

AWS Academy Machine Learning Foundations

模組 4：預測簡介



歡迎學習“模組 4：預測簡介”。

模組概覽



小節目錄

1. 預測概覽
2. 處理時間序列資料
3. 使用 Amazon Forecast
4. 引導式實驗
5. 模塊總結

演示

- 使用 Amazon Forecast 創建預測

引導式實驗

- 使用 Amazon Forecast 創建預測



知識測驗

本模組包含以下主題：

- 預測概覽
- 處理時間序列資料
- 使用 Amazon Forecast
- 引導式實驗
- 模塊總結

此外，本模組包含一項引導式動手實驗，教您如何在處理時間序列資料時使用 Amazon Forecast。

最後，您需要完成一個知識測驗，以測試您對本模組中涵蓋的關鍵概念的理解程度。

模組目標



學完本模組後，您應該能夠：

- 描述使用 Amazon Forecast 可解決的業務問題
- 描述使用時間序列資料時面對的挑戰
- 列出使用 Amazon Forecast 創建預測所需的步驟
- 使用 Amazon Forecast 進行預測

完成本模組後，您應能夠：

- 描述使用 Amazon Forecast 可解決的業務問題
- 描述使用時間序列資料時面對的挑戰
- 列出使用 Amazon Forecast 創建預測所需的步驟
- 使用 Amazon Forecast 進行預測

模組 4：預測簡介

第 1 節：預測概覽



介紹“第 1 節：預測概覽”。

先回顧一下預測的含義，然後瞭解一些預測使用案例。

- 根據歷史資料預測未來值
 - 可以是單變數或多變數
- 常見模式 –
 - 趨勢：增加、減少或停滯的模式
 - 季節性：基於季節的模式
 - 週期性：其他重複模式
 - 不規則：似乎具有隨機性的模式



趨勢資料



季節性數據



週期性數據



不規則數據

預測是機器學習的一個重要領域。它之所以重要，是因為許多預測未來結果的機會都基於歷史資料。其中許多機會都涉及時間元件。儘管時間元件增加了更多資訊，但與其他類型的預測相比，也使得時間序列問題更加難以處理。

您可以將時間序列資料分為兩大類。第一類是**單變量**，這意味著它只有一個變數。第二類是**多變數**，這意味著它具有多個變數。除了這兩個類別，大多數時間序列資料集還遵循以下模式：

- 趨勢 – 顯示值隨著時間的推移而增加、減少或保持不變的模式
- 季節性 – 基於一年中的季節的重複模式
- 週期性 – 重複模式的其他形式
- 不規則 – 資料隨著時間的推移而發生變化，但變化似乎具有隨機性或沒有明顯的模式

預測使用案例



銷售和需求預測



庫存預測



能源消耗



天氣預報

您可以對各種領域使用預測，一些比較常見的應用包括：

- 行銷應用，如銷售預測或需求預測。
- 庫存管理系統，預測所需的庫存水準。通常，此類預測包含有關交貨時間的資訊。
- 能源消耗，確定何時何地需要能源。
- 政府天氣預報系統，以及農業等商業應用。

模組 4：預測簡介

第 2 節：處理時間序列資料



介紹 “第 2 節：處理時間序列資料”

使用時間序列資料會帶來一些獨特的挑戰，現在我們來瞭解一下。

時間序列資料

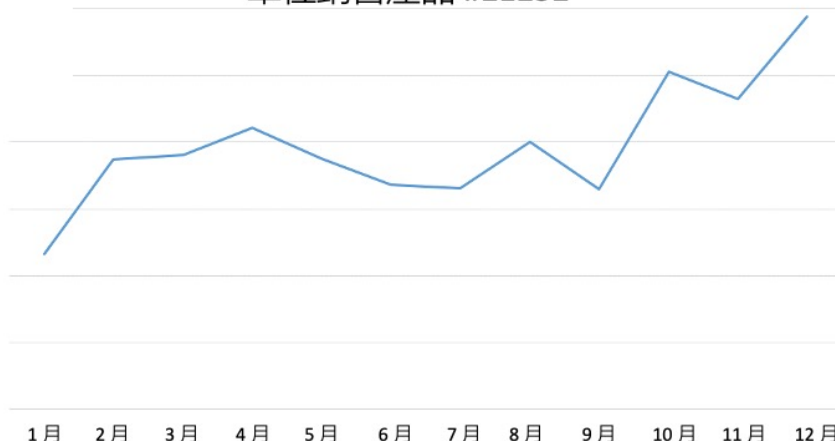


時間序列資料 按時間 順序捕獲

相關數據為時間序列資料（例如，價格或促銷）提供資訊支援

中繼資料可能也要用來解釋預測，例如，品牌名稱或類別

單位銷售產品 #21232



時間序列資料是在指定的時間段內按時間順序捕獲的資料。將時間引入機器學習模型會產生積極的影響，因為模型可以隨著時間的變化從資料點變化中推導出意義。時間序列資料往往是相關的，這意味著資料點之間存在依賴關係。

因為牽扯到回歸問題，並且回歸假定資料點是獨立的，所以您必須採取某種方法來處理資料依賴關係。這種方法的目的是提高預測的有效性。

除了時間序列資料，您還可以添加相關資料來增強預測模型。例如，對於有關零售的預測，您可以包含有關所售產品的資訊（如商品標識或售價）。這些資訊是單位時間段內出售的件數之外的資訊。

第三類資料是有關資料集的中繼資料。例如，對於零售資料集，您可能希望包含中繼資料以對結果進行分組，如品牌名稱。再比如，中繼資料可能包括音樂或視頻類型。

時間和日期挑戰



不完整和多變的時間戳記

yyyy-mm-dd HH:MM:SS	包含時間
yyyy-dd-mm	年、日、月
yyyy-mm-dd	年、月、日
yyyy-mm	無日期
ss:s	秒
mm-dd	無年份

UTC、當地和時區：

時間是否採用 UTC 格式？

2020-06-02T13:15:30Z

擁有的資料越多越好。對於多個資料來源，您可能遇到的一個難題是資料時間戳記問題。您會發現時間戳記格式有所差異，以及其他挑戰，比如資料不完整。在某些情況下，您可能要推斷缺失資料。例如，假設您的某些資料包含月份和日期，但不包含年份。假設資料在資料庫中按月數排序，並在 12 之後重複。在這種情況下，如果知道資料何時開始，則可以添加年份。您可以根據資料的順序推斷未來幾年。

許多資料都以世界協調時間 (UTC) 格式存儲，但並非所有資料都採用 UTC 格式。您應該檢查時間戳記是當地時間還是世界時間。

有時，時間戳記並不代表您所認為的時間。例如，假設您有一個關於在汽車修理廠維修的汽車的資料庫。時間戳記是否指示汽車到達、維修完畢或提走的時間？或者是否指示最後一條記錄錄入系統的時間？

如果您試圖對患者每小時的卡路里攝入量建模，但您僅擁有每日的資料，則必須調整目標時間表。

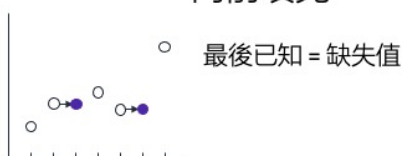
您的資料中可能沒有顯示時間戳記。您可能還會採用其他方式來推斷時間序列，具體取決於資料和領域。例如，您的圖像中可能具有波長測量值或向量。

世界各地的夏令時各不相同。由於夏令時，某些時間在其時區每年會出現兩次。

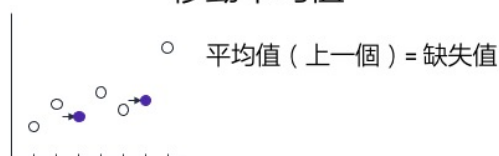
時間序列處理：缺失數據



向前填充



移動平均值

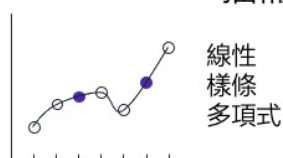


向後填充



危險！Lookahead 風險！

插補



注意：零有時是理想的填充值

在實際的預測問題中，一個常見的情況是原始資料中缺失值。缺失值使得模型更加難以生成預測。零售業主要面臨的是對缺貨情況進行需求預測。如果某商品缺貨，則當天的銷售額為零。如果根據這些零銷售值生成預測，這樣的預測將不正確。

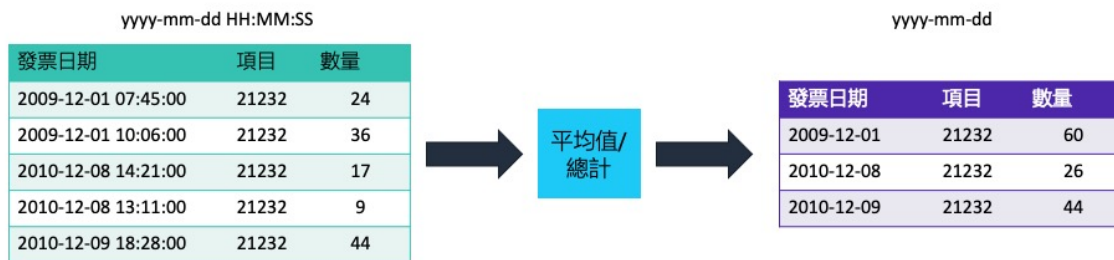
出於各種原因，可以將缺失值標記為缺失。由於沒有業務，或者可能由於衡量誤差，會出現缺失值的情況。也許監控特定資料的服務無法正常工作，或者無法進行正確測量。

可以通過以下幾種方法計算缺失資料：

- 向前填充 – 使用上一個已知值作為缺失值。
- 移動平均值 – 使用最近幾個已知值的平均值來計算缺失值。
- 向後填充 – 使用下一個已知值作為缺失值。請注意，用未來計算過去存在潛在危險，這並不是良好的預測。這種做法稱為 *Lookahead*，應加以避免。
- 插補 – 其本質上是使用方程式來計算缺失值。

您還可以選擇使用零填充。這種方式通常用於零售業，在該領域中不能計算缺失的銷售資料。缺失資料表示當天沒有訂單。在這種情況下，最好是調查原因，不能填充缺失的值。

時間序列處理：下採樣

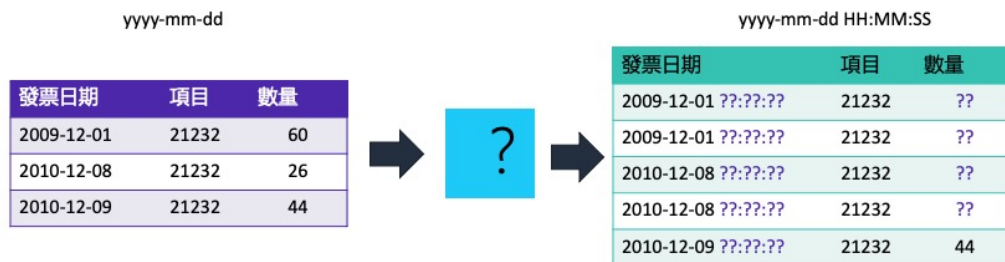


您可能會以不同的頻率獲取資料。例如，您可能擁有包含確切記錄銷售時間戳記的銷售資料。但庫存資料可能僅包含年、月和日的庫存水準。如果資料獲取頻率不同於其他資料集，或者與您的問題不相符，則可能需要下採樣。

下採樣意味著從細微性級別較高的時間移向細微性級別較低的时间。本示例將每小時資料集轉換為每日資料集。

在下採樣時，必須確定如何合併這些值。對於銷售資料，對數量進行求和最為合理。如果是溫度資料，則可能需要計算平均值。瞭解資料有助於確定最佳行動方案。

時間序列處理：上採樣



上採樣的原因：

- 匹配不同時間序列
- 不規則的時間序列
- 領域知識

下採樣的反方向是上採樣。上採樣的問題在於，大多數情況下很難實現。假設您希望將銷售資料從每日銷售額上採樣至每小時銷售額。除非有其他資料來源可供參考，否則您將無法從每日銷售額轉換為每小時銷售額。在某些情況下，必須使用其他資料或知識。例如，如果必須匹配另一時間序列的頻率，則可能會用到不規則的時間序列或特定的領域知識。在這些情況下，必須注意如何轉換資料。對於零售示例，最佳做法是在當天指定時間創建一個訂單。對於溫度，您可以將每日溫度複製到每個小時時段中，或者使用某種公式來計算曲線。

時間序列處理：平滑數據



為什麼要平滑？

- 數據準備
- 可視化

平滑對結果有何影響？

- 使用更為乾淨的資料建模
- 模型相容性
- 生產改進

在資料科學領域，異常值可能是一個難題。時間序列資料也是如此。

如果您在查看銷售資料時，發現一個訂單中的商品數量異常多，您可能不希望在預測計算中包含該訂單。訂單大小可能永遠不會重複。消除這些異常值和反常情況的過程稱為平滑。

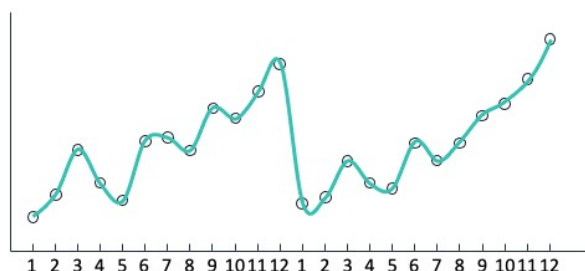
平滑資料有助於處理異常值和其他反常情況。出於以下原因，您可能會考慮使用平滑。

- 資料準備 – 刪除誤差值和異常值
- 可視化 – 降低圖中的雜訊

瞭解為什麼要平滑資料及其可能產生的影響。您想要的結果可能是降低雜訊並創建更好的模型。但是，考慮以下問題也同樣重要：平滑是否會損害模型？模型是否會出現雜訊資料？是否也可以在生產環境中平滑資料？

季節性

- 季節性頻率
 - 每小時、每天、每季度、每年
 - 春季、夏季、秋季、冬季
 - 重大節日銷售、寒假
- 合併假期



資料季節性是指以穩定觀察頻率進行的任何形式的重複觀察。例如，在銷售領域，您通常會看到季度末和第 4 季度的銷售額更高。在消費零售業，這種模式在第 4 季度出現的更多。請注意，同一資料集中的資料可能會呈現多種季節性。

很多時候，您可能希望將季節性資訊納入預測中。對於銷售來說，本地假期是一個很好的例子。

時間序列相關性



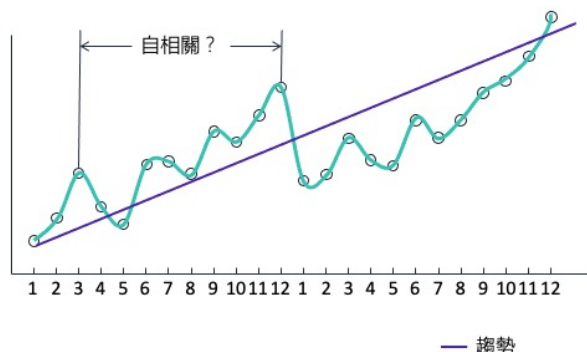
相關性並不表示因果關係。

在解釋自己的資料時要小心，而對於相關性，要謹慎對待，您一定不希望對沒有現實意義的相關性採取行動。作為實驗，假如您生成兩個數字介於 0 和 1 之間的隨機時間序列資料集。您會發現它們的相關性較低。但如果對兩個資料集引入相同的斜率，則會發現其相關性較強。

有關相關性的更多資訊，請參閱[假性相關](#)。許多相關性都繪製出來，基本沒有什麼意義。

平穩性、趨勢和自相關

- 平穩性
 - 系統的穩定性如何？
 - 過去能否預示未來？
- 趨勢
 - 相關性問題
- 自相關
 - 時間點如何呈線性相關
- 影響演算法選擇



瞭解系統的穩定性非常重要。穩定性或平穩性可以說明系統過去的行為對未來的行為影響有多大。穩定性較差的系統無法有效預測未來。

通常，您希望確定時間序列的趨勢。但根據趨勢調整序列可能會導致難以將序列與針對趨勢調整的其他序列進行比較。趨勢可能會主導序列的值，這會導致您高估兩個序列之間的相關性。這一現象在上一主題中有所體現。有關發生這種情況的原因的詳細說明，請參閱[在時間序列分析中避免常見錯誤](#)。

自相關是您所面臨的時間序列資料問題中比較特殊的一個。正如您在其他機器學習問題中看到的，構建 ML 模型的目的是將信號與雜訊分開。自相關是雜訊的一種形式，因為單獨的觀察結果並不相互獨立。

含有自相關的時間序列可能會誇大生成的模型的準確性。您在本模組中看到的一些演算法有助於校正自相關。

這些因素以及季節性會影響您選擇用於生成預測的模型。一些演算法可以處理季節性和自相關問題，而有些演算法則不能。

使用 Pandas 獲取時間序列資料



- 時間感知索引

```
dataframe['2010-01-04']
```

```
dataframe['2010-02':'2010-03']
```

```
dataframe['weekday_name'] = dataframe.index.weekday_name
```

- GroupBy 和重採樣操作

- `dataframe.groupby('StockCode')`

```
dataframe.groupby('StockCode').resample('D').sum()
```

- 自相關

```
dataframe['Quantity'].autocorr()
```

Pandas 庫在開發時考慮了財務資料分析。因此，它擅長處理時間序列資料。

您可以將 Pandas DataFrame 的索引設置為 *datetime*，這樣就能使用日期和時間來選擇資料。您可以使用包含部分日期的範圍。也可以提取日期部分，比如 *year*、*month*、*weekday_name* 等。

對於分組和重採樣任務，Pandas 具有內置的函數來完成這兩項任務。

最後，Pandas 讓您可以深入瞭解自相關。

有關 Pandas 和時間序列資料的更多資訊，請參閱[時間序列/日期功能](#)在 Pandas 文檔中...

時間序列演算法

- 自回歸積分移動平均模型 (ARIMA)
- DeepAR+
- 指數平滑法 (ETS)
- 非參數時間序列 (NPTS)
- Prophet



構建預測應用程式的其中一個任務是選擇適當的演算法。您所使用的資料集類型以及該資料集的特徵決定了您選擇何種演算法。

Amazon Forecast 支援以下 5 種演算法。

- 自回歸積分移動平均模型 (ARIMA) – 此演算法可消除可能會影響觀察結果模式的自相關。
- DeepAR+ – 一種用於預測一維時間序列的監督式學習演算法。它使用遞迴神經網路在多個時間序列上訓練模型。
- 指數平滑法 (ETS)：該演算法適用於具有季節性的資料集。它對所有觀察結果使用加權平均值。權重會隨著時間的推移而減小。
- 非參數時間序列 (NPTS) – 預測基於對過去觀察結果的抽樣。具有專用於季節性資料集和氣候資料集的版本。
- Prophet – 貝葉斯時間序列模型。該模型適用於時間跨度較大、具有缺失資料或較大異常值的資料集。

第 2 節要點



- 時間序列資料為序列化資料
- 時間挑戰 –
 - 不同的格式
 - 缺失數據
 - 季節性
 - 相關性
- Pandas 庫提供對時間序列資料的支援
- 借助 Amazon Forecast ,
您可以從 5 種演算法中進行選擇

本模組中這部分內容的要點包括：

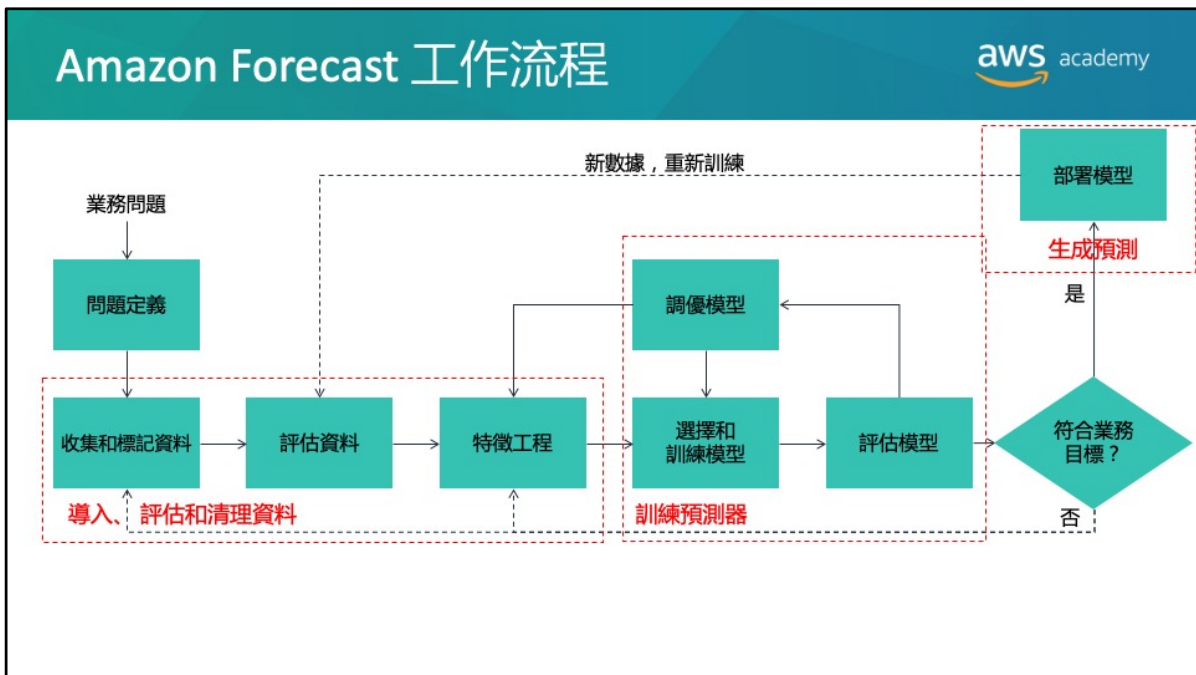
- 時間序列資料是包含時間因素的序列化資料，這使其不同於常規資料集
- 一些時間方面的挑戰包括 –
 - 處理不同的時間格式
 - 通過下採樣、上採樣和平滑處理缺失資料
 - 處理季節性，如工作日和年度週期
 - 避免不良相關
- Pandas 庫通過處理時間的函數提供對時間序列資料的支援
- 借助 Amazon Forecast，您可以從 5 種演算法中進行選擇 –
 - ARIMA
 - DeepAR+
 - ETS
 - NPTS
 - Prophet

模組 4：預測簡介

第 3 節：使用 Amazon Forecast



介紹 “第 3 節：使用 Amazon Forecast”



生成預測時，您可以應用在本課程中所用的機器學習開發管道。

- 導入資料 – 您必須盡可能多地導入擁有的資料，包括歷史資料和相關資料。在使用資料訓練模型之前，應該先進行一些基本評估和特徵工程。
- 訓練預測器 – 要訓練預測器，必須選擇一種演算法。如果您不確定哪種演算法最適合您的資料，可以選擇 *AutoML* 作為您的演算法，由 Amazon Forecast 來選擇。您還必須為資料選擇一個領域，但如果不確定哪個領域最適合，可以選擇自訂領域。領域具有其所需的特定資料類型。有關更多資訊，請參閱 Amazon Forecast 文檔中的預定義資料集領域和資料集類型。
- 生成預測 – 在擁有訓練良好的模型之後，就可以通過輸入資料集組來使用模型進行預測。生成預測後，您可以查詢預測，也可以將其匯出到 Amazon Simple Storage Service (Amazon S3) 存儲桶。您還可以選擇在匯出前對預測中的資料進行加密。

Amazon Forecast 概覽



使用 Amazon Forecast 的總體過程是導入歷史資料和相關資料。Amazon Forecast 將檢查資料、識別關鍵資料並選擇適當的演算法。它使用該演算法對自訂模型進行訓練和優化，並生成預測器。您可以通過將預測器應用於資料集來創建預測。然後，您可以在 AWS 控制台中檢索這些預測，也可以將這些預測匯出為以逗號分隔的檔。您還可以使用應用程式設計發展介面 (API) 和命令列介面 (CLI) 命令來創建和檢索預測。

支持的領域



- 零售
- 庫存計畫
- Amazon EC2 容量
- 勞動力
- Web 流量
- 指標
- 自定義



在使用 Amazon Forecast 時，應選擇適當的領域。您可以從以下列表中選擇：

- 零售 – 產品需求
- 庫存計畫 – 原材料要求
- EC2 容量 – Amazon Elastic Compute Cloud (Amazon EC2) 的容量需求
- 勞動力 – 工作負載預測
- Web 流量 – 預計到達一個或多個網站的流量
- 指標 – 預測指標，如收入、銷售額或現金流
- 自定義 – 無法映射到先前某個領域的領域的預測

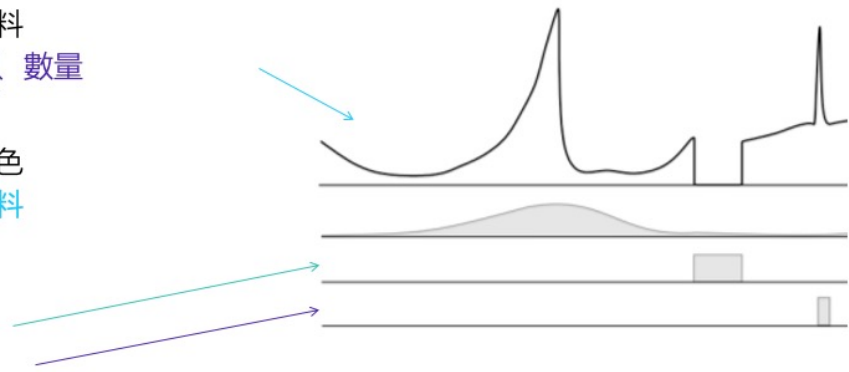
選擇領域之後，可以提高預測器的效率。每個領域都具有構建預測器時所提供的特定資料類型。例如，零售領域包含商品識別字、觀察結果的時間戳記、商品的銷量以及指定的時間戳記。

有關支援的領域的更多資訊，請參閱[預定義資料集域和資料集類型](#)主題

零售預測示例



- 時間序列資料
 - 事務性銷售資料
 - 時間戳、項目、數量
- 元數據
 - 類別、專案顏色
 - 項目、中繼資料
- 相關數據
 - 時間序列
 - 存貨資料
 - 促銷數據
 - 時間戳、項目、價格



以下示例顯示了零售需求預測所需的資料。

對於時間序列，必須提供：

- 時間戳記 – 事務發生的時間，最好採用 UTC 格式
- 商品 – 商品的商品 ID
- 數量 – 出售的商品數量

例如，商品的中繼資料可能包含類別或商品顏色。返回時間序列資料的連結僅包含商品 ID，因為商品的中繼資料通常不會變更。

銷售價格或其他促銷資料是可以創建更實用的預測的相關資料示例。要將此連結回商品，必須包含時間戳記和商品 ID。

- 時間序列資料
 - 網頁 ID
 - 每月的頁面流覽量
 - 時間戳
- 相關和中繼資料
 - 頁面類別
 - 地理識別字



以下示例顯示了 Web 流量預測所需的資料類型。

對於時間序列，必須提供：

- 網頁 ID
- 每月的頁面流覽量
- 時間戳

創建更實用的預測的相關資料包括：

- 頁面類別（如導航或內容類別別）
- Web 用戶端的地理識別字

您可能還需要以下中繼資料：

- 區域
- 促銷信息

選擇 Amazon Forecast 演算法



您可以從以下演算法清單中選擇：

- ARIMA
- DeepAR+
- ETS
- NPTS
- Prophet



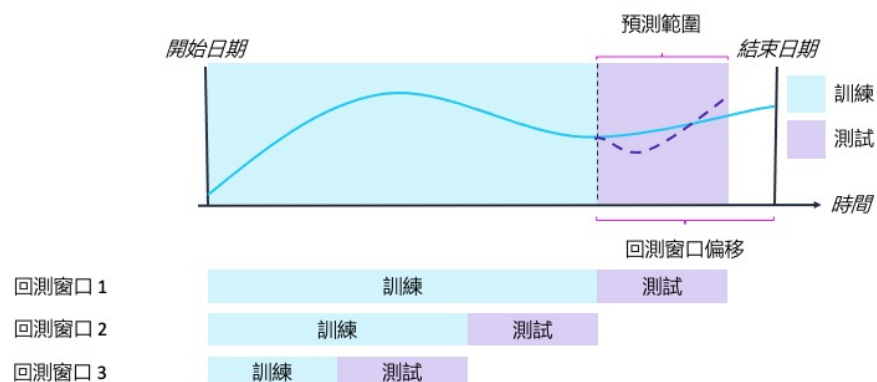
Amazon Forecast 預測器使用演算法來訓練模型。然後，通過輸入資料集組來使用模型進行預測。為幫助您上手使用，Amazon Forecast 提供了以下預定義演算法：

- ARIMA
- DeepAR+
- ETS
- NPTS
- Prophet

您還可以使用 *AutoML* 功能，該功能會嘗試所有演算法以查看哪種演算法最適合預測資料。

有關 Amazon Forecast 演算法的更多資訊，請參閱 AWS 文檔中的[選擇 Amazon Forecast 演算法](#)主題。

預測器準確性指標基於回測。



在準備資料進行機器學習訓練時，通常要保留資料，以便在驗證和評價模型時使用。您保留的資料通常是可用資料的隨機樣本。對於時間序列資料，由於時間之間的相關性，必須以不同的方式處理資料。

導入資料時，Amazon Forecast 會將其分為訓練和測試資料集，如圖所示。訓練資料用於訓練模型，然後針對保留的資料進行測試。您可以指定多個回測窗口，這將多次劃分資料、訓練模型並使用指標來確定哪種模型的效果最佳。預設回測窗口為 1。

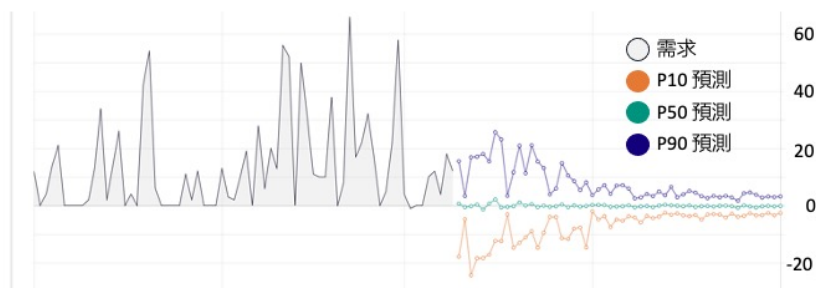
您可以在創建預測器時通過設置 `BackTestWindowOffset` 參數來更改 Amazon Forecast 劃分資料的方式。如果您未設置此值，則演算法將使用預設值。

訓練模型後，您需要衡量其準確性，接下來您將學習這部分內容。

評估指標：wQuantileLoss



- 為 10%、50% 和 90% 分位數分位數
- wQuantileLoss 是一組中每個分位數的平均誤差 –
 - 最適合誤差變化較大的模型



第一個 Amazon Forecast 評估指標是加權分位數損失 (wQuantileLoss)。當 Amazon Forecast 創建預測時，它將提供三個不同分位數的概率預測 – 10%、50% 和 90%。這些預測分位元數顯示了與每種預測關聯的不確定性。

P10 分位數預測，在 10% 的時間內，true 值將低於預測值。例如，假設您是零售商。您想預測冬季手套的產品需求，這些手套只在秋季和冬季暢銷。假設您沒有足夠的存儲空間，並且投資資本的成本較高，或者擔心冬季手套的庫存積壓。那麼，您可以使用 P10 分位數來訂購數量相對較少的冬季手套。您知道，P10 預測只有 10% 的時間會高估您的冬季手套的需求量，所以 90% 的時間冬季手套會賣光。

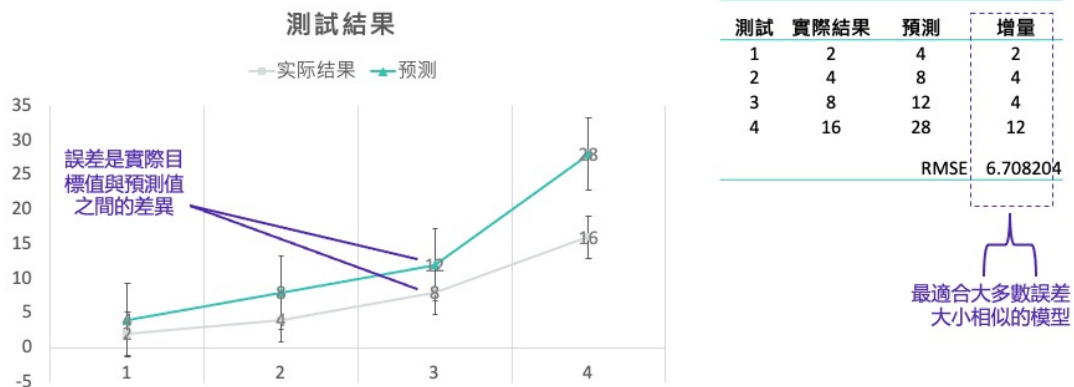
P50 分位數預測，在 50% 的時間內，true 值將低於預測值。繼續以冬季手套為例，假設您知道有適度的手套需求量，並且不會擔心存貨積壓。那麼，您可以選擇使用 P50 分位數訂購手套。

P90 分位數預測，在 90% 的時間內，true 值將低於預測值。假設您確定手套庫存不足將導致巨額收入損失，例如不銷售手套的成本極高或投資資本的成本較低。那麼，在這種情況下，您可以選擇使用 P90 分位數訂購手套。

Amazon Forecast 還會計算每個分位數的相關損失（誤差）。加權分位數損失 (wQuantileLoss) 計算某個分位數的預測與任一方向的實際需求的差距。wQuantileLoss 指標越低意味著模型的預測越可靠。

均方根誤差 (RMSE)

RMSE 是指誤差的平方。



均方根誤差 (RMSE) 是評估預測可靠性的另一種方法。與 wQuantileLoss 類似，RMSE 會計算預測值與實際測試資料的差距。

RMSE 會查找資料集中的實際目標值與該時間段的預測值之間的差異，然後求差異的平方。本例展示了如何計算 RMSE。RMSE 值表示預測誤差的標準差。當誤差的大小大致相同時（即沒有過多的異常值），該測試有助於提高預測有效性。RMSE 指標越低意味著模型的預測越可靠。

模型準確性示例



網路零售鞋商想要預測其無法履行
AnyCompany 品牌鞋子訂單的頻率。

Amazon Forecast 預測每月需求量为 1,000 雙

- P10：10% 的時間，訂購 880 雙以下
- P50：50% 的時間，訂購 1,050 雙以下
- P90：90% 的時間，訂購 1,200 雙以下



P10 = 880	預測 = 1000
P50 = 1050	
P90 = 1200	

本示例顯示了網路零售商如何使用準確性指標來評估預測。該零售商想要預測某品牌鞋子的銷售需求。他們將該品牌的銷售記錄輸入 Amazon Forecast 來創建預測器。

預測器根據 P10、P50 和 P90 值預測的需求為 1000 雙。wQuantileLoss 值指示在 10% 的時間內 (P10)，銷量低於 880 雙。然後，在 50% 的時間內 (P50)，銷量低於 1050 雙。最後，在 90% 的時間內 (P90)，銷量低於 1200 雙。然後，零售商可以根據這些值來確定要維持的庫存水準。最終決定基於他們對無法履行訂單或存貨過剩的風險評估。

演示：使用 Amazon Forecast 創建預測



您的講師現在將演示如何使用 Amazon Forecast 來創建預測，或者為您提供演示錄影的存取權限。

第 3 節要點



- 您可以將 Amazon Forecast 用於時間序列資料
- 架構特定於領域
- 數據可以包括 –
 - 時間序列資料
 - 元數據
 - 相關數據
- 考慮時間問題，將資料分為訓練和測試資料
- 使用 RMSE 和 wQuantileLoss 指標評估模型

本模組中這部分內容的要點包括：

- 您可以使用 Amazon Forecast 來訓練和使用時間序列資料模型
- 有針對零售和 EC2 容量規劃等領域定義的特定架構，或者您也可以使用自訂架構
- 您至少需要提供時間序列資料，也可以提供中繼資料和相關資料以將移動資訊添加到模型中
- 與大多數監督式機器學習問題一樣，您的資料也分為訓練和測試資料，但這種劃分考慮了時間元素
- 使用 RMSE 和 wQuantileLoss 指標來評估模型的效率

模組 4 – 引導式 實驗：使用 Amazon Forecast 創建預測

aws academy



現在，您要完成“模組 4 – 引導式實驗：使用 Amazon Forecast 創建預測”。

模組 4：預測簡介

模組總結



現在，我們來回顧和總結一下本模組，然後進行知識測驗。

模組總結



總體來說，您在本模組中學習了如何：

- 描述使用 Amazon Forecast 可解決的業務問題
- 描述使用時間序列資料時面對的挑戰
- 列出使用 Amazon Forecast 創建預測所需的步驟
- 使用 Amazon Forecast 進行預測

總體來說，您在本模組中學習了如何：

- 描述使用 Amazon Forecast 可解決的業務問題
- 描述使用時間序列資料時面對的挑戰
- 列出使用 Amazon Forecast 創建預測所需的步驟
- 使用 Amazon Forecast 進行預測



現在可以完成本模組的知識測驗。

其他資源



- [Amazon Forecast 文檔](#)
- [Amazon Forecast 產品頁面](#)
- [如何不使用機器學習進行時間序列預測](#)
- [時間序列預測原理 Amazon Forecast 白皮書](#)

如果您想瞭解有關本模組所涵蓋主題的更多資訊，下面這些其他資源可能會有所幫助：

- [Amazon Forecast 文檔](#)
- [Amazon Forecast 產品頁面](#)
- [如何不使用機器學習進行時間序列預測](#)
- [時間序列預測原理 Amazon Forecast 白皮書](#)

謝謝

© 2020 Amazon Web Services, Inc. 或其附屬公司。保留所有權利。未經 Amazon Web Services, Inc. 事先書面許可，不得複製或轉載本文的部分或全部內容。禁止因商業目的複製、出借或出售本文。如有對本課程的糾正或回饋意見，請發送電子郵件至：aws-course-feedback@amazon.com。如有其他任何問題，請與我們聯繫：<https://aws.amazon.com/contact-us/aws-training/>。所有商標均為各自所有者的財產。



感謝您的參與！