



このコンテンツは公開から3年以上経過しており内容が古い可能性があります
最新情報については[サービス別資料](#)もしくはサービスのドキュメントをご確認ください

[AWS Black Belt Online Seminar]

Amazon Forecast

サービスカットシリーズ

Partner Solutions Architect

小田桐 昂史

2020/01/28

AWS 公式 Webinar

<https://amzn.to/JPWebinar>



過去資料

<https://amzn.to/JPArchive>



自己紹介

小田桐 昂史 (おだぎり こうし)

所属

パートナーソリューションアーキテクト

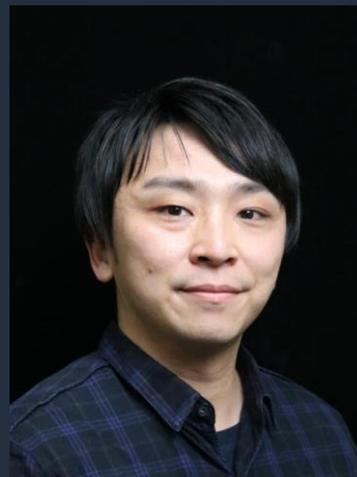
好きな AWS のサービス



Amazon Forecast



Amazon Sumerian



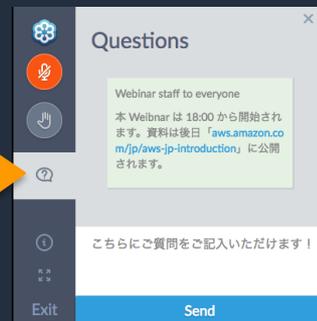
AWS Black Belt Online Seminar とは

「サービス別」「ソリューション別」「業種別」のそれぞれのテーマに分かれて、アマゾンウェブ サービス ジャパン株式会社が主催するオンラインセミナーシリーズです。

質問を投げることができます！

- 書き込んだ質問は、主催者にしか見えません
- 今後のロードマップに関するご質問は
お答えできませんのでご了承下さい

- ① 吹き出しをクリック
- ② 質問を入力
- ③ Sendをクリック



Twitter ハッシュタグは以下をご利用ください
#awsblackbelt

内容についての注意点

- 本資料では2020年1月28日時点のサービス内容および価格についてご説明しています。最新の情報はAWS公式ウェブサイト(<http://aws.amazon.com>)にてご確認ください。
- 資料作成には十分注意しておりますが、資料内の価格とAWS公式ウェブサイト記載の価格に相違があった場合、AWS公式ウェブサイトの価格を優先とさせていただきます。
- 価格は税抜表記となっております。日本居住者のお客様には別途消費税をご請求させていただきます。
- AWS does not offer binding price quotes. AWS pricing is publicly available and is subject to change in accordance with the AWS Customer Agreement available at <http://aws.amazon.com/agreement/>. Any pricing information included in this document is provided only as an estimate of usage charges for AWS services based on certain information that you have provided. Monthly charges will be based on your actual use of AWS services, and may vary from the estimates provided.

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ



Amazon Forecast

Amazon.com と同じテクノロジーに基づいた、機械学習の経験が不要な
時系列予測サービス

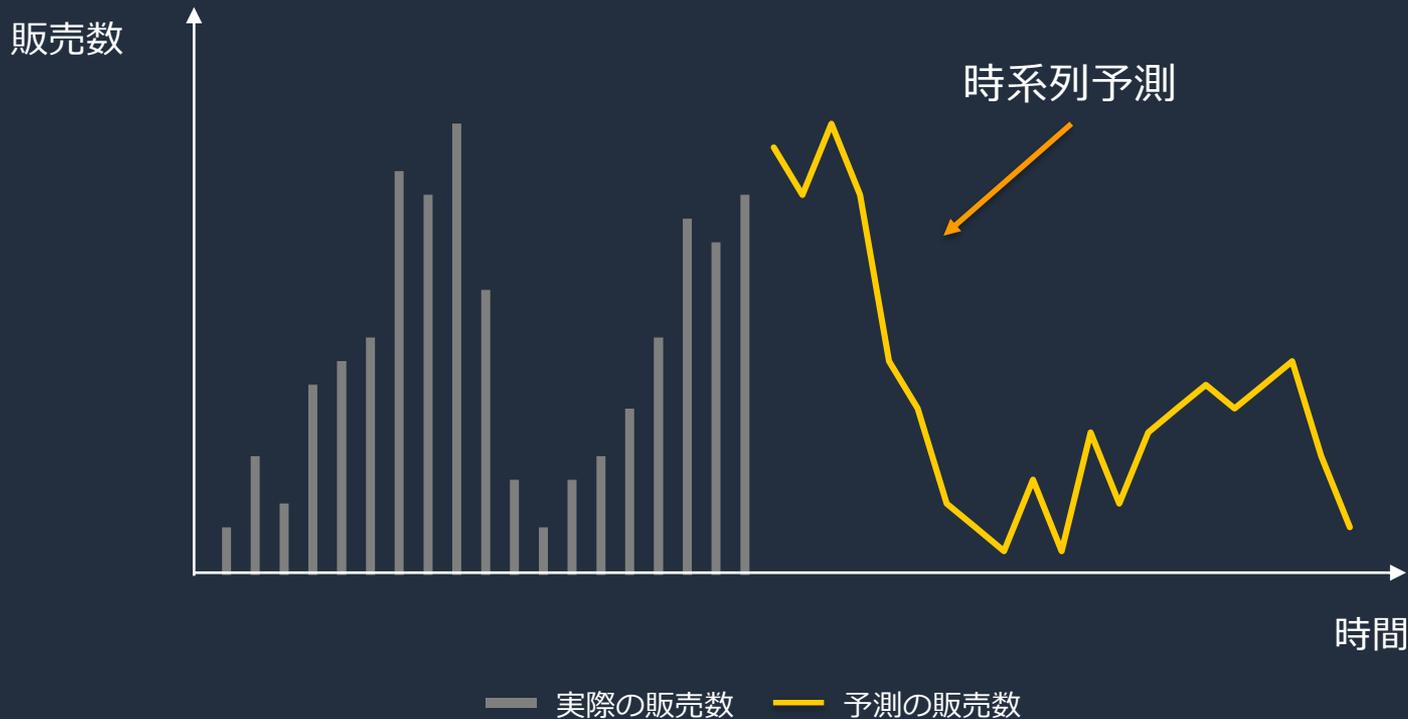
時系列予測とは

販売数の遷移



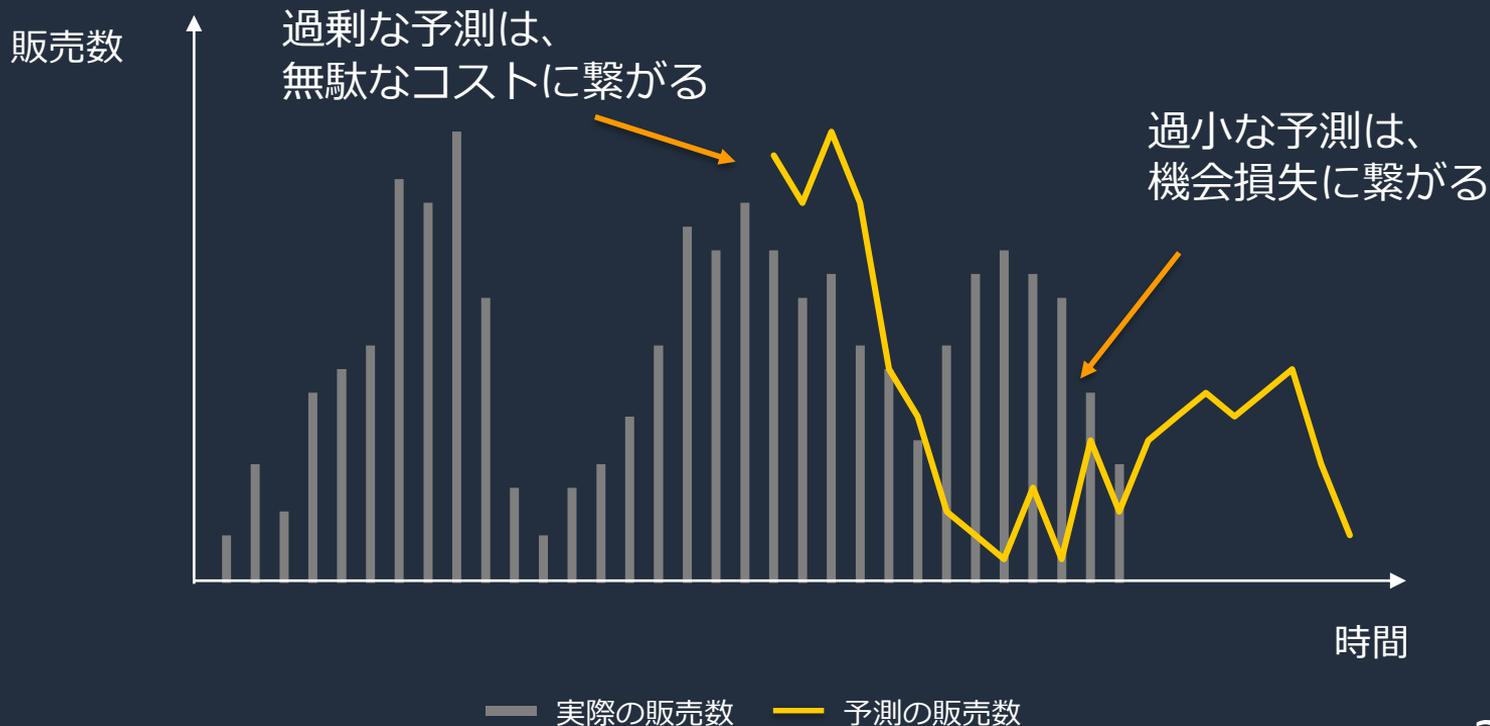
時系列予測とは

販売数の遷移



時系列予測とは

販売数の遷移



時系列予測における従来の課題



季節によって特徴的なデータ傾向（季節性）を扱えない



商品の特徴といったメタデータの扱いができない



価格・祝日・プロモーションといった、予測精度に影響を与える関連データを組み込めない



過去のデータが存在しない、新製品に関する予測が困難

Amazon Forecast



Amazon.com における予測技術に関する20年の経験



5種類のアлゴリズムを提供しており、そのうちの1つの DeepAR+ は深層学習を利用



1,000以上の時系列データにおいて、深層学習によって精度が最大50%改善

Amazon Forecast

Amazon Forecast

過去データ

Timestamp	Demand	itemId
2020/01/01 10:00	2	client_0
2020/01/01 11:00	23	client_1
2020/01/01 12:00	0	client_2
2020/01/01 13:00	100	client_3
2020/01/01 14:00	75	client_4
2020/01/01 15:00	200	client_5
2020/01/01 16:00	6	client_6
2020/01/01 17:00	7	client_7
2020/01/01 18:00	50	client_8
2020/01/01 19:00	60	client_9



データの読み込み
データの解析
特徴判定
アルゴリズム
ハイパーパラメータの選択
モデル学習
モデル最適化
モデルのホスト

予測値



フルマネージドで提供

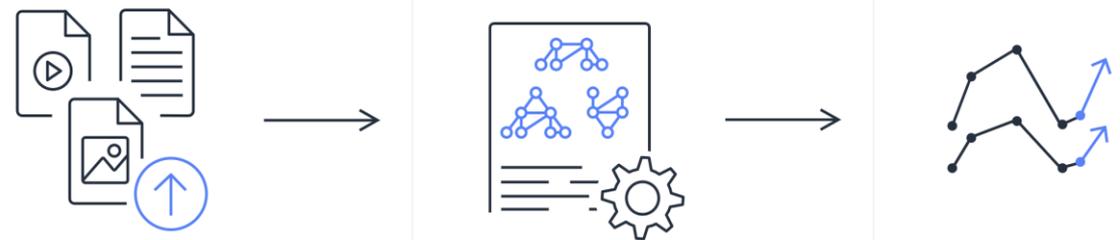
Amazon Forecast

Amazon Forecast ×

Amazon Forecast > Dataset groups > util_power_forecastdemo_dsg > Dashboard

Dashboard Info

Overview



Import your data
Datasets are required to train predictors, which are then used to generate forecasts.

[View datasets](#)

Train a predictor
Train a predictor, a custom model with underlying infrastructure that Amazon Forecast trains on your datasets.

[View predictors](#)

[Train predictor](#)

Generate forecasts
Generate forecasts using your trained predictors.

[Lookup forecast](#)

[Create a forecast](#)

Dataset groups

- util_power_forecastdemo_dsg
 - Dashboard**
 - Datasets
 - Predictors
 - Forecasts
 - Forecast lookup

ユースケース事例



Product Demand Planning

- ウェブサイトや特定の店舗、ロケーションで販売されている製品に対する需要予測
- サプライチェーンの需要予測



Financial planning

- 地域毎やサービス毎のキャッシュフロー、売上、利益、経費の予測



Resource planning

- コールセンターのオペレータ数、契約社員、ITスタッフ、需要を満たすのに必要なエネルギー
- サーバーへのウェブトラフィック、AWS の使用状況、IoT センサーの使用状況

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

Amazon Forecast のワークフロー概要 **全体像**

手順	説明
データセットグループ (Dataset Groups)/データセット (Datasets) の準備	予測を行うためのデータセット・データセット群。予測対象の時系列データ、マッピング定義であるスキーマを用意し、Amazon Forecast に読み込み、トレーニング（学習）の準備を行う。データセットは、メインの時系列データ、関連する時系列データ、メタデータの3種類からなる。各種データセットはデータセットグループでまとめることができる。
予測子 (Predictors) の作成	読み込んだデータに対して、アルゴリズムや予測する期間を設定し、トレーニングを行い、予測子（モデル）を作成する。また、予測精度に対する評価も行う。
予測 (Forecasts) の作成	トレーニングされたモデルのエンドポイントを作成し、時系列データの予測を生成する。

Amazon Forecast のワークフロー概要



Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備



Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備

③ スキーマの定義

```
1 "schema": {
2   "dimensions": [
3     {
4       "name": "Time",
5       "type": "Time",
6       "granularity": "Hourly",
7       "period": "P1H",
8       "start": "2014-01-01T00:00:00",
9       "end": "2014-01-01T00:00:00"
10    },
11    {
12      "name": "Client",
13      "type": "Categorical",
14      "values": ["client_0", "client_1"]
15    }
16  ],
17  "target": "Usage"
18 }
```

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備

③ スキーマの定義

```
1 "schema": {
2   "dimensions": [
3     {
4       "name": "Time",
5       "type": "Time",
6       "granularity": "Hourly",
7       "start": "2014-01-01T00:00:00",
8       "end": "2014-01-01T00:00:00"
9     },
10    {
11      "name": "Client",
12      "type": "Categorical",
13      "values": ["client_0", "client_1"]
14    }
15  ],
16  "target": "Forecast"
17 }
```

④ データセットの作成

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備

③ スキーマの定義

```
{
  "schema": {
    "name": "client_0",
    "type": "ITEM_METADATA",
    "description": "client_0"
  },
  "schema": {
    "name": "client_1",
    "type": "ITEM_METADATA",
    "description": "client_1"
  },
  "schema": {
    "name": "client_0",
    "type": "RELATED_TIME_SERIES",
    "description": "client_0"
  }
}
```

④ データセットの作成

⑤ 各種データのインポート

- TARGET_TIME_SERIES
- ITEM_METADATA
- RELATED_TIME_SERIES

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備

予測子の作成

③ スキーマの定義

```
{
  "schema": {
    "name": "client_0",
    "description": "client_0",
    "properties": {
      "time": {
        "type": "string",
        "format": "date-time"
      },
      "value": {
        "type": "float"
      },
      "client_id": {
        "type": "string"
      }
    }
  }
}
```

④ データセットの作成

⑤ 各種データのインポート

- TARGET_TIME_SERIES
- ITEM_METADATA
- RELATED_TIME_SERIES

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備

③ スキーマの定義

```
{
  "name": "client",
  "type": "string",
  "required": true,
  "description": "Client ID"
},
{
  "name": "target",
  "type": "float",
  "required": true,
  "description": "Target value"
},
{
  "name": "related_time_series",
  "type": "string",
  "required": false,
  "description": "Related time series name"
}
```

④ データセットの作成

⑤ 各種データのインポート

- TARGET_TIME_SERIES
- ITEM_METADATA
- RELATED_TIME_SERIES

予測子の作成

⑥ 予測子の作成

- 予測期間の設定
- アルゴリズムの選択
- パラメータの設定

いわゆるトレーニングが実行され、モデルが作成される

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備

③ スキーマの定義

```
{
  "name": "client",
  "type": "string",
  "required": true,
  "description": "Client ID"
},
{
  "name": "target",
  "type": "float",
  "required": true,
  "description": "Target value"
},
{
  "name": "related_time_series",
  "type": "string",
  "required": false,
  "description": "Related time series name"
}
```

④ データセットの作成

⑤ 各種データのインポート

- TARGET_TIME_SERIES
- ITEM_METADATA
- RELATED_TIME_SERIES

予測子の作成

⑥ 予測子の作成

- 予測期間の設定
- アルゴリズムの選択
- パラメータの設定

いわゆるトレーニングが実行され、モデルが作成される

⑦ 予測子の評価

各指標を参考に評価を実施

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備

③ スキーマの定義

```
{
  "name": "client",
  "type": "string",
  "required": true,
  "description": "Client ID"
},
{
  "name": "target",
  "type": "float",
  "required": true,
  "description": "Target value"
},
{
  "name": "related",
  "type": "float",
  "required": false,
  "description": "Related value"
}
```

④ データセットの作成

⑤ 各種データのインポート

- TARGET_TIME_SERIES
- ITEM_METADATA
- RELATED_TIME_SERIES

予測子の作成

⑥ 予測子の作成

- 予測期間の設定
- アルゴリズムの選択
- パラメータの設定

いわゆるトレーニングが実行され、モデルが作成される

⑦ 予測子の評価

各指標を参考に評価を実施

予測の作成

Amazon Forecast のワークフロー概要

Amazon Forecast



Amazon S3

① 各種CSVファイルの準備

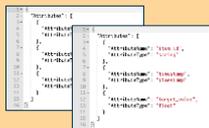
```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

② データセットグループの作成

データセットグループ

データの準備

③ スキーマの定義



④ データセットの作成

⑤ 各種データのインポート

- TARGET_TIME_SERIES
- ITEM_METADATA
- RELATED_TIME_SERIES

予測子の作成

⑥ 予測子の作成

- 予測期間の設定
- アルゴリズムの選択
- パラメータの設定

いわゆるトレーニングが実行され、モデルが作成される

⑦ 予測子の評価

各指標を参考に評価を実施

予測の作成

⑧ 予測の作成

学習済みモデルを利用し、予測（推論エンドポイント）を作成する

予測はAPI経由での取得、または、マネジメントコンソール上で確認することができる

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

データの準備 概要

- CSVファイルを格納する S3 バケットの作成
- CSVファイルの格納
- データセットグループの作成
- データセットの作成
 - スキーマの定義
 - データのインポート

```
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 2.53807106598985, client_0
2014-01-01 01:00:00, 23.648648648648624, client_1
2014-01-01 02:00:00, 9.648648648612345, client_0
```

CSVファイル

アップロード



Amazon S3

スキーマ定義
(属性のマッピング情報)

```
{
  "attributes": [
    {
      "attributeName": "timestamp",
      "attributeType": "timestamp"
    },
    {
      "attributeName": "item_id",
      "attributeType": "string"
    }
  ]
}
```

データセット読み込み



Amazon Forecast

データの準備 データセット

- 時系列データまたはメタデータ
- カンマ区切り値 (CSV) ファイル形式
 - タイムスタンプは yyyy-MM-dd (例 : 2020-01-28) と、オプションで HH:mm:ss 形式 (例 : 2020-01-28 11:22:33) をサポート
- JSON 形式で定義されるスキーマを用意する
 - データタイプとして integer, float, string, timestamp をサポート
- データセットには、データセットタイプ、データセットドメインを関連づけ、これらによりデータセットに含めるフィールドとスキーマが決定する

データの準備 データセットタイプ

- 予測子をトレーニングするには、TARGET_TIME_SERIES とオプションで、RELATED_TIME_SERIES、ITEM_METADATAを作成し、データセットグループに追加する
 - TARGET_TIME_SERIES データセットは必須
- 作成したデータセット毎に、データセットタイプ、データセットドメインを関連づける

データセットタイプ	用途	説明
TARGET_TIME_SERIES	全てのデータセットドメインで必要	予測したいデータに関する過去の時系列データ 例：売上データ
RELATED_TIME_SERIES	補完的な時系列データを利用するデータセットドメインで利用可能	予測したいデータに関連する補完的な時系列データ 例：価格、キャンペーンデータ、天気
ITEM_METADATA	メタデータ情報を利用するデータセットドメインで利用可能	予測したいデータに関連する静的なメタデータ 例：商品カテゴリ（色、サイズ、ブランド）

データの準備 データセットドメイン

- データセットグループ毎にデータセットドメインを定義する
- データセットドメインは、予測するユースケースの定義
- データセットドメイン毎に必要なデータセットタイプは異なる

Create dataset group Info

Dataset groups are containers for all your datasets.

Dataset group details

Dataset group name
The name can help you distinguish this dataset group from other dataset groups on the dataset groups dashboard.

The dataset group name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and

Forecasting domain Info
A forecasting domain defines a forecasting use case. You can choose a predefined domain, or you can create your own domain.

Retail
This is a predefined domain for forecasting demand for a retailer.

Custom
Choose this domain if none of the other domains are applicable to your forecasting needs.

Inventory planning
Forecast demand for raw materials and determine how much inventory of a particular item to stock.

Ec2 capacity
Forecast your Amazon Elastic Compute Cloud (Amazon EC2) capacity.

Work force
Use this domain to plan and identify the amount of work force you require.

Web traffic
Forecast web traffic to a web property or a set of web properties.

Metrics
This domain is for forecasting metrics such as revenue, sales, and cashflow.

Cancel **Next**

データの準備 データセットドメイン

データセットドメイン	説明
RETAIL ドメイン	小売の需要予測
INVENTORY_PLANNING ドメイン	サプライチェーンとインベントリの計画
EC2_CAPACITY ドメイン	Amazon Elastic Compute Cloud (Amazon EC2) のキャパシティの予測
WORK_FORCE ドメイン	従業員の計画
WEB_TRAFFIC ドメイン	ウェブトラフィックの見積もり
METRICS ドメイン	収益およびキャッシュフローなどの予測メトリクス
CUSTOM ドメイン	その他すべての時系列予測のタイプ

https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/howitworks-domains-ds-types.html

データの準備 データセットドメイン

RETAIL ドメインの場合

TARGET_TIME_SERIES

フィールド	説明
item_id *	商品の固有ID
timestamp *	時間
demand *	販売数
location	販売された場所

RELATED_TIME_SERIES

フィールド	説明
item_id *	商品の固有ID
timestamp *	時間
price	timestamp時の商品価格
webpage_hits	timestamp時のウェブヒット数
stockout_days	商品の在庫がなくなるまでの日数
inventory_onhand	在庫数
revenue	売上
in_stock	在庫があるかどうかのフラグ
promotion_applied	timestamp時にプロモーションがあったかどうか

ITEM_METADATA

フィールド	説明
item_id *	商品の固有ID
category	商品カテゴリー
brand	商品ブランド
color	商品の色
genre	商品のジャンル

* 必須項目

https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/retail-domain.html

データの準備 TARGET_TIME_SERIES

- すべてのドメインで必須
- 予測したいデータに関する過去の時系列データ
- ID, timestamp, 予測値のフィールド（ディメンション）は必須
- 予測フィールドを最大10個指定することができる
 - フィールドはロケーションやストア名など時間の経過とともに変化しないもの
 - フィールドごとの予測をすることが可能
 - オプションすべてのフィールドは string 型

データの準備 TARGET_TIME_SERIES

例 (RETAIL ドメインの場合)

- item_id (商品のID)
- timestamp (タイムスタンプ)
- demand (タイムスタンプ時の販売数)
- location (販売場所)

```
2019-01-01,socks,NYC,25
2019-01-05,socks,SFO,45
2019-02-01,shoes,ORD,10
2019-06-01,socks,NYC,100
2019-06-05,socks,SFO,5
2019-07-01,shoes,ORD,50
...
```

CSV

```
{
  "attributes": [
    {
      "attributeName": "timestamp",
      "attributeType": "timestamp"
    },
    {
      "attributeName": "item_id",
      "attributeType": "string"
    },
    {
      "attributeName": "location",
      "attributeType": "string"
    },
    {
      "attributeName": "demand",
      "attributeType": "float"
    }
  ]
}
```

スキーマ

データの準備 TARGET_TIME_SERIES

例 (RETAIL ドメインの場合)

- item_id (商品のID)
- timestamp (タイムスタンプ)
- demand (タイムスタンプ時の販売数)
- location (販売場所)

```
2019-01-01,socks,NYC,25
2019-01-05,socks,SFO,45
2019-02-01,shoes,ORD,10
2019-06-01,socks,NYC,100
2019-06-05,socks,SFO,5
2019-07-01,shoes,ORD,50
...
```

CSV

```
{
  "attributes": [
    {
      "attributeName": "timestamp",
      "attributeType": "timestamp"
    },
    {
      "attributeName": "item_id",
      "attributeType": "string"
    },
    {
      "attributeName": "location",
      "attributeType": "string"
    },
    {
      "attributeName": "demand",
      "attributeType": "float"
    }
  ]
}
```

スキーマ

データの準備 TARGET_TIME_SERIES

例 (RETAIL ドメインの場合)

- item_id (商品のID)
- timestamp (タイムスタンプ)
- demand (タイムスタンプ時の販売数)
- location (販売場所)

```
2019-01-01,socks,NYC,25
2019-01-05,socks,SFO,45
2019-02-01,shoes,ORD,10
2019-06-01,socks,NYC,100
2019-06-05,socks,SFO,5
2019-07-01,shoes,ORD,50
...
```

CSV

```
{
  "attributes": [
    {
      "attributeName": "timestamp",
      "attributeType": "timestamp"
    },
    {
      "attributeName": "item_id",
      "attributeType": "string"
    },
    {
      "attributeName": "location",
      "attributeType": "string"
    },
    {
      "attributeName": "demand",
      "attributeType": "float"
    }
  ]
}
```

スキーマ

データの準備 TARGET_TIME_SERIES

例 (RETAIL ドメインの場合)

- item_id (商品のID)
- timestamp (タイムスタンプ)
- demand (タイムスタンプ時の販売数)
- location (販売場所)

```
2019-01-01,socks,NYC,25
2019-01-05,socks,SFO,45
2019-02-01,shoes,ORD,10
2019-06-01,socks,NYC,100
2019-06-05,socks,SFO,5
2019-07-01,shoes,ORD,50
...
```

CSV

```
{
  "attributes": [
    {
      "attributeName": "timestamp",
      "attributeType": "timestamp"
    },
    {
      "attributeName": "item_id",
      "attributeType": "string"
    },
    {
      "attributeName": "location",
      "attributeType": "string"
    },
    {
      "attributeName": "demand",
      "attributeType": "float"
    }
  ]
}
```

スキーマ

データの準備 TARGET_TIME_SERIES

前処理 (Featurization)

- TARGET_TIME_SERIES のみ、学習の前に前処理をかけることができる
- 時間間隔の調整
 - 1時間おきに予測する場合、1時間総和、平均等も取得可能
- 欠損値の補完
 - 0埋めや nan (欠損のまま) を選ぶことができる

時間	データ
2018-03-03 01:00:00	100
2018-03-03 02:20:00	50
2018-03-03 02:45:00	20
2018-03-03 04:00:00	120

時間	データ
2018-03-03 01:00:00	100
2018-03-03 02:00:00	70
2018-03-03 03:00:00	空
2018-03-03 04:00:00	120

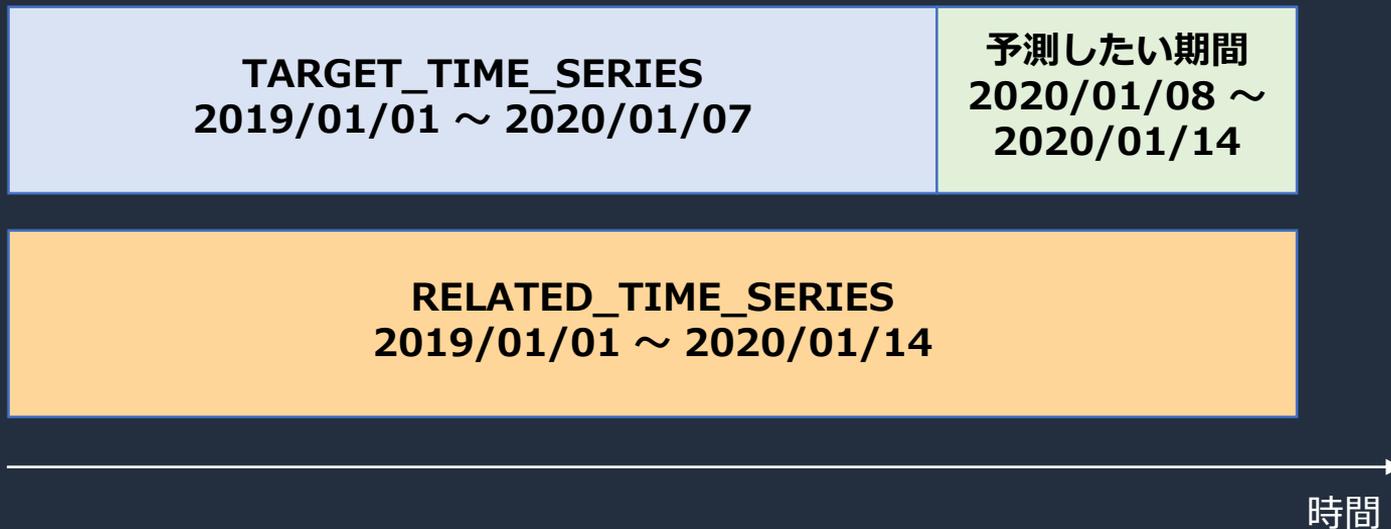
https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/howitworks-datasets-groups.html#howitworks-data-alignment

データの準備 RELATED_TIME_SERIES

- TARGET_TIME_SERIES に含まれていない時系列データ
- 最大10個の予測フィールド (TARGET_TIME_SERIESと同じ)
- 最大13個の関係時系列データ
 - 関係時系列データは float 型または int 型
 - 例：天気を示す場合
 - is_sunny, is_cloudy, is_rainy の関係時系列として定義するのを推奨
- RELATED_TIME_SERIES のデータ頻度は、予測の頻度と合わせる必要がある
 - https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/related-time-series-datasets.html

データの準備 RELATED_TIME_SERIES

- TARGET_TIME_SERIES の期間 + 予測したい期間の関連データが必要



データの準備 RELATED_TIME_SERIES

例 (RETAIL ドメインの場合)

- item_id (商品のID)
- timestamp (タイムスタンプ)
- price (商品価格)
- promotion_applied (プロモーションフラグ)

```
2019-01-01,socks,10,0
2019-01-05,socks,10,0
2019-01-03,socks,15,0
...
2019-06-01,socks,10,0
...
2019-07-01,socks,50,1
...
2019-01-01,shoes,10,0
2019-01-05,shoes,15,0
...
2019-07-01,shoes,10,0
```

CSV

```
{
  "attributes": [
    {
      "attributeName": "timestamp",
      "attributeType": "timestamp"
    },
    {
      "attributeName": "item_id",
      "attributeType": "string"
    },
    {
      "attributeName": "price",
      "attributeType": "float"
    },
    {
      "attributeName": "promotion_applied",
      "attributeType": "integer"
    }
  ]
}
```

スキーマ

データの準備 ITEM_METADATA

- 必須項目 item_id を含む最大10フィールド定義可能
- すべての値は string 型

例 (RETAIL ドメインの場合)

- item_id (商品のID)
- category (商品カテゴリ)
- brand (ブランド) など

```
socks, socks, brandA
socks1, socks, brandB
...
shoes, shoes, brandC
shoes1, shoes, brandD
...
```

CSV

```
{
  "attributes": [
    {
      "attributeName": "item_id",
      "attributeType": "string"
    },
    {
      "attributeName": "category",
      "attributeType": "string"
    },
    {
      "attributeName": "brand",
      "attributeType": "string"
    }
  ]
}
```

スキーマ

データの準備 データセットの作成

- CreateDataset API または、マネジメントコンソール上から可能
- 時系列データの場合はデータの頻度を指定する
- スキーマを定義する

Create target time series dataset [Info](#)

Dataset details

Dataset name
The name can help you distinguish this dataset from other datasets on your Datasets dashboard.

Frequency of your data
This is the frequency at which entries are registered into your data file.

Your data entries have a time interval of

To help Amazon Forecast understand the fields in your data, you must define the schema. Specify the attributes in the schema.

```
1- {
2-   "Attributes": [
3-     {
4-       "AttributeName": "item_id",
5-       "AttributeType": "string"
6-     },
7-     {
8-       "AttributeName": "timestamp",
9-       "AttributeType": "timestamp"
10-    },
11-    {
12-      "AttributeName": "demand",
13-      "AttributeType": "float"
14-    }
15-   ]
16- }
```

データの頻度

スキーマの定義

データの準備 データのインポート

Import target time series data [Info](#)

Dataset import details

Dataset import name
The name can help you distinguish this dataset import from other imports on your dataset detail page.

The dataset import name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and .

Timestamp format [Info](#)
This is the format of the timestamp in your dataset. The format that you enter here must match the format in your data file.

IAM Role [Info](#)
Dataset groups require permissions from IAM to read your dataset files in S3. Choose or create a role using this control.

Data location [Info](#)
The location is the path to the file in your S3 bucket that contains your data.

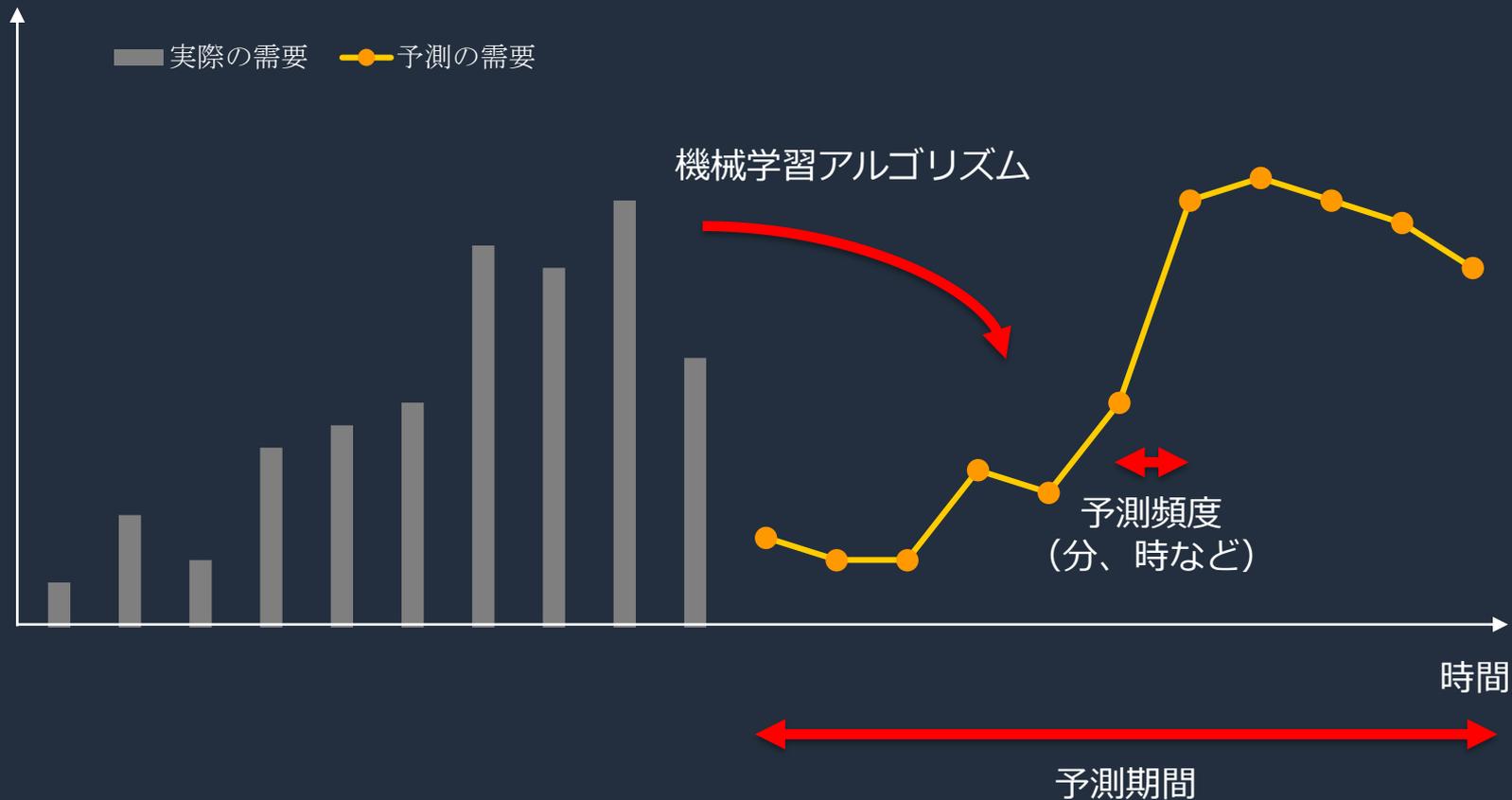
Your files must be in CSV format.

- CSVファイルのタイムスタンプのフォーマットを指定
- CSVデータが格納されている S3 のパスとアクセスするためのロールを指定する
- CSVファイルを指定した場合、指定ファイルのみ読み込む
- S3 バケットの名前が指定された場合は、バケット内のすべてのCSVファイルをインポートする (最大 10,000 ファイル)
- 指定例：
 - s3://bucket-name/example-object.csv
 - s3://bucket-name/prefix/
 - s3://bucket-name

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

予測子の作成 概要



予測子の作成 トレーニング

- CreatePredictor API または、マネジメントコンソール上から可能
- 予測期間 (Forecast Horizon) の指定
- 予測頻度 (Forecast Frequency) の指定
- アルゴリズムの選択
- その他オプションを指定

- トレーニングの実行
- 予測子の評価

https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/howit-works-predictor.html

Train predictor Info

Amazon Forecast uses an algorithm to train a predictor on the data in your dataset group.

Predictor details

Predictor name
The name can help you distinguish this predictor from your other predictors.

The predictor name must have 1 to 63 characters. Valid characters: a-z, A-Z, 0-9, and _

Forecast horizon Info
This number tells Amazon Forecast how far into the future to predict your data at the specified forecast frequency.

Forecast frequency
This is the frequency at which your forecasts are generated.

Your forecast frequency is

Algorithm selection Info
An algorithm is used to train your predictor.

Automatic (AutoML)
Let Amazon Forecast choose the right algorithm for your dataset.

Manual
Explore the algorithms and choose one.

Forecast dimensions - optional
Item id is used in training by default. Select additional keys you would like to use to generate a forecast. These keys are fields in your dataset.

Country for holidays - optional
The holiday calendar you want to include for model training

Number of backtest windows - optional Info
This is the number of times that the algorithm splits the input data for use in training and evaluation.

Backtest window offset - optional Info
This is the point in the dataset where you want to split the data for model training and evaluation.

▶ **Advanced configurations**
Set advanced configurations for your predictor and forecasts.

予測子の作成 トレーニング

- 予測頻度 (Forecast Frequency)
 - 予測をするデータの頻度
 - 分 (minutes)、時間 (hour)、日 (day)、週 (week)、月 (month)、年 (year) から選択可能
- 予測期間 (Forecast Horizon)
 - 予測をする期間。ステップ数。
 - ただし、500 ステップまたはTARGET_TIME_SERIES データセットの長さの 1/3 の小さい方まで

例： Forecast Frequency を Hour に Forecast Horizon を 168 と指定した場合

$168/24 \text{ hour} = 7$ 日間先まで時間の頻度で予測

予測子の作成 アルゴリズムの種類と概要

名前	利用されるデータセットタイプ			概要
	TARGET_TIME_SERIES	RELATED_TIME_SERIES	ITEM_METADATA	
ARIMA	○			自己回帰和分移動平均。時系列予測に一般的に使用される局所統計アルゴリズム。定常時系列にマッピングできるデータセットに特に活用できる。
ETS	○			指数平滑法。時系列予測に一般的に使用される局所統計アルゴリズム。傾向や、季節性を含むデータセットに特に活用できる。
NPTS	○	○		ノンパラメトリック時系列。過去の観測からサンプリングすることで、特定の時系列の未来の価値分布を予測する。断続的なスパイクデータを含む場合に活用できる。
Prophet	○	○		局所的なベイズ構造時系列モデル。長期にわたる傾向や循環性、季節性、重要だが不規則に起きるイベントなどのデータセットに特に活用できる。
DeepAR+	○	○	○	再帰型ニューラルネットワーク (RNN) を使用した、教師あり学習アルゴリズム。予測する時系列に関するデータや特徴データが多くデータセットに含まれる場合に優れている。HPO (ハイパーパラメータの最適化) をサポートしている。

https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/aws-forecast-choosing-recipes.html

予測子の作成 **DeepAR+**

- 再帰型ニューラルネットワーク (RNN) を使用した時系列予測のための教師あり学習アルゴリズム
- 関連時系列や、カテゴリ別特徴を扱うことができる

最適化のベストプラクティス

- 常に時系列全体を入力とする
- 予測期間に非常に大きい値 (> 400) を指定することを避ける
 - 期間を伸ばすには 1分間隔から1時間間隔と集約する
- 多くの時系列で DeepAR+ モデルをトレーニングする

https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/aws-forecast-recipe-deeparplus.html

予測子の作成 AutoML

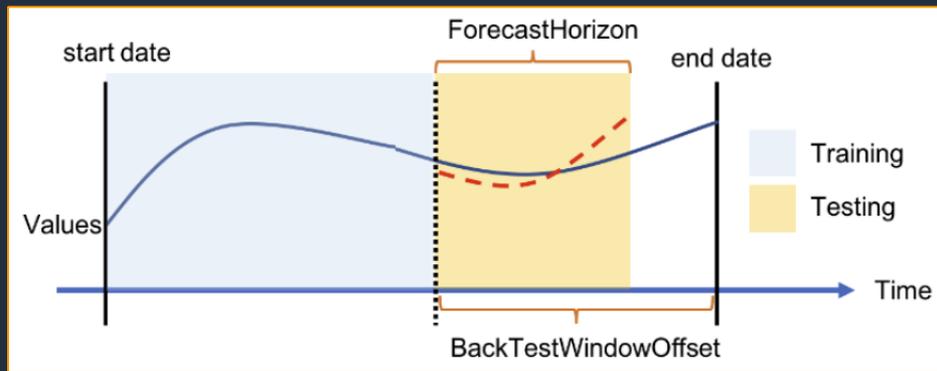
- AutoML は全てのアルゴリズムを評価し、データセットに基づいて最適なアルゴリズムを Amazon Forecast が自動で選択
- 評価された各アルゴリズムのメトリクスは予測子の詳細ページから確認できる

Algorithm metrics Info			
Evaluation type SUMMARY	Test window start -	Test window end -	
wQL[0.5] 0.5990	RMSE 12.9328	wQL[0.9] 0.2411	
Evaluation type COMPUTED	Test window start Sat, 28 Dec 2019 00:00:00 GMT	Test window end Mon, 30 Dec 2019 00:00:00 GMT	
wQL[0.5] 0.5990	RMSE 12.9328	wQL[0.9] 0.2411	wQL[0.1] 0.1933

NPTS arn:aws:forecast::algorithm/NPTS (Winning algorithm) ▲
NPTS arn:aws:forecast::algorithm/NPTS (Winning algorithm)
Deep_AR_Plus arn:aws:forecast::algorithm/Deep_AR_Plus
ARIMA arn:aws:forecast::algorithm/ARIMA
Prophet arn:aws:forecast::algorithm/Prophet
ETS arn:aws:forecast::algorithm/ETS

予測子の作成 予測子の評価

- Amazon Forecast は、入力データとテストデータセットの2つのデータセットに自動的に分割し、バックテストを使用してメトリクスを生成



- トレーニング後、Amazon Forecast は、二乗平均平方根誤差 (RMSE) と重み付き分位点損失 (wQuantileLoss) を計算する
- GetAccuracyMetrics API またはマネジメントコンソールから確認

https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/metrics.html

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

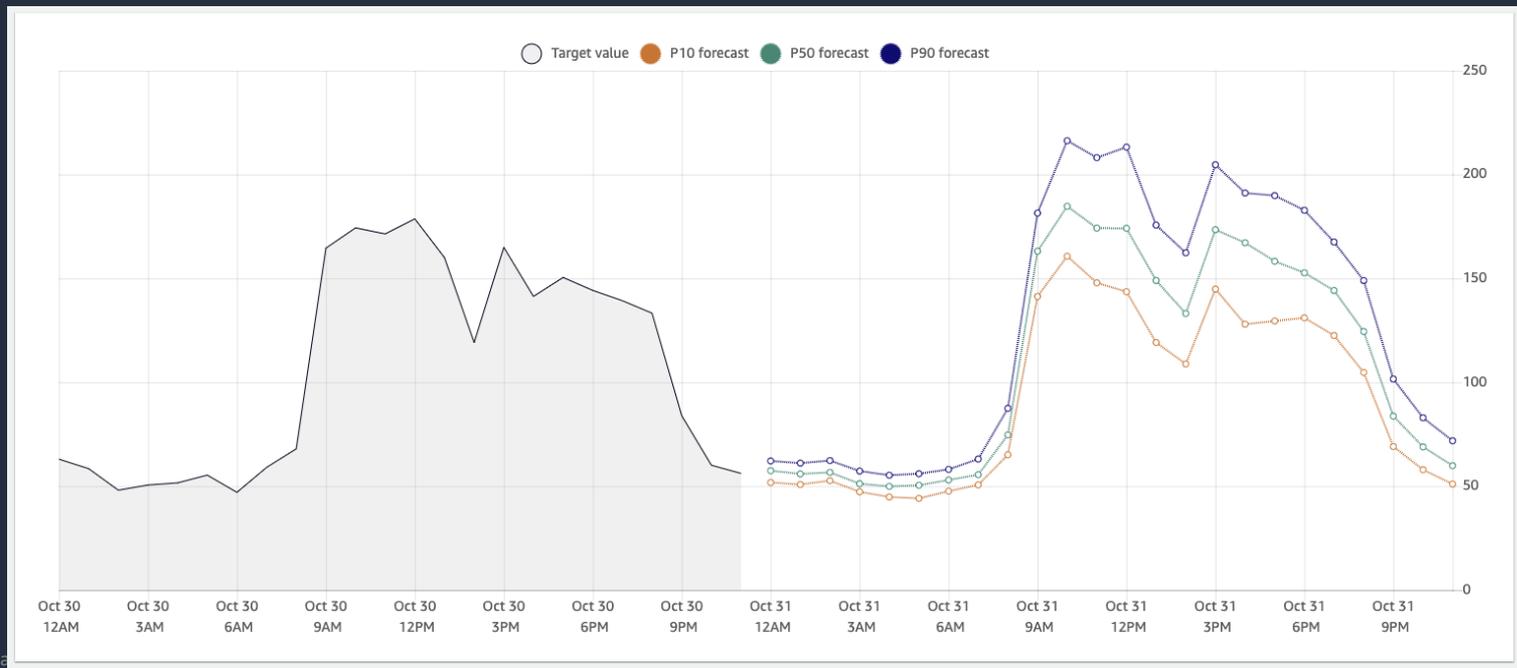
予測の作成



https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/API_forecastquery_QueryForecast.html
https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/API_CreateForecastExportJob.html

予測の作成 予測結果の確認

- API での予測結果の確認以外に、マネジメントコンソール上の Forecast lookup から簡単に確認することもできる
- 最大5つまでの分位点を指定することができる



予測の作成 予測のサイクル

- データセットを追加でインポートする
- 学習済みの予測子を指定し、予測を作成する（予測子の再学習は不要）

データセットのインポート

学習データ
2019/01/01 ~
2020/01/07

追加データ
2020/01/08 ~
2020/01/14

予測子のトレーニング

予測子

予測の作成

予測データ
2020/01/08 ~
2020/01/14

予測データ
2020/01/15 ~
2020/01/21

予測の作成 予測のサイクル

- データセットを追加でインポートする
- 学習済みの予測子を指定し、予測を作成する（予測子の再学習は不要）
- RELATED_TIME_SERIES を利用している場合は、RELATED_TIME_SERIES の追加インポートも必要
- ただし、ある程度の期間で予測子の再学習は必要



本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

セキュリティ

- Amazon Key Management Service (KMS) による暗号化
- アイデンティティベースのポリシー(IAM ポリシー) のサポート

例 :

```
{  
  "Version":"2012-10-17",  
  "Statement":[  
    {  
      "Effect":"Allow",  
      "Action":[  
        "forecast:DescribeForecast",  
      ],  
      "Resource":"arn:aws:forecast:region:acct-id:predictor/predictor-name"  
    }  
  ]  
}
```

JSON

- AWS CloudTrail によるAPIコールの監視
- Amazon CloudWatch メトリクスのサポート

https://docs.aws.amazon.com/ja_jp/forecast/latest/dg/security.html

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

料金体系

Amazon Forecast は、予測数、保存およびトレーニングのために使用したストレージ量、トレーニングにかかった時間に対して課金される

対象	料金
予測数	0.60 USD/1,000 forecasts
データストレージ	0.088 USD/GB
トレーニング時間	0.24 USD/時間

無料利用枠

- Amazon Forecast を利用開始時点より最初の 2 ヶ月は、以下が提供される
 - 予測数：毎月最大 10,000 Forecasts
 - データストレージ：毎月最大 10 GB
 - トレーニング時間：毎月最大 10 時間

料金体系

1ヶ月の料金例：

2,000アイテムを50店舗で販売しているデータに対し、各店舗の各アイテムに1時間間隔分の予測を行なうため、5GBのデータを利用し、20時間のトレーニング時間がかかった場合。

- 予測数
 - $2,000 \times 50 / 1,000 \times 3 \text{ (quantiles)} \times 0.60 \text{ USD} = 180 \text{ USD}$
- データストレージ
 - $5 \text{ GB} \times 0.088 \text{ USD} = 0.44 \text{ USD}$
- トレーニング時間
 - $20 \times 0.24 \text{ USD} = 4.80 \text{ USD}$

合計コスト = 180 USD + 0.44 USD + 4.80 USD = 185.24 USD

<https://aws.amazon.com/jp/forecast/pricing/>

提供リージョン

- 東京
- 北バージニア
- オハイオ
- オレゴン
- シンガポール
- アイルランド

本日のアジェンダ

- Amazon Forecast とは
- Amazon Forecast の使い方
 - Amazon Forecast のワークフロー概要
 - データの準備
 - 予測子の作成
 - 予測の作成
 - デモ
- セキュリティ
- 料金体系
- まとめ

まとめ

- Amazon Forecast は、Amazon.com と同様の技術を用いた、様々な時系列データの予測を簡単に行うことのできる機械学習サービス
- Amazon Forecast は、事前定義済みアルゴリズム、および、AutoML も提供されているため、機械学習の専門知識がなくても、予測を行うことが可能

参考資料

- 公式ページ
 - <https://aws.amazon.com/jp/forecast/>
- ドキュメントやチュートリアルなどのリソース
 - <https://aws.amazon.com/jp/forecast/resources/>
- サンプル
 - <https://github.com/aws-samples/amazon-forecast-samples>
- 事例
 - <https://aws.amazon.com/jp/forecast/customers/>

補足

Amazon の機械学習への取り組み、各種 AWS 機械学習サービスの位置付けに関しては、以下の資料もご参照ください。



The slide features a dark green background with a repeating geometric pattern of squares. At the top left is the AWS logo. The main title is "[AWS Black Belt Online Seminar] Amazon AI Service" in white, with "サービスキャストシリーズ" below it. At the bottom left, it lists "Solutions Architect 長谷川 仁志" and "2019/9/24". In the center, there are two QR codes: the left one is labeled "AWS 公式 Webinar" with the URL "https://amzn.to/JPWebinar", and the right one is labeled "過去資料" with the URL "https://amzn.to/JP&archive". The AWS logo is also present at the bottom right.

aws

[AWS Black Belt Online Seminar]
Amazon AI Service
サービスキャストシリーズ

Solutions Architect 長谷川 仁志
2019/9/24

AWS 公式 Webinar
<https://amzn.to/JPWebinar>

過去資料
<https://amzn.to/JP&archive>

aws

[Amazon AI Service](#)

AWS Black Belt Online Seminar @ 2019.9.24

Q&A

お答えできなかったご質問については

AWS Japan Blog 「<https://aws.amazon.com/jp/blogs/news/>」にて
後日掲載します。

AWS の日本語資料の場所「AWS 資料」で検索



日本担当チームへお問い合わせ サポート 日本語 ▼ アカウント ▼

コンソールにサインイン

製品 ソリューション 料金 ドキュメント 学習 パートナー AWS Marketplace その他 🔍

AWS クラウドサービス活用資料集トップ

アマゾン ウェブ サービス (AWS) は安全なクラウドサービスプラットフォームで、ビジネスのスケールと成長をサポートする処理能力、データベースストレージ、およびその他多種多様な機能を提供します。お客様は必要なサービスを選択し、必要な分だけご利用いただけます。それらを活用するために役立つ日本語資料、動画コンテンツを多数ご提供しております。(本サイトは主に、AWS Webinar で使用した資料およびオンデマンドセミナー情報を掲載しています。)

[AWS Webinar お申込 »](#)

[AWS 初心者向け »](#)

[業種・ソリューション別資料 »](#)

[サービス別資料 »](#)

<https://amzn.to/JPArchive>



AWS Well-Architected 個別技術相談会

毎週“W-A個別技術相談会”を実施中

- AWSのソリューションアーキテクト(SA)に
対策などを相談することも可能

- **申込みはイベント告知サイトから**

(<https://aws.amazon.com/jp/about-aws/events/>)

AWS イベント

で[検索]



ご視聴ありがとうございました

AWS 公式 Webinar

<https://amzn.to/JPWebinar>



過去資料

<https://amzn.to/JPArchive>

