資料庫系統。例如，在具備資料驅動決策能力的多數企業中，預測在自動化和最佳化營運程序方面扮演著重要角色。產品供需的預測可用於最佳庫存管理、員工排班和拓撲規劃，且一般來說，對絕大多數的供應鏈最佳化環節而言都是關鍵技術。

下圖包含預測問題的摘要，以顯現模式的觀測時間序列為基礎 (在此範例中為季節性)，對指定的期間建立預測。水平軸代表從過去 (左) 到未來 (右) 的時間。垂直軸代表測量單位。基於過去 (藍色線段) 到垂直黑線識別未來 (紅色線段)，是預測任務。



預測任務概觀

嘗試解決預測問題前的考量事項

在解決預測問題之前最需要了解的重要問題是：

- 您需要解決預測問題嗎？
- 你為何要解決預測問題？

由於時間序列資料無所不在，我們隨處都可輕易發現預測問題。然而，關鍵問題在於是否真的需要解決預測問題，或者，您是否大可規避這類問題且不犧牲有效的業務決策。提出此問題是很重要的，因為就科學上而言，預測是機器學習中最困難的問題之一。

以線上零售商的產品推薦為例。我們可將這類產品推薦問題設想為預測問題，其中，對於每個客戶/存貨單位 (SKU) 配對，您要預測該名客戶將購買之特定品項的單位數量。此問題表述有許多好處。其中一項好處是明確將時間要素納入考量，因此您可以根據客戶的購買模式推薦產品。

但是，產品推薦問題很少被制定為預測問題，因為解決此類預測問題要比直接解決建議問題困難得多 (例如，因為客戶 SKU 層級的資訊稀少，問題的規模較大)。因此，在思考預測應用程式時，請務必考慮預測的下游使用，以及是否可使用替代方法來解決此問題。

在這些情況下，[Amazon Personalize](#) 將有所幫助。Amazon Personalize 是一款機器學習服務，能讓開發人員輕鬆利用他們的應用程式為客戶提供個人化建議。

確認您需要解決預測問題後，下一個浮現的問題是，為何要解決預測問題？在許多企業的設想下，預測通常只是為了達到目的的手段。以零售業的需求預測為例，預測可用來進行庫存管理決策。預測問題通常會輸入到決策問題中，繼而成為最佳化問題模型的一部分。

舉例來說，這種決策問題包括要購買的單位數量或處理現有庫存的最佳方法。其他商業預測問題包括預測伺服器容量或預測製造業的原料/零件需求。這些預測可作為其他程序的輸入 (無論是用於上述決策問題，還是用於案例模擬)，繼而用於沒有明確模型的規劃。所謂的預測本身並非目的，是有例外情況的。例如，在財務預測中，預測會直接用來建立財務儲備或提交給投資者。

若要了解預測的目的，請考量下列問題：

- 您應預測多久之後的未來？
- 您多久要產生一次預測？
- 您是否應深入研究預測的某些具體層面？

案例研究：電子商務企業的零售需求預測問題

為了更詳細地說明預測概念，請考量電子商務企業線上銷售產品的案例。最佳化供應鏈方面的決策 (例如庫存管理) 對於此企業的核心競爭力至關重要，因為這有助於在適當的履約地點備妥準確數量的產品。基本上，這意味著更短的出貨時間和價格優勢方面的選擇餘裕，從而造就更高的客戶滿意度。供應鏈軟體系統的主要輸入，是型錄中每項產品的需求預測或潛在銷售量預測。此預測可運用在重要的下游決策，主要包括：

- 宏觀層面規劃 (策略性預測)：就企業整體而言，總銷售額/收入方面的預計增長為何？企業應在哪些區域展現積極性或更為積極？人力應如何配置？
- 需求 (或庫存) 預測：每項產品在每個地點的預期銷售數量為何？
- 促銷活動 (戰術性預測)：促銷活動應如何進行？產品是否應清倉減價？

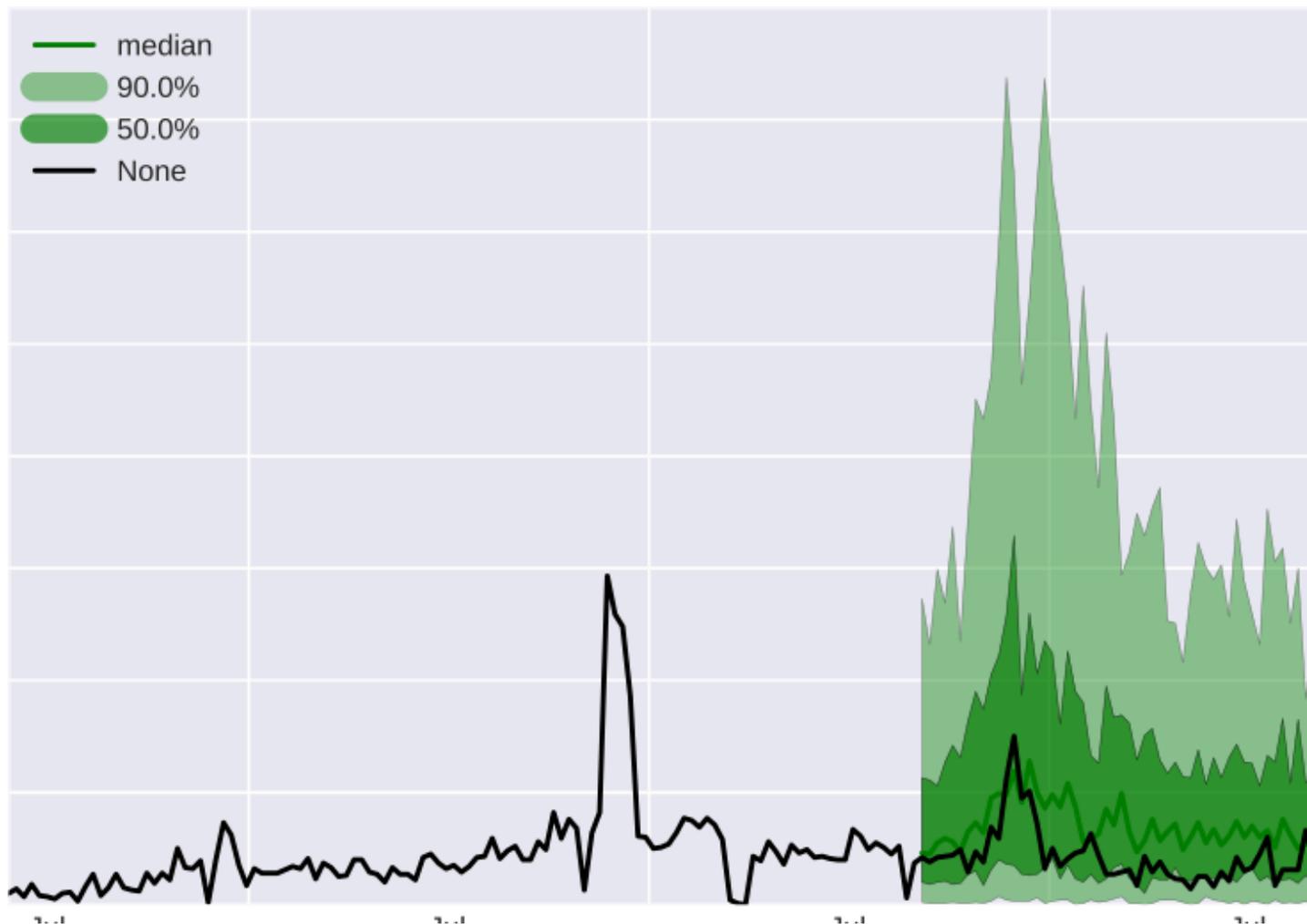
案例研究的其餘部分側重於第二個問題，這是營運預測問題 (Januschowski 和 Kolassa, 2019) 的一部分。本文件以主要考量為基調：資料、模型 (預測器)、推論 (預測) 和生產化。

對於此案例研究請牢記，預測問題只是達到目的的手段。儘管預測對企業而言至關重要，但更重要的是下游供應鏈決策。在我們的案例研究中，這些決定是由自動化購買系統進行的，而這些系統採用來自運營研究的數學最佳化模型。這些系統會嘗試將企業的預期成本降到最低。

關鍵字是預期，這意味著預測不僅應涵蓋一個可能的未來，還應涵蓋所有可能的未來，並根據特定結果的可能性使用適當的加權。為此，下游決策的關鍵推動因子將是預測值的完整分佈，而非僅止於點預測。下圖顯示機率預測 (也稱為密度預測)。請注意，您可以從這個機率預測中輕鬆導出單點預測 (最有可能的未來)，但是從點預測導出機率預測將會困難得多。

有了機率預測，您即可從中獲得不同的統計資料，並適當自訂結果以利做出您想要的決策。電子商務企業可能會有多項他們絕不想缺貨的主要產品。在此案例中，我們使用高分位數 (例如，第 90 個百分位數)，其實質意義為產品有 90% 的時間有庫存。對於其他產品，例如較容易找到替代品的產品 (如鉛筆)，使用較低的百分位數可能較合適。

在 Amazon Forecast 中，您可以機率預測輕鬆獲得不同的分位數。



機率預測圖例

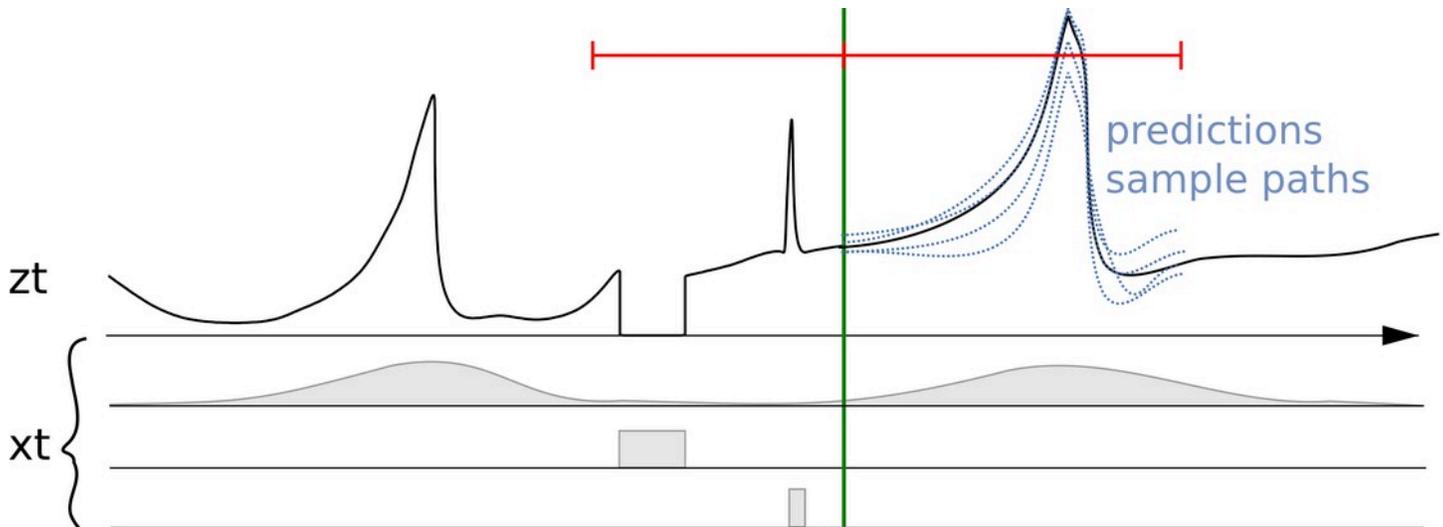
在上圖中，黑線是實際值；深綠色線是預測分佈的中位數；深綠色陰影區域是您預期 50% 的值所將隸屬的預測區間；淺綠色區域是您預期 90% 的實際值所將隸屬的預測區間。

以下幾節說明解決此企業的預測問題所涉及的步驟，包括：

- [資料收集與彙總 \(步驟 1\)](#)
- [資料準備 \(步驟 2\)](#)
- [建立預測器 \(步驟 3\)](#)
- [評估預測器 \(步驟 4\)](#)
- [自動產生預測 \(步驟 5\)](#)

步驟 1：收集和彙總資料

下圖顯示預測問題的心理模型。目標是預測時間序列 z_t 的未來，並盡可能使用較多的相關資訊，使預測盡可能準確。因此，首要步驟是盡可能收集較多的正確資料。



時間序列 z_t 以及相關聯的特徵或共變數 (x_t) 和多個預測

在上圖中，垂直線右側顯示了多個預測。這些預測是機率預測分佈的樣本 (或者，可反過來表示機率預測)。

零售企業應記錄的主要資訊是：

- 交易銷售資料 — 例如，存貨單位 (SKU)、地點、時間戳記和銷售數量。
- SKU 品項詳細資料 — 品項的中繼資料。範例包括顏色、部門、大小等等。
- 價格資料 — 附有時間戳記的各個品項的價格時間序列。
- 促銷資訊資料 — 不同類型的促銷活動，可能是針對一系列的品項 (類別) 或附有時間戳記的個別品項。
- 庫存資訊資料 — 在每個時間單位，指出某個 SKU 是有庫存/可購買還是缺貨的資訊。
- 位置資料 — 品項或銷售在給定時間點的所在位置可以表示為字串 `location_id` 或 `store_id`，或是表示為實際地理位置。地理位置可以是國家/地區代碼加上五位數的郵遞區號或 `latitude_longitude` 座標。位置被視為交易銷售的一個「維度」。

在 [Amazon Forecast](#) 中，要預測之數量的歷史資料稱為目標時間序列 (TTS)。就零售企業而言，TTS 是交易銷售資料。其他歷史資料 (與每筆銷售交易同時得知) 稱為相關時間序列 (RTS)。零售企業的 RTS 會包含價格、促銷和庫存變數。

請注意，庫存資訊很重要，因為此問題的重點在於預測需求而不是銷售量，但企業僅記錄銷售量。某 SKU 缺貨時，銷售量會低於潛在需求，因此了解並記錄這類缺貨事件何時發生，是非常重要的。

其他需考量的資料集包括網頁造訪次數、搜尋詞彙的詳細資料、社交媒體和天氣資訊。資料在過去和未來都可供使用往往是很重要的，如此才能在模型中使用該資料。這是許多預測模型和回測的要求 (說明請見 [步驟 4：評估預測](#) 一節)。

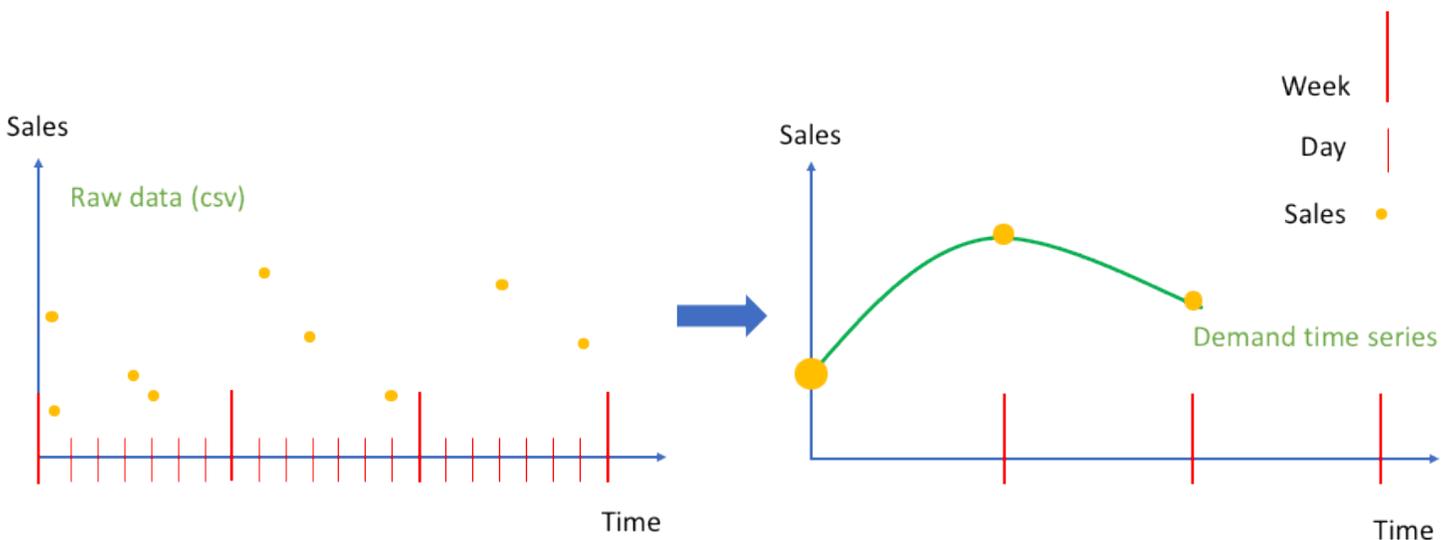
就某些預測問題而言，原始資料的頻率自然會與預測問題的頻率相符。其範例包括伺服器磁碟區的要求，當您要以分鐘為頻率進行預測時，就會按分鐘對磁碟區取樣。

資料常會以更精細的頻率記錄，或就是在時間範圍內以任意時間戳記記錄，但預測問題的精細程度較粗糙。這是零售案例研究中的常見情況，銷售資料通常會記錄為交易資料；例如，格式中包含以高精細程度指出銷售發生時間的時間戳記。在預測使用案例中可能不需要這種低精細程度，將此資料彙總到每小時或每日銷售額可能會較為適當。此處的彙總層級對應於下游問題；例如，庫存管理或資源規劃。

範例

在下圖中，左圖顯示可在 Amazon Forecast 中輸入為逗號分隔值 (CSV) 檔案的原始客戶銷售資料的範例。在此範例中，銷售資料定義於較精細的每日時間格線上，而問題是要以較粗略的時間格線預測未來的每週需求。Amazon Forecast 會在 `create_predictor` API 呼叫中執行給定週內每日值的彙總。

結果會將原始資料轉換成格式良好、具有固定每週頻率的時間序列集合。右圖說明使用預設加總和彙總方法在目標時間序列上進行此彙總的情形。其他彙總方法包括平均值、最大值、最小值或選擇單一點 (例如，第一點)。彙總精細程度和方法必須適當選擇，使其最為符合資料的商業使用案例所需。在此範例中，彙總值會與每週彙總相符。使用者可以使用 `create_predictor` API `FeaturizationConfig` 參數的 `FeaturizationMethodParameters` 索引鍵來設定其他彙總方法。



將原始銷售資料以事件的形式 (左) 彙總為等間距時間序列 (右)

步驟 2：準備資料

有了可用的原始資料後，您必須處理複雜性 (例如遺漏的資料)，並確定您為預測模型準備的資料能夠確切擷取預期的解釋。

如何處理遺漏的資料

在實際的預測問題中，常見的情況是原始資料中有遺漏值存在。時間序列中有遺漏值，表示具有指定頻率的每個時間點的真實對應值無法供進一步處理之用。值被標示為遺漏可能有多種原因。

發生遺漏值的原因可能是沒有交易，或可能的測量錯誤 (例如，因為監視特定資料的服務未正常運作，或因為測量無法正確進行)。在零售案例研究中，第二種原因的常見範例是需求預測中的缺貨狀況，這表示需求不等於當天的銷售額。

當服務達到限制時 (例如，特定 [AWS 區域](#) 中的 [Amazon EC2](#) 執行個體都在忙碌中)，雲端運算案例中也會發生類似的效應。另一個遺漏值的範例，是產品或服務尚未推出或停止生產時。

遺漏值也可以由特徵處理元件插入，以透過填補確保時間序列的長度相等。遺漏值若遍佈到某種程度，可能會嚴重影響模型的準確性。

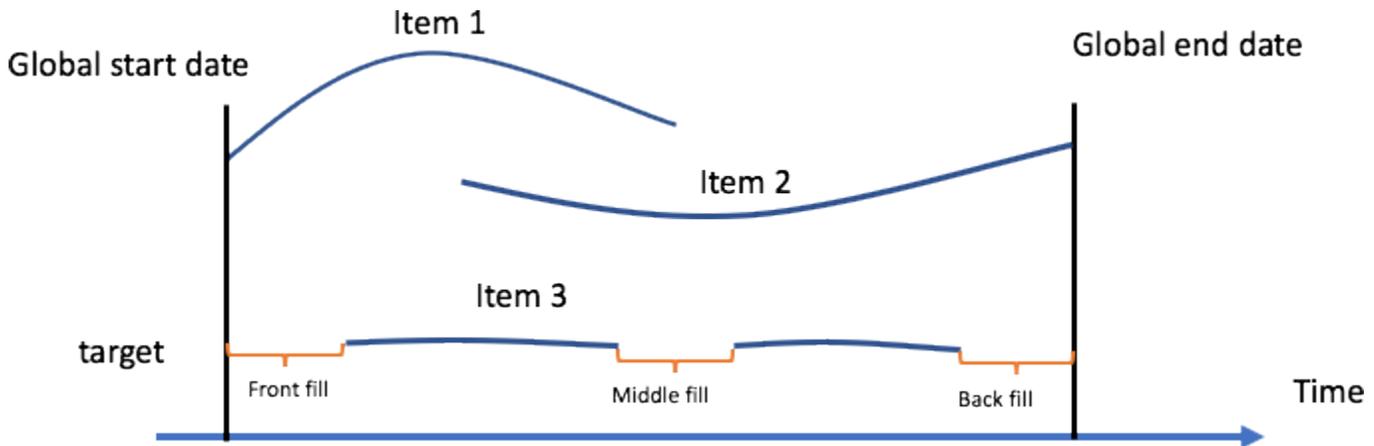
範例 1

填入是將標準化值新增至資料集中的遺失項目的程序。下圖針對由三個品項組成的資料集中的品項 2，說明了在 Amazon Forecast 中處理遺漏值的不同策略 — 前端、中間、往回和未來填入。

Amazon Forecast 對於目標和相關時間序列均支援填入。全域開始日期定義為您的資料集中所有品項的開始日期之中最早的開始日期。在下列範例中，全域開始日期是品項 1 的開始日期。同樣地，全域結束日期定義為所有品項的時間序列最晚的結束日期，即品項 2 的結束日期。

前端填入會填入從特定時間序列的開頭到全域開始日期之間的每個值。截至本文件發佈時，Amazon Forecast 並未開啟任何前端填入，並允許所有時間序列從不同的時間點開始。中間填入表示在時間序列中間 (例如，在品項的開始日期和結束日期之間) 填入的值，往回填入則會在該時間序列的最後一個日期到全域結束日期之間填入。

對於目標時間序列，中間和往回填入方法的預設填入邏輯為零。未來填入 (僅適用於相關時間序列) 會填入品項的全域結束日期與客戶指定的預測期間之間的任何遺漏值。將相關時間序列資料集與 [Prophet](#) 和 [DeepAR+](#) 一起使用時需要未來值，對 [CNN-QR](#) 而言則是選用的。



Amazon Forecast 中的遺漏價值處理策略

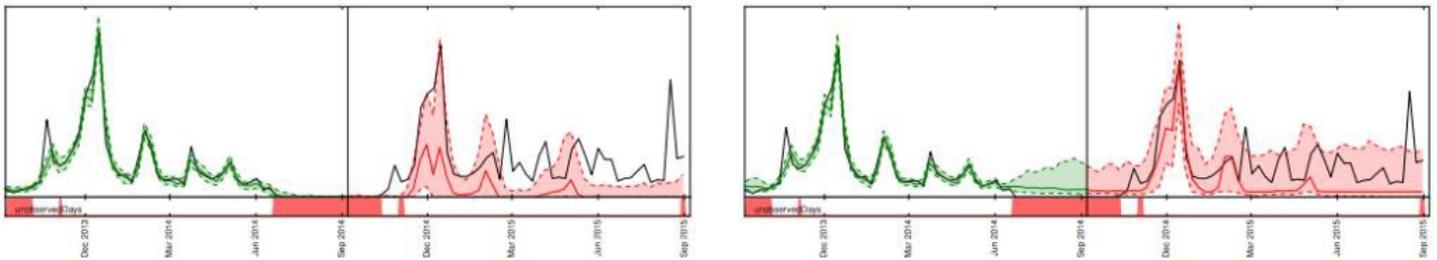
在上圖中，全域開始日期表示所有品項的開始日期之中最早的開始日期，而全域結束日期則表示所有品項的結束日期之中最晚的結束日期。預測期間是 Forecast 就目標值提供預測的期間。

可供貨品項的交易資料顯示銷售額為零，是零售研究中常見的情況。這些值會被視為真的零，並且用於指標評估元件中。Amazon Forecast 可讓使用者識別實際上遺漏的值，並將其編碼為非數字 (NaN) 以供演算法處理。本文接下來將探討為何這兩種案例有所不同，及其各自的適用時機。

在零售案例研究中，零售商銷售了零個可供貨品項的資訊，會與銷售了零個不可供貨品項的資訊不同；這可分成在該品項的非上市期間 (例如，在該品項推出前或停賣後) 或在其上市期間 (例如，部分缺貨，或沒有此時間範圍內的銷售資料記錄) 兩種情況。預設的零填入適用於前述的第一個案例。在第二個案例中，即使對應的目標值通常為零，但在標示為遺漏的值中還會傳達其他資訊。最佳實務是保留有遺漏資料的資訊，而不要捨棄此資訊。請參閱下列範例，了解保留該資訊有何重要性。

Amazon Forecast 支援值、平均值、中位數、最小值和最大值等其他填入邏輯。對於相關時間序列 (例如價格或促銷)，並沒有針對中間、往回或未來填入方法指定的預設值，因為正確的遺漏值邏輯會隨著屬性類型與使用案例而不同。支援相關時間序列的填入邏輯包括零、值、平均值、中位數、最小值和最大值。

若要執行遺漏值填入，請指定您在呼叫 [CreatePredictor](#) 操作時要實作的填入類型。填入邏輯指定於 [FeaturizationMethod](#) 物件中。例如，若某值不代表目標時間序列中不可供貨產品的銷售額為零，在對該值進行編碼時，請將填入類型設定為 NaN，以將值標示為真正的遺漏值。與零填入不同，使用 NaN 編碼的值會被視為真正的遺漏值，而不會用於指標評估元件中。



對同一品項進行預測時使用 0 填入與 NaN 填入的效果

在上方的左圖中，垂直黑線左側的值填入 0，產生了低估偏差的預測 (垂直黑線右側)。在右圖中，這些值標示為 NaN，而產生了適當的預測。

範例 2

上圖說明對線性狀態空間模型 (例如 [ARIMA](#) 或 [ETS](#)) 正確處理遺漏值的重要性。其中繪製了部分缺貨之品項的需求預測。訓練區域以綠色顯示在左圖中，預測範圍在右側面板中以紅色顯示，真正的目標顯示為黑色。中位數、p10 和 p90 預測分別顯示為紅線和陰影區域。底部顯示以紅色標示的缺貨商品 (80% 的資料)。在左圖中，缺貨區域遭到忽略，並以 0 填入。

這導致預測模型假設有很多需要預測，因而致使預測過低。在右圖中，缺貨區域被視為真正的遺漏觀察值，缺貨區域中的需求變成未定值。將缺貨品項的遺漏值適當標示為 NaN 後，您會在此圖中看到預測範圍內沒有低估偏差。Amazon Forecast 填補了資料中的這些缺漏，讓您可以輕鬆地正確處理遺漏的資料，而無須明確修改所有的輸入資料。

特徵化和相關時間序列的概念

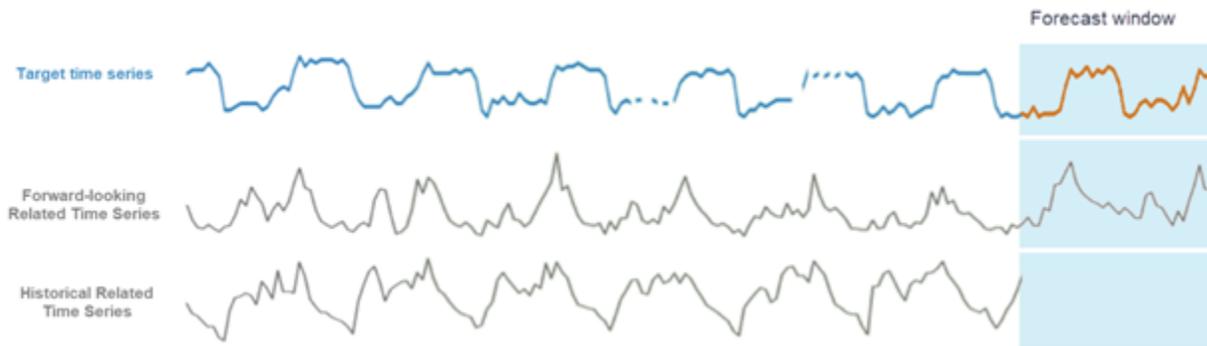
Amazon Forecast 可讓使用者輸入相關資料，以利提高特定支援預測模型的準確性。這項資料有兩種類型：相關時間序列或靜態項目中繼資料。

Note

中繼資料和相關資料在機器學習中稱為特徵，在統計學中則稱為共變數。

相關時間序列是與目標值有一定關聯的時間序列，應提供某種程度的統計強度來預測目標值，因為它們以直觀的詞彙提供解釋 (如需範例，請參閱 [Amazon Forecast：大規模預測時間序列](#))。與目標時間序列不同，相關時間序列是過去可能對目標時間序列產生影響的已知值，且未來可能會有已知值。

在 Amazon Forecast 中，您可以新增兩種類型的相關時間序列：歷史時間序列和前瞻性時間序列。歷史相關時間序列包含預測期間之前的資料點，而不包含未來預測期間內的任何資料點。前瞻性相關時間序列包含預測期間之前和期間內的資料點。



透過 Amazon Forecast 使用相關時間序列的不同方法

範例 3

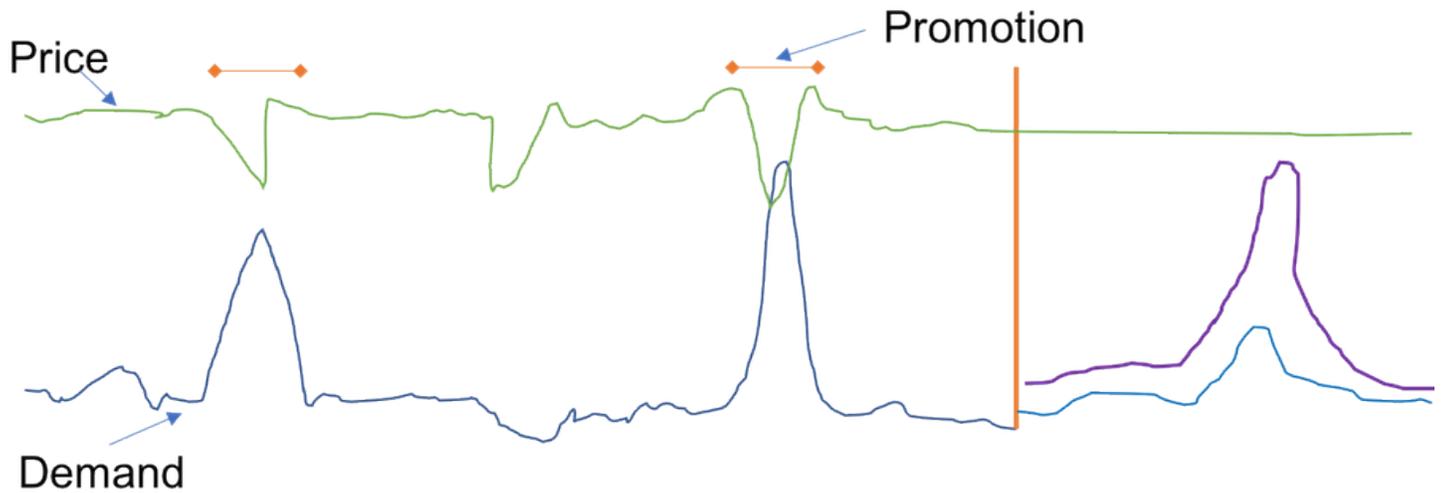
下圖顯示如何使用相關時間序列來預測暢銷書未來需求的範例。藍線代表目標時間序列中的需求。價格顯示為綠線。垂直線代表預測的開始日期，垂直線右側顯示兩個分位數的預測。

此範例使用前瞻性相關時間序列，此序列與目標時間序列的預測精細程度保持一致，且未來在預測開始日期到預測期間遞增的預測開始日期 (預測結束日期) 之間的範圍內，全程 (或多數時間) 都是已知的。

下圖也顯示價格是合用的特徵，因為您可以看到價格下降與產品銷售量增加之間的相關性。相關時間序列可透過個別的 CSV 檔案提供給 Amazon Forecast，其中包含品項 SKU、時間戳記和相關時間序列值 (在此案例中為價格)。

Amazon Forecast 支援彙總方法 (例如目標時間序列的平均值和總和)，但不支援相關時間序列。例如，將每日價格與每週價格相加沒有什麼意義，每日促銷也是如此。

Amazon Forecast 可納入內建的特徵化資料集，將[天氣](#)和[假日](#)資訊自動整合到模型中 (請參閱 [SupplementaryFeature](#))。天氣資訊和假日對零售需求可能有很大的影響。



特定商品的銷售量 (在垂直紅線左側顯示為藍色)

項目中繼資料 (也稱為類別變數) 是可輸入至 Amazon Forecast 的其他有用特徵 (如需範例, 請參閱 [Amazon Forecast : 大規模預測時間序列](#))。類別變數與相關時間序列的主要區別在於, 類別變數是靜態的, 不會隨著時間而變更。常見的零售例子包括品項的顏色、書籍的類別, 以及電視是否為智慧電視的二元指標。假設類似的 SKU 會有類似的銷售量, 您可以透過深度學習演算法來了解存貨單位 (SKU) 之間的相似性, 從而取得這項資訊。由於此中繼資料沒有時間相依性, 項目中繼資料 CSV 檔案中的每一列僅包含品項 SKU 和對應的類別標籤或說明。

步驟 3：建立預測器

有兩種方式可以建立預測器：執行 [AutoML](#)，或手動從六種內建 Amazon Forecast 演算法中擇一使用。截至撰寫本文件為止，在執行 AutoML 時，Amazon Forecast 會自動測試這六種內建演算法，並選擇第 10 個、第 50 個 (中位數) 和第 90 個分位數的平均分位數損失最低的演算法。

Amazon Forecast 提供四種本機模型：

- [整合移動平均自回歸模型 \(ARIMA\)](#)
- [指數平滑法 \(ETS\)](#)
- [非參數時間序列 \(NPTS\)](#)
- [Prophet](#)

本機模型是一種預測方法，會對每個個別的時間序列 (或特定品項/維度組合) 採用單一模型，然後使用該模型來推算未來的時間序列。

ARIMA 和 ETS 是 R 預測套件中常用本機模型的可擴展版本。NPTS 是 Amazon 開發的一種本機方法，與其他本機模型有顯著的差異。與藉由重複最後一個值或適當季節性的值來提供點預測的簡單季節性預測器不同，NPTS 會產生機率預測。NPTS 使用固定時間指數，其中，上一個指數 ($T - 1$) 或上一季 ($T - \tau$) 是時階 T 的預測。演算法會在集合 $\{0, \dots, T - 1\}$ 中隨機取樣時間指數 (t)，以產生目前時階 T 的樣本。NPTS 對於含有許多零的間歇性 (有時也稱為稀疏) 時間序列尤有效用。Forecast 還包含 Prophet 的 Python 實作，這是一種貝葉斯結構的時間序列模型。

Amazon Forecast 提供兩種全域深度學習演算法：

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

全域模型會對資料集中的整個時間序列集合中訓練單一模型。在一組橫截面單位間有類似的時間序列時，此模型特別有用。例如，不同產品、伺服器負載和網頁要求之需求的時間序列分組。

一般而言，隨著時間序列數量增加，CNN-QR 和 DeepAR+ 的功效也會提升。但本機模型不一定是如此。深度學習模型也可在 (幾乎) 沒有歷史銷售資料的情況下用來產生新 SKU 的預測。這稱為 [冷啟動預測](#)。

	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)	✔	✔	✘	✘	✘	✘

比較 Amazon Forecast 中可用的演算法

如需相關時間序列的詳細資訊，請參閱[相關時間序列](#)。

步驟 4：評估預測器

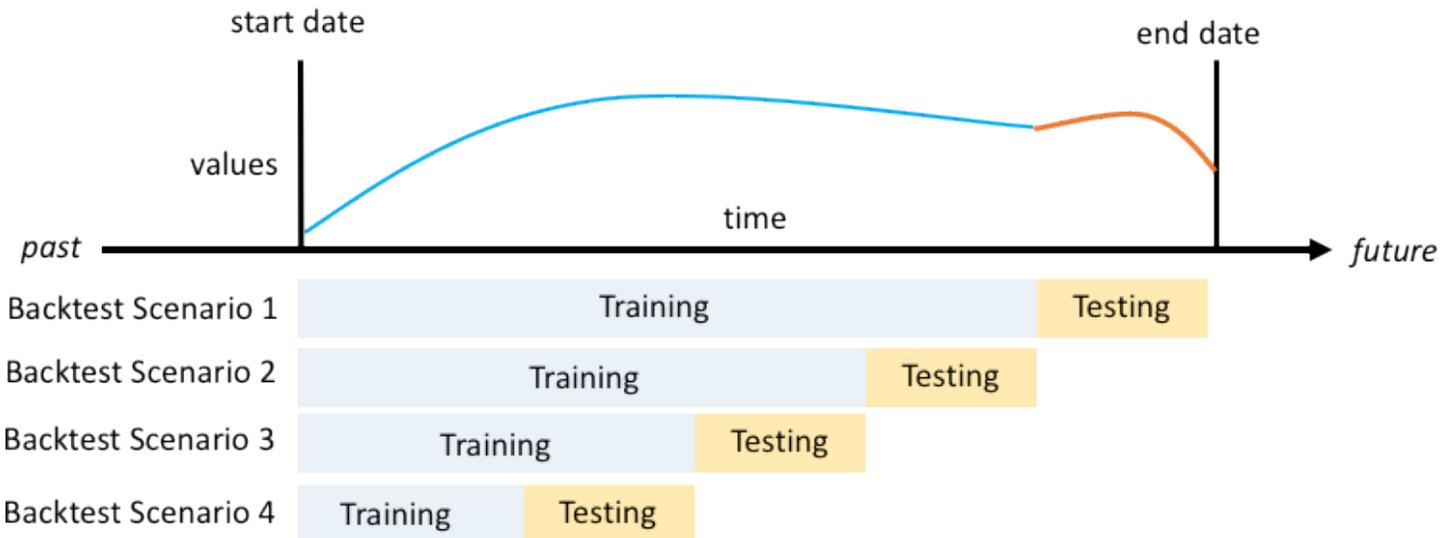
機器學習中的常見工作流程包括以訓練集訓練一組模型或模型的組合，以及在鑑效資料集上評估其準確性。本節討論如何分割歷史資料，以及要使用哪些指標來評估時間序列預測的模型。對於預測，回測技術是評估預測準確性的主要工具。

回測

適當的評估和回測架構，是藉由機器學習的應用獲致成功的首要因素之一。憑藉對模型的成功回測，您對於模型的未來預測能力將更有信心。此外，您可以透過超參數最佳化 (HPO) 來調整模型、學習模型組合，以及啟用中繼學習和 AutoML。

時間序列預測中的時間特性，使其在評估和回測方法方面與應用機器學習的其他領域有所不同。在 ML 任務中，若要藉由回測評估預測誤差，通常可按項目來分割資料集。例如，若要在影像相關任務中進行交叉驗證，您可以訓練某個百分比的圖片，然後將其他部分用於測試和驗證。在預測中，您主要需按時間分割 (而在較低程度上按項目分割)，以確保您不會將訓練集的資訊流入測試集或驗證集中，且您將盡可能真實模擬生產案例。

按時間分割必須謹慎，因為您要選擇的是多個時間點，而不是單一時間點。若是單一時間點，準確性將過於依賴於藉由分割點而定義的預測開始日期。滾動式預測評估可讓您在多個時間點進行一系列分割，並輸出平均結果，而產生更穩健可靠的回測結果。下圖說明四種不同的回測分割。



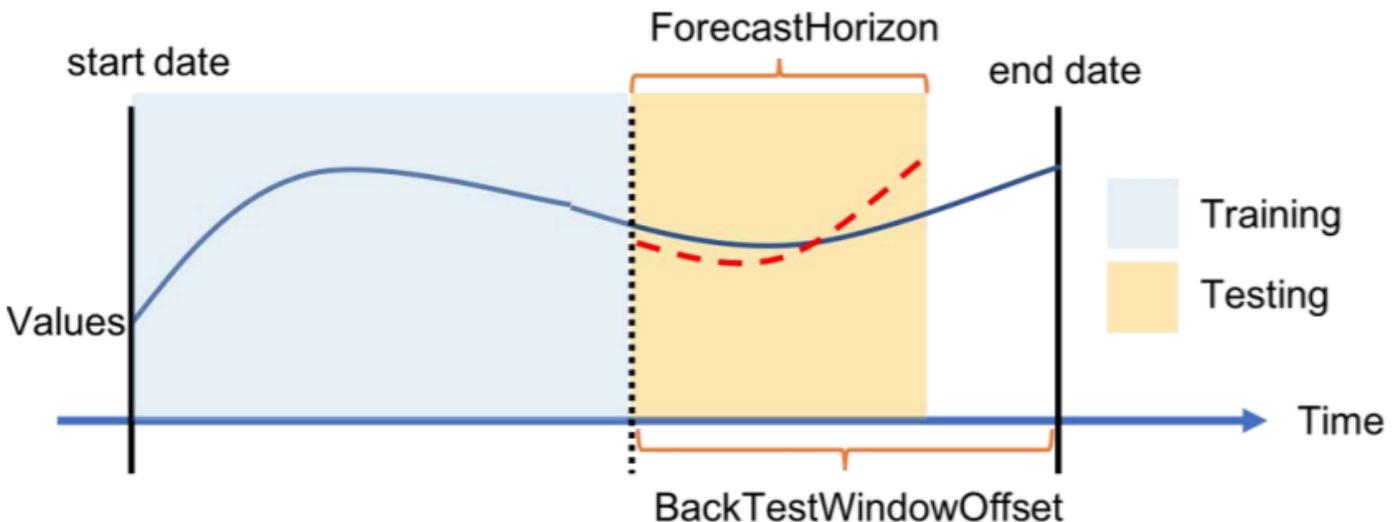
此圖例顯示訓練集大小增加、但測試集大小不變的四種不同的回測案例

在上圖中，所有回測案例在其整個過程中都有可用的資料，而能夠根據實際值評估預測值。

需要多個回測時段的原因是，現實世界中的多數時間序列通常是非靜止的。案例研究中的電子商務企業將總部設在北美洲，其產品需求許多都是來自 Q4 旺季的推動，尤其是感恩節前後和聖誕節之前的高峰。在 Q4 購物旺季，時間序列的變化性會高於一年中的其餘時間。有了多個回測時段，您即可在更平衡的設定中評估預測模型。

下圖針對每個回測案例顯示 Amazon Forecast 術語中的基本要素。Amazon Forecast 會自動將資料分割到訓練資料集和測試資料集。Amazon Forecast 會使用在 `create_predictor` API 中指定為參數的 `BackTestWindowOffset` 參數來決定如何分割輸入資料，或使用其 `ForecastHorizon` 的預設值來決定。

在下圖中，`BackTestWindowOffset` 和 `ForecastHorizon` 參數不相等，您會看到較為典型的第一種情況。`BackTestWindowOffset` 參數會定義虛擬預測開始日期，在下圖中顯示為虛線垂直線。這可用來回答以下假設問題：如果模型在這一天部署，將會有何預測？`ForecastHorizon` 會定義從虛擬預測開始日期到預測的時階數。



此圖例顯示單一回測案例及其在 Amazon Forecast 中的組態

Amazon Forecast 可匯出在回測期間產生的預測值和準確性指標。匯出的資料可用來就特定的時間點和分位數評估特定項目。

預測分位數和準確性指標

預測分位數可提供預測的上限和下限。例如，對於使用預測類型 0.1 (P10)、0.5 (P50) 和 0.9 (P90) 產生的值範圍，我們可稱之為 P50 預測的 80% 信賴區間。藉由產生對 P10、P50 和 P90 的預測，您可以預期實際值有 80% 的時間會落在這些界限內。

本白皮書將在[步驟 5](#) 中進一步討論分位數。

Amazon Forecast 會使用加權分位數損失 (wQL)、均方根誤差 (RMSE) 和加權絕對百分比誤差 (WAPE) 準確性指標來評估回測期間的預測器。

加權分位數損失 (wQL)

加權分位數損失 (wQL) 誤差指標會測量模型預測在指定分位數上的準確性。這在預測不足和過度預測的成本不同時尤有效用。設定 wQL 函數的權重 (τ)，會自動納入用於預測不足和過度預測的不同懲罰。

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

wQL 函數

其中：

- τ — {0.01, 0.02, ..., 0.99} 集合中的一個分位數
- $q_{i,t}(\tau)$ — 模型預測的 τ -quantile (τ 分位數)。
- $y_{i,t}$ — 在點 (i,t) 處的觀測值

加權絕對百分比誤差 (WAPE)

加權絕對百分比誤差 (WAPE) 是測量模型準確性的常用指標。它會測量預測值與觀測值的整體偏差。

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

WAPE

其中：

- $y_{i,t}$ - 在點 (i,t) 處的觀測值
- $\hat{y}_{i,t}$ - 在點 (i,t) 處的預測值

Forecast 會使用平均值預測作為預測值 $\hat{y}_{i,t}$ 。

均方根誤差 (RMSE)

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

均方根誤差 (RMSE) 是測量模型準確性的常用指標。和 WAPE 一樣，它會測量估計值與觀察值的整體偏差。

其中：

- $y_{i,t}$ - 在點 (i,t) 處的觀測值
- $\hat{y}_{i,t}$ - 在點 (i,t) 處的預測值
- nT - 測試集中的資料點數量

Forecast 會使用平均值預測作為預測值 $\hat{y}_{i,t}$ 。計算預測器指標時， nT 是回測時段中的資料點數量。

WAPE 和 RMSE 的相關問題

在大多數的情況下，可在內部或從其他預測工具產生的點預測，應會與 p50 分位數或平均值預測相符。對於 WAPE 和 RMSE，Amazon Forecast 會使用平均值預測來表示預測值 (yhat)。

若 $w\text{QL}[\tau]$ 方程式中的 $\tau = 0.5$ ，則兩個權重會相等，且 $w\text{QL}[0.5]$ 會降至點預測常用的加權絕對百分比誤差 (WAPE)：

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

其中， $y\text{hat} = q(0.5)$ 是運算預測。 $w\text{QL}$ 公式中使用縮放係數 2 來取消 0.5 係數，以取得精確的 $\text{WAPE}[\text{median}]$ 運算式。

請注意，上述 WAPE 定義與平均絕對百分比誤差 ([MAPE](#)) 的通用解釋不同。差異在於分母。上述 WAPE 的定義方式迴避了除以 0 的問題，這是實際情況下常見的問題，例如案例研究中的電子商務企業，常會有給定 SKU 在給定日期售出 0 單位的情形。

不同於使用加權分位數損失指標， τ 不等於 0.5 時，每個分位數中的固有偏差無法藉由 WAPE (其權重相等) 之類的計算得出。WAPE 的缺點還包括它不是對稱的、對微小數字的百分比誤差會過度膨脹，以及只是一個逐點指標。

RMSE 是 WAPE 中的誤差項平方，在其他 ML 應用程式中則是常見的誤差指標。RMSE 指標較適用於個別誤差具有一致幅度的模型，因為較大的誤差變化將會超比例推升 RMSE。由於平方誤差，一些不當預測的值在其他方面皆良好的預測中可能會推升 RMSE。此外，由於平方項，較小的誤差項在 RMSE 中的權重會低於 WAPE 中的權重。

準確性指標可用來進行預測的定量評估。對大規模的比較 (方法 A 整體上是否優於方法 B) 而言，這一點格外重要。不過，以個別 SKU 的視覺效果予以補強，往往也很重要。

步驟 5：產生並使用預測進行決策

一旦有模型達到您特定使用案例所需的準確性閾值 (由回測判定)，即可執行最後一個步驟：部署模型並產生預測。若要在 Amazon Forecast 中部署模型，您必須執行 `Create_Forecast` API。此動作會託管藉由對整個歷史資料集進行訓練而建立的模型 (不同於 `Create_Predictor` 只會將資料分割到訓練集和測試集中)。在預測期間內產生的模型預測隨後將有兩種運用方式：

- 您可以從 [AWS CLI](#) 或直接透過 [AWS Management Console](#) 使用 `Query_Forecast` API，來查詢特定項目的預測 (藉由指定項目，或指定項目/維度的組合)。
- 您可以使用 `Create_Forecast_Export_Job` API，對所有分位數的所有項目/維度組合產生預測。此 API 會產生一個 CSV 檔案，並安全地存放在您選擇的 [Amazon Simple Storage Service](#) (Amazon S3) 位置。然後，您可以使用 CSV 文件中的資料，並將其導入您用於決策的下游系統中。例如，您現有的供應鏈系統可以直接從 Amazon Forecast 擷取輸出，以協助您對特定 SKU 的製造做出相關決策。

機率預測

Amazon Forecast 可就不同的分位數產生預測，這在預測不足和過度預測的成本不同時尤有效用。與預測器訓練階段類似，同樣可對 p_1 與 p_{99} 之間的分位數產生機率預測。

根據預設，Amazon Forecast 會就預測器訓練期間使用的相同分位數產生預測。如果在預測器訓練期間未指定分位數，依預設將會在 p_{10} 、 p_{50} 和 p_{90} 產生預測。

就 p_{10} 預測而言，預期真正的值有 10% 的時間會低於預測值，並且可使用 $wQL[0.1]$ 指標來評估其準確性。這表示 P_{10} 預測有 90% 的時間是預測不足的，且若用於庫存時，品項將有 90% 的時間處於售罄狀態。沒有大量儲存空間，或投入資金的成本過高時， P_{10} 預測可能有其效用。

Note

分位數預測的正式定義是 $\Pr(\text{實際值} \leq \text{分位數 } q \text{ 的預測}) = q$ 。就技術上而言，分位數就是百分位數/100。統計學家傾向於表述為「 P_{90} 分位數層級」，因為比「分位數 0.9」容易表述。例如， P_{90} 分位數層級預測表示可以預期實際值將有 90% 的時間小於預測。具體來說，如果時間為 t_1 ，分位數層級為 0.9，預測值為 30，這表示如果您模擬了 1,000 次，時間 t_1 的實際值預期將小於 30 (模擬 900 次)，若模擬 100 次，實際值預期會超過 30。

另一方面，P90 預測是指有 90% 的時間過度預測，這在不銷售某個品項的機會成本極高或投入資金成本較低時，將有其效用。就雜貨店來說，P90 預測可用在牛奶或衛生紙之類的商品上，因為店家絕不樂見這類商品售罄，也不會介意架上總是有些存貨。

就 p50 預測 (通常也稱為中位數預測) 而言，預期真正的值有 50% 的時間會低於預測值，並且可使用 $wQL[0.5]$ 指標來評估其準確性。若無庫存過多的顧慮，且對特定品項有適量的需求，p50 分位數預測可能有其效用。

視覺效果

Amazon Forecast 支援直接在 AWS Management Console 中繪製預測。此外，您還可以利用完整的 Python 資料科學堆疊 (請參閱 [Amazon Forecast 範例](#))。Amazon Forecast 支援透過 `ExportForecastJob` API 將預測匯出為 CSV 檔案，讓使用者可在自己選擇的分析工具中呈現預測。



Amazon Forecast 控制台針對不同的分位數提供的視覺效果

預測工作流程和 API 摘要

下表說明預測工作流程的每個步驟及其各自對應的 Amazon Forecast API。

表 1：預測步驟和 Amazon Forecast API

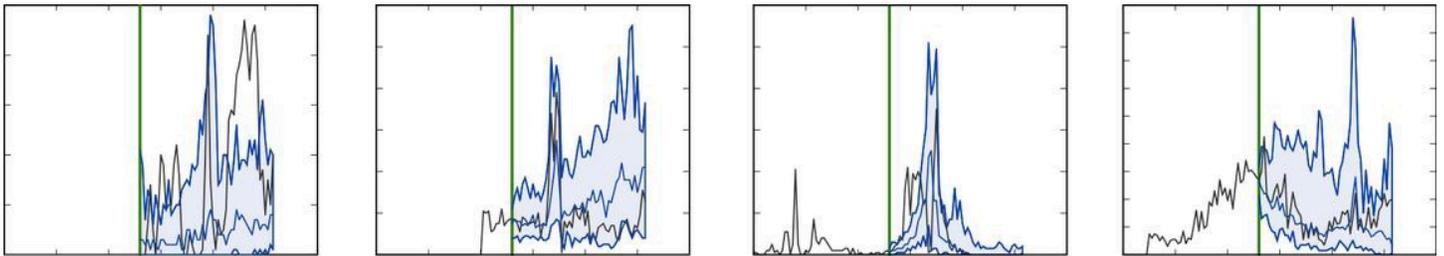
步驟	API	API 函數
步驟 1：收集和彙總資料 步驟 2：準備資料	Create_Dataset_Group、Create_Dataset、Create_Dataset_Import_Job	<ol style="list-style-type: none"> 1. 建立問題的高階領域 (零售、指標等等)。 2. 定義不同資料集 (目標、相關、項目中繼資料) 的結構描述。 3. 將資料從 Amazon S3 匯入 Amazon Forecast 中。
步驟 3：建立預測器 步驟 4：評估預測器	Create_Predictor	<ol style="list-style-type: none"> 1. 執行 ETL。 2. 將資料分割到訓練集/測試集中，並訓練模型。 3. (選擇性) 使用 Create_predictor_backtest_Export_job 將回測結果匯出為 CSV，用以計算項目層級指標。
步驟 5：產生並使用預測進行決策	Create_Forecast	<ol style="list-style-type: none"> 1. 訓練/託管模型。 2. 對相關的特定分位數 (例如，任何介於 1 到 99 之間的整數，包括平均值) 產生預測期間內的預測。
	Query_Forecast、Create_Forecast_Export_Job	可讓您取用由 Create_Forecast 建立的預測

在常見案例中使用 Amazon Forecast

您也可以根據外部變數 (例如價格或促銷) 的變更產生不同的預測，以執行假設分析。例如，在電子商務案例研究範例中，您可以根據您可能規劃的促銷建立不同的預測。您可以先預測產品有 10% 折扣時的需求，再預測 20% 折扣時的需求，以了解您需要多少庫存量來滿足產品需求。為此，您可以根據相關的情況設定獨特的資料集群組，並更新各群組中的相關時間序列。

此外，您也可以對沒有歷史記錄的品項產生預測 (有時稱為冷啟動問題)。此方法需要使用 DeepAR+ 或 CNN-QR 以及中繼資料 (例如項目中繼資料資料集) 建立預測器，以產生新品項的預測。

下圖顯示四種不同 SKU 出現在實際營運預測問題中時的範例。



四種不同 SKU 出現在實際營運預測問題中時的範例。

在上圖中，觀察到的實際值位於垂直線左側，預測以藍色顯示於垂直線右側，與黑色的實際值相比較。請注意，每個 SKU 的歷史記錄 (位於垂直線左側)，並不表示在綠線右側會有同等的演變。

將預測導入生產中

完成端對端 Amazon Forecast 工作流程後，務必要釐清 `Create_Predictor` 與 `Create_Forecast` API 的主要差異及其各自的適用時機。

前者主要在概念驗證期間用來評估模型準確性/指標，後者則在生產環境中用來產生預測。

進入生產環境後，`Create_Predictor` 就無須在每次需要產生預測時都執行，而是僅在因資料有所變更或基於預先建立的頻率 (例如，每兩週或每個月) 而需重新訓練模型時，才需要執行。由於資料集已使用新資料進行更新，因此只需執行 `Create_Forecast` 以產生新預測期間的預測。

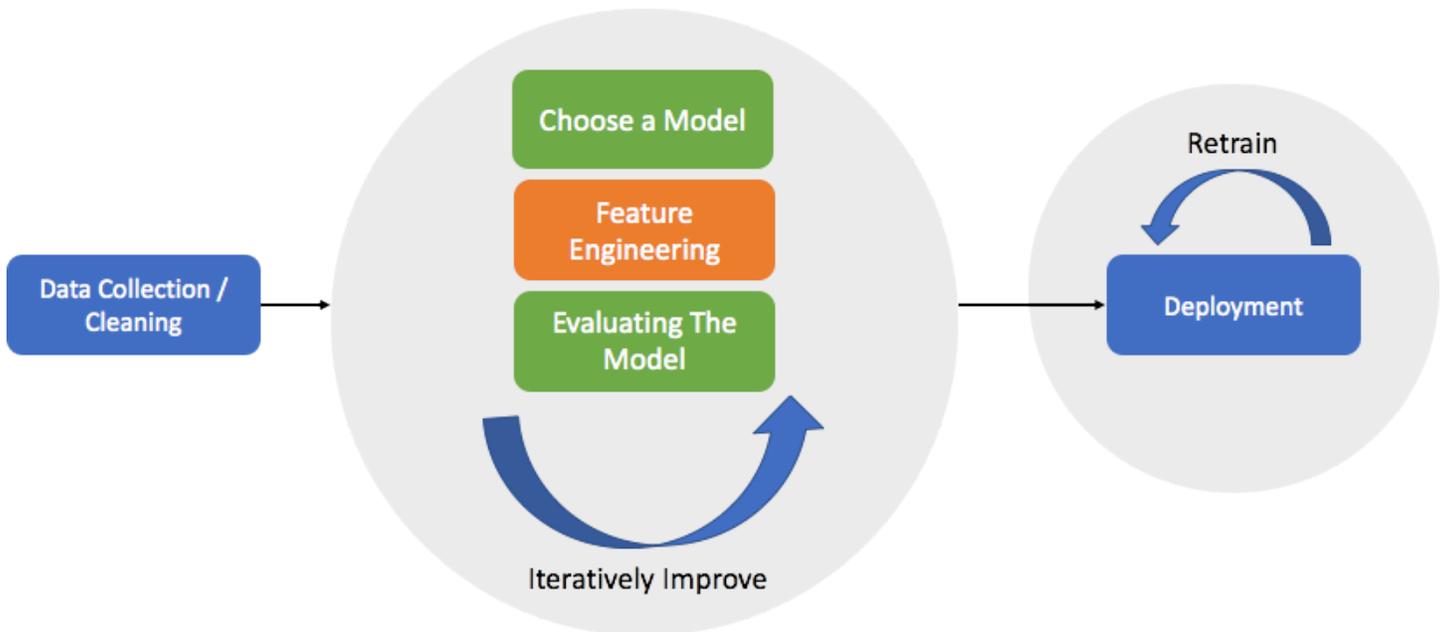
在生產環境中，您還需要自動執行資料集匯入和預測操作，以便產生滾動式的新預測。現今，您可以搭配使用 [Amazon CloudWatch Events](#) 日誌、[AWS Step Functions](#) 和 [AWS Lambda](#) 函數進行 cron 工作的設定，以達成此目的。[cron](#) 工作設定完成後，將自動執行 Amazon Forecast API 呼叫，以進行匯入/重新訓練或產生預測。最後，請務必管理資源並定期予以刪除，這樣就不會超出服務規定的[系統限制](#)。請參閱這篇包含 [Amazon Redshift](#) 的[部落格文章](#)，以進一步了解如何設定排程工作。

結論

[Januschowski 和 Kolassa \(2019\)](#) 提供了與企業需要進行的決策 (包括策略、戰術和運營決策) 一致的預測問題分類。每個決策層級都有相對應的預測任務。

營運和戰術性預測問題的特點是包含大量資料，且通常需要高度的自動化。不同的預測方法可以因應這些問題。本機預測方法通常適用於策略性預測問題，以深度學習為基礎的方法適用於營運預測問題，而對介於兩者之間的問題，則可能需要一些實驗。雖然本白皮書討論營運預測問題，但 Amazon Forecast 並非僅侷限於其提供的模型，而是包含提供策略性以及營運和戰術性預測問題的模型。

運營預測問題解決程序可分成若干基本步驟，從資料收集和準備開始，後續是模型的建置和部署。通常，將其視為迭代程序而非線性程序，是最有用的。例如，在進一步理解模型和使用案例後，回到資料收集階段可能是合理的。模型開發本身也是高度迭代的。



將預測模型導入生產的簡化開發程序。

作者群

此文件的作者包括：

- AI Vertical Services 資深機器學習科學家 Yuyang Wang
- AI Vertical Services 應用科學家 Danielle Robinson
- ML 應用科學經理 Tim Januschowski
- AI Vertical Services 資深產品經理 Namita Das
- 資深 AI/ML 專家解決方案架構師 Christy Bergman
- AI/ML 技術文件撰寫人員 Kris Tonthat

深入閱讀

如需時間序列預測和深度學習方法的其他資訊，請參閱：

- [Amazon Forecast 文件](#)
- [Amazon Forecast 正式版本部落格](#)
- [現已在 Amazon SageMaker 中提供：更精確的 DeepAR 演算法](#)
- [Amazon SageMaker DeepAR 現在支援遺漏值、分類和時間序列特徵，以及通用頻率](#)
- [Amazon Forecast 現在可以使用卷積神經網路 \(CNN\) 來訓練預測模型，速度可加快 2 倍，準確性可提高 30%](#)
- [Amazon Forecast 現在支援個別項目的準確性測量](#)
- [透過 Amazon Forecast 測量預測模型準確性，以最佳化您的業務目標](#)
- [Amazon Forecast 氣象指數 — 自動納入當地天氣以提高預測模型的準確性](#)
- [關於時間序列預測模型的科學論文](#)
- [Amazon Forecast 樣本 GitHub 頁面](#)
- [AWS 架構中心](#)

附錄 A：常見問答集

問：如何開始使用 Amazon Forecast？

1. 首先，您需要一個 AWS 帳戶。
2. 接著，在 [AWS Management Console](#) 中開啟 Forecast 服務，建立資料集群組，然後將 .csv 檔案匯入至目標時間序列資料集 (必要)。入門所需的最低資料是您要預測之數量的歷史資料，例如，每個家庭每個時間戳記的用電量。
3. 最後，藉由執行 [CreatePredictor](#) 建立一個模型，並執行 [CreateForecast](#) 以產生結果。如需詳細資訊，請參閱 [入門](#) 文件頁面。

另請參閱 [GitHub 簡介和最佳實務指南](#)。

問：Amazon Forecast 適合我嗎？

並非所有機器學習問題都是預測問題。首先應該自問：「我的業務問題是否在陳述中包含時間序列？」例如，您是否只需要未來特定時間和日期的特定值？預測不適用於一般、靜態 (特定日期/時間無關緊要) 的問題，例如詐騙偵測或向使用者推薦電影片名。靜態問題有比這快得多的解決方案。

除了有時間序列資料以外，資料本身還應該是「密集」的，且具有較長的歷史。下表總結了這一點：

表 2 — 條件和 Amazon Forecast 演算法類別

條件	Amazon Forecast 演算法類別
含有多達 500 萬個時間序列的大型資料集，具有類似的基礎模式 + 季節性效果 + 相關資料。每個時間序列均應有很長的歷史 (理想情況下，如果嘗試擷取年度事件，則應超過兩年)，且每個時間序列應有超過 300 個資料點 (理想情況下至少要有 1K)。	Amazon Forecast 專屬的深度學習 DeepAR +、CNN-QR
含有 1-100 個時間序列的小型資料集，其中大多數的時間序列具有超過 300 個資料點 + 季節性效果 + 相關資料。	Prophet

條件	Amazon Forecast 演算法類別
含有 1-10 個時間序列的小型資料集，其中大多數的時間序列具有超過 300 個資料點 + 季節性效果。	ETS、ARIMA
含有 1-10 個時間序列的間歇 (包含許多 0 的稀疏) 資料集，其中大多數時間序列具有超過 300 個資料點。	Amazon Forecast 專屬的 NPTS
含有 1-10 個時間序列的小型資料集 (一般或稀疏)，其中大多數的時間序列具有超過 300 個資料點。	資料對 Amazon Forecast 而言太小。請改為在 Excel 或傳統的統計模型 ARIMA 和 Prophet 中嘗試使用 ETS。

最佳實務是在第一次訓練資料時，在您的預測器中使用 AutoML 模式進行訓練。AutoML 會自動執行所有演算法 (DL 演算法會在開啟 HPO 的情況下執行)，以了解何種演算法最適合您的資料。

問：如何認定資料遺漏？在何種程度下將無法產生合理的預測？

這有可能是指資料的記錄出現問題，或資料的彙總層級過低或過高。通則是 Forecast 長度不可超過訓練資料的 1/3。

除了遺漏資料的數量外，另一項考量因素是遺漏資料的原因。您可以將所有 0 轉換為空值，讓 Amazon Forecast 代為執行自動推算遺漏值的繁重工作。Amazon Forecast 會自動偵測遺漏值是因為新產品的導入 (冷啟動) 還是生命週期結束的產品所導致的。您可以使用數個遺漏值邏輯，包括值、中位數、最小值、最大值、零、平均值和 NaN (僅限目標時間序列)。請參閱[空值填入語法的文件](#)。

- "frontfill" - (僅限 TTS) 是指新的或冷啟動的品項，以及在品項開始有任何歷史記錄之前要如何處理空值
- "middlefill" — 是指時間序列值中間的空值
- "backfill" — 是指生命週期結束的品項，以及在品項停止銷售後要如何處理空值
- "futurefill" — (僅限 RTS) 是指在訓練資料結束後發生的空值

問：我的輸入歷史資料沒有負值，但我在需求預測中看到負值。為什麼會這樣？如何避免這種情況？

對於 NPTS (以非負資料進行訓練) 和 DeepAR (使用負二項式相似度函數) 以外的所有模型，不保證會產生正數。解決方案是變更為前述模型之一，或將預測值截斷為非負值。

問：分位數的準確性指標為何有所不同？既然模型相同，誤差不也應相同嗎？

有關加權如何依存於分位數的更多說明，請參閱[加權分位數損失 \(wQL\)](#)。

假設你對以下三個不同的分位數進行預測：p10、p50、p90。這三個預測本身就是隨機變數。準確性是在每個分位數層級的實際值與預測之間個別計算的。您可能會看到如下的“wQL” (加權分位數損失) 表格。wQL 值彼此之間沒有確定性關係。(記住，損失就是誤差，所以是無序的；然而，分位數預測是有序的)。因此，舉例來說，p90 wQL 不一定會大於 p50 wQL。

表 3 — 預測分位數範例

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

問：如何改善預測準確性？

預測準確性取決於是否有適當數量和質量的適當資料可供使用。如果準確性不如預期，了解問題有多少可預測性 (或資料的隨機/雜訊/靜止程度) 可能會有幫助。其他需要考量的因素包括評估不同的模型、超參數設定，以及使用相關的時間序列和項目中繼資料資料集來併入其他特徵。如需具體建議，請參閱[GitHub 上的這份最佳實務文件](#)。

問：有個演算法非常適合運用在我的使用案例，但 Amazon Forecast 中並未提供。我該怎麼做？

Amazon Forecast 團隊將很樂意為您提供此使用案例的協助。請將電子郵件寄到 <amazonforecast-poc@amazon.com>，與 Amazon Forecast 的服務團隊聯繫。

附錄 B：參考資料

[Tim Januschowski 與 Stephan Kolassa。商業預測問題的分類。前瞻性：國際應用預測雜誌。2019](#)

[David Salinas、Valentin Flunkert、Jan Gasthaus 和 Tim Januschowski。DeepAR：利用自回歸遞歸網路進行機率預測。國際預測雜誌。2019](#)

[Jan Gasthaus、Konstantinos Benidis、Yuyang Wang、Syama Sundar Rangapuram、David Salinas、Valentin Flunkert 和 Tim Januschowski。利用曲線分位數函數 RNN 進行機率預測。第 22 屆人工智慧與統計國際會議。2019](#)

[Tim Januschowski、Jan Gasthaus、Yuyang Wang、David Salinas、Valentin Flunkert、Michael Bohlke-Schneider 和 Laurent Callot。對預測方法進行分類的條件。國際預測雜誌。2019 \(需要登入\)](#)

[Tim Januschowski、Jan Gasthaus、Yuyang Wang、Syama Rangapuram 和 Laurent Callot。預測的深度學習。前瞻性：國際應用預測雜誌。2018](#)

[Tim Januschowski、Jan Gasthaus、Yuyang Wang、Syama Sundar Rangapuram 和 Laurent Callot。預測的深度學習：當前趨勢和挑戰。前瞻性：國際應用預測雜誌。2018](#)

[Joos-Hendrik Bose、Valentin Flunkert、Jan Gasthaus、Tim Januschowski、Dustin Lange、David Salinas、Sebastian Schelter、Matthias Seeger 和 Yuyang Wang。大規模機率需求預測。VLDB Endowment 紀事。2017](#)

文件歷史記錄

若要收到此白皮書更新的通知，請訂閱 RSS 摘要。

變更	描述	日期
白皮書已更新	更新。	2021 年 9 月 1 日
初版	白皮書初始出版。	2020 年 2 月 4 日

Note

若要訂閱 RSS 更新，您必須為正在使用的瀏覽器啟用 RSS 外掛程式。

聲明

客戶應負責對本文件中的資訊自行進行獨立評估。本文件：(a) 僅供參考之用，(b) 代表目前的 AWS 產品供應與實務，如有變更恕不另行通知，以及 (c) 不構成 AWS 及其附屬公司、供應商或授權人的任何承諾或保證。AWS 產品或服務以「現況」提供，不提供任何明示或暗示的擔保、主張或條件。AWS 對其客戶之責任與義務，應受 AWS 協議之約束，且本文件並不屬於 AWS 與其客戶間之任何協議的一部分，亦非上述協議之修改。

© 2021 Amazon Web Services, Inc. 或其關係企業。保留所有權利。

AWS 詞彙表

如需最新 AWS 術語的清單，請參閱 AWS 一般參考資料中的 [AWS 詞彙表](#)。