

未来を予測する！ Amazon Forecast の Tips とユースケース

大淵 麻莉

アマゾン ウェブ サービス株式会社
機械学習ソリューションアーキテクト

自己紹介



大淵 麻莉（おおぶち まり）

機械学習ソリューションアーキテクト

- 前職はカメラメーカーで画像処理と機械学習

好きなAWSサービスは

Amazon Rekognition Custom Labels

AWS のミッション

すべての開発者の手に機械学習を


AWS の機械学習サービス

幅広く機械学習サービスを提供







AI サービス: 機械学習の深い知識なしに利用可能

画像  Amazon Rekognition	音声   Amazon Polly Amazon Transcribe + Medical	テキスト    Amazon Comprehend Amazon Translate Amazon Textract + Medical	NEW! 検索  Amazon Kendra	チャットボット  Amazon Lex	レコメンド  Amazon Personalize	時系列予測  Amazon Forecast	NEW! 不正検知  Amazon Fraud Detector	NEW! コードレビュー  Amazon CodeGuru	NEW! コンタクトセンタ  Contact Lens For Amazon Connect
---	---	--	---	--	--	---	---	--	--

ML サービス: 機械学習のプロセス全体を効率化するマネージドサービス

 Amazon SageMaker	Ground Truth	Augmented AI	ML Marketplace	Amazon SageMaker Studio IDE NEW!								Neo
				Built-in algorithms	NEW! Notebooks	NEW! Experiments	Model training & tuning	NEW! Debugger	NEW! Autopilot	Model hosting	NEW! Model Monitor	

ML フレームワークとインフラストラクチャ: 機械学習の環境を自在に構築して利用

 TensorFlow	 mxnet	 GLUON	 Keras	Deep learning AMIs & containers	GPUs & CPUs	Amazon Elastic Inference	AWS Inference	FPGA
 PYTORCH	 HOROVOD							

AWS の機械学習サービス

幅広く機械学習サービスを提供

AI サービス: 機械学習の深い知識なしに利用可能

画像 Amazon Rekognition	音声 Amazon Polly Amazon Transcribe + Medical	テキスト Amazon Comprehend + Medical Amazon Translate	NEW! 検索 Amazon Kendra	チャットボット Amazon Lex	レコメンド Amazon Personalize	時系列予測 Amazon Forecast	NEW! 不正検知 Amazon Fraud Detector	NEW! コードレビュー Amazon CodeGuru	NEW! コンタクトセンタ Contact Lens For Amazon Connect
--------------------------	--	--	-----------------------------	-----------------------	-----------------------------	--------------------------	---------------------------------------	------------------------------------	--

ML サービス: 機械学習のプロセス全体を効率化するマネージドサービス

Amazon SageMaker	Ground Truth	Augmented AI	ML Marketplace	Build-in algorithms	NEW! Neo	NEW! Notebook	NEW! Experiments	Model training & tuning	Amazon SageMaker Studio IDE NEW!	NEW! Debugger	NEW! Autopilot	Model hosting	NEW! Model Monitor	Neo
------------------	--------------	--------------	----------------	---------------------	-------------	------------------	---------------------	-------------------------	-------------------------------------	------------------	-------------------	---------------	-----------------------	-----

ML フレームワークとインフラストラクチャ: 機械学習の環境を自在に構築して利用

TensorFlow	mxnet	PyTorch	Horovod	Amazon Forecast	Deep learning AMIs & containers	GPUs & CPUs	Amazon Elastic Inference	AWS Inferentia	FPGA
------------	-------	---------	---------	-----------------	------------------------------------	-------------	--------------------------	----------------	------

このセッションのゴール

ハンズオンで Amazon Forecast を触ったことはあるけれど、
自社でどうやって使ったら良いかわからずモヤモヤ している方が、
こういうときに Amazon Forecast を使えば良いのか！と理解できる。

自社データで Amazon Forecast を使って時系列予測を行なっているけれど、
予測モデルの性能評価や予測精度向上の方法に悩んでいる 方が、
問題解決のヒントを得て明日から手を動かせるようになる。

Agenda

時系列データ予測とは

Amazon Forecast のしくみ

Amazon Forecast を活用するための Tips

まとめ

時系列データ予測とは？

過去の時系列データから、未来の値を予測する

タイムスタンプ	アイテムID	販売数
⋮	⋮	⋮
2020/6/24	Item_1	3
2020/6/24	Item_2	12
2020/6/25	Item_1	2
2020/6/25	Item_2	15
2020/6/26	Item_1	?
2020/6/26	Item_2	?

ここを予測！

なぜ予測したいのか？

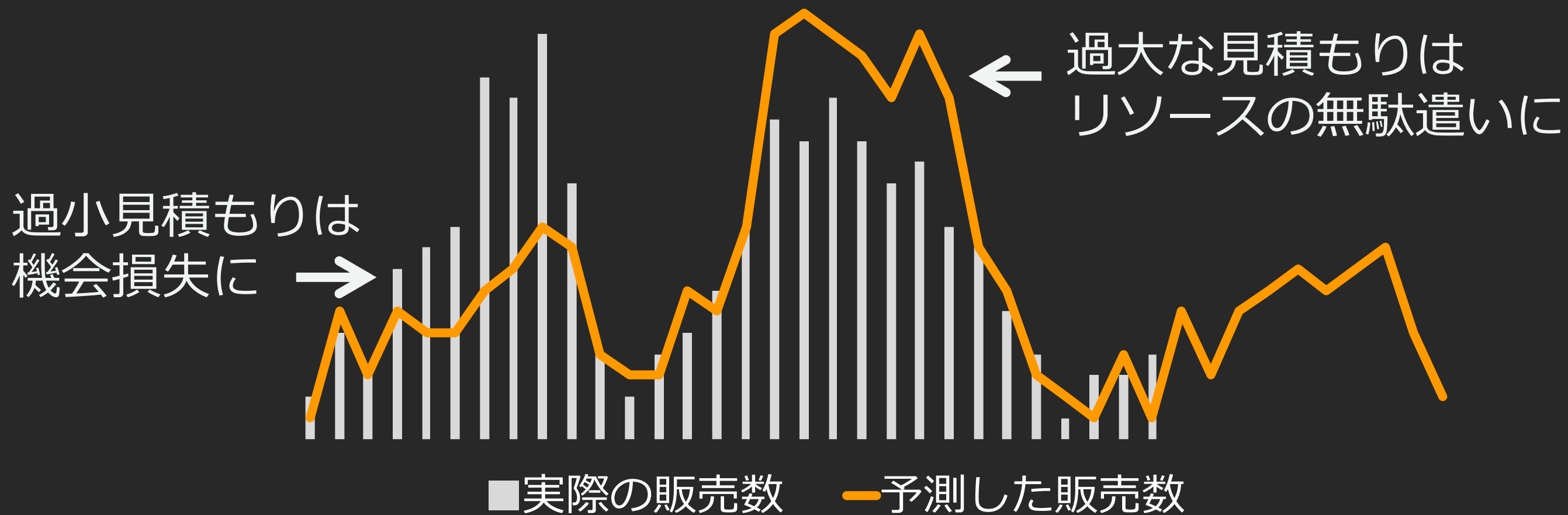
未来を予測できれば、

コスト削減などビジネス価値を生むために

これから何をすべきかがわかる

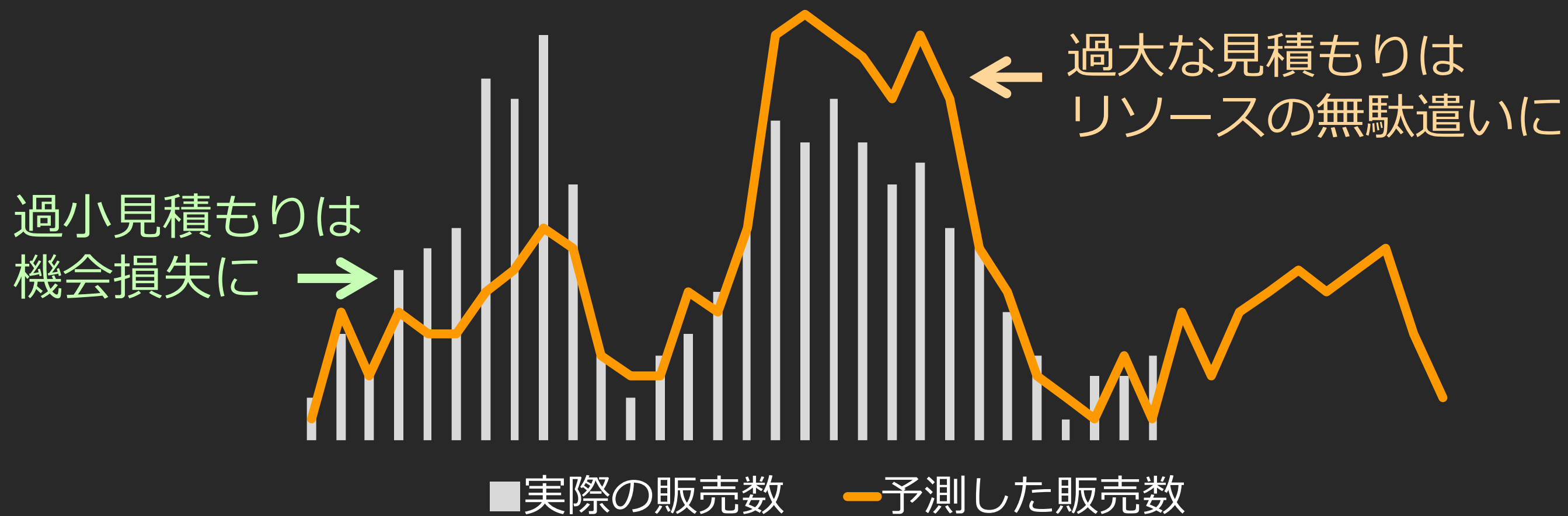
時系列データ予測の難しさ

商品販売数の予測の場合



時系列データ予測の難しさ

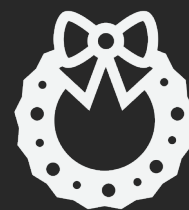
商品販売数の予測の場合



時系列データ予測の課題



季節性への対応



価格・祝日・
プロモーションなどの
関連データの考慮



商品ごとの特徴など
メタデータの考慮



過去のデータがない
新商品の予測

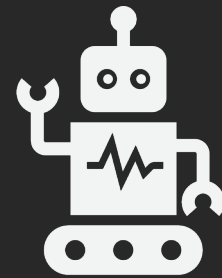
Amazon Forecast の特徴

Amazon.com で利用されている技術を利用して、
将来のビジネスの状況を予測する



フルマネージド

データパイプライン、学習、
予測を自動でセットアップ



AutoML

5種類のアлゴリズム
を自動で試行



簡単

機械学習の知識は不要

CasaOne様: 在庫管理

ビジネス課題:

- 都市間で家具を早く移動させるために需要を予測
- 倉庫における余剰在庫を削減



需要が多い
倉庫スペースをとらない



需要が少ない
倉庫スペースをとる

AffordableTours様: 労働力の管理

ビジネス課題:

- お客様の問い合わせに十分対応するためにリソースを最適化
- 顧客体験を一貫して確保する
- オペレーションコストの削減

Award-winning Service



WINNER OF DISTINCTION
Better Business Bureau (BBB)
May 16, 2019



2018 PINNACLE WINNER
Better Business Bureau (BBB)
May 23, 2018



2018 SALES ACHIEVEMENT
Ensemble Travel Group
Jul 15, 2019



Club 180 Beyond Ordinary Top Partner
Windstar Cruises
Jan 8, 2019

Winner of BBB's Pinnacle Winner

As an award winning travel agency, AffordableTours.com has been ranked as an A+ BBB accredited company since 2001, and also has earned the highest BBB honor, the Pinnacle Award in 2006, 2007, 2011, 2017, 2018. We have also been recognized over the years with the BBB "Award for Excellence" and "Winner of Distinction" every year since 2004, making us a BBB award winner every year for our on-going commitment to our valued clients.

One of the Largest sellers of Escorted Travel

AffordableTours.com's mission remains the same as it was when they started the travel company in 1999: to provide high-quality vacations at low prices, while giving each and every guest personalized attention. Due to our long standing relationship and high sales volume with our travel partners, we earn exclusive pricing and in turn we pass the savings on to our clients.

Personalized service

When you purchase with AffordableTours.com, you will be assigned a dedicated travel consultant who will assist you plan your vacation. Our travel consultants are among the best and the brightest in the industry. We consistently recruit qualified staff and train them to offer our clients incredibly attentive service. Our travel consultants are known for their attention to detail and rapid response to any of your questions or concerns.



Cruise Connoisseur Club
Oceania Cruises
Oct 4, 2018



Excellence Award
Carnival
Nov 6, 2018



Limited Edition Print
Globus
Nov 16, 2018



Most Improved Partner
Apple Vacations
Apr 10, 2018

Call Us For a Free Quote **1-800-935-2620**

There are many friendly consultants at AffordableTours.com that are ready to answer any of your questions or assist you with your reservation.



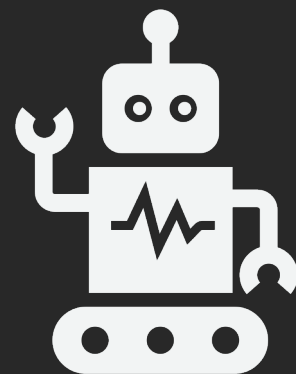
Amazon Forecast のしくみ

Amazon Forecast を使う 3ステップ



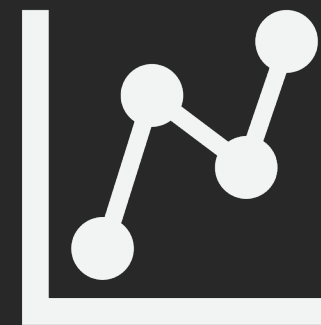
データ読み込み

Import Dataset



モデルの自動学習

Create Predictors



時系列データ予測

Create Forecasts



使用可能なデータ

ターゲット時系列 (必須)

予測したいデータの履歴

- ・商品ごとの販売数
- ・商品の使用量、など

関連時系列

時間変動する関連データの履歴

- ・商品の価格
- ・プロモーション実施有無、など

項目メタデータ

項目の属性や特徴

- ・商品のジャンル
- ・商品の提供場所、など

組み込みデータセット

組み込みの時系列データ

- ・休日



使用可能なデータ

ターゲット時系列（必須）

予測したいデータの履歴

- 商品ごとの販売数
- 商品の使用量
- コールセンター電話数
- 消費電力
- 物流センターの出荷数

など

タイムスタンプ	アイテムID	販売数
2020/6/24	Item_1	3
2020/6/24	Item_2	12
2020/6/25	Item_1	2
2020/6/25	Item_2	15
⋮	⋮	⋮



使用可能なデータ

関連時系列

時間変動する関連データの履歴

- 商品の価格
- プロモーション実施有無
- 気温
- 在庫数
- 売り上げ金額

など

タイムスタンプ	アイテムID	PR
2020/6/24	Item_1	0
2020/6/24	Item_2	1
2020/6/25	Item_1	0
2020/6/25	Item_2	1
⋮	⋮	⋮



使用可能なデータ

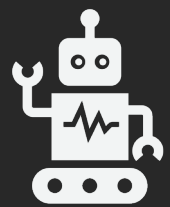
項目メタデータ

項目の属性や特徴

- 商品のジャンル
- 商品の提供場所
- 商品の価格
- 商品の生産地

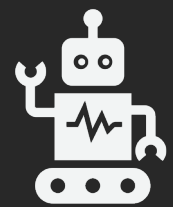
など

アイテムID	カテゴリ	色
Item_1	Tシャツ	青
Item_2	靴	緑
Item_3	手袋	黄色
Item_4	メガネ	赤
⋮	⋮	⋮



使用可能なアルゴリズム

	概要	使用するデータセット
ARIMA	比較的小さなデータセットで使用。項目ごとに独立して予測。定常的なデータの予測が得意。	ターゲット時系列
ETS	比較的小さなデータセットで使用。項目ごとに独立して予測。トレンド、季節性、残差を認識。	ターゲット時系列
NPTS	項目ごとに独立して予測。 ゼロが続くデータにも対応可能。	ターゲット時系列 関連時系列
Prophet	項目ごとに独立して予測。トレンド、季節性、周期性、休日の影響などに対応可能。	ターゲット時系列 関連時系列
DeepAR+	予測したい項目数が多数ある場合、項目間の相互関係も考慮した予測が可能。 新商品の予測（コールドスタート）にも対応。	ターゲット時系列 関連時系列 項目メタデータ



モデル学習における AutoML 機能

AutoML を有効にすることで、以下が自動で実施される

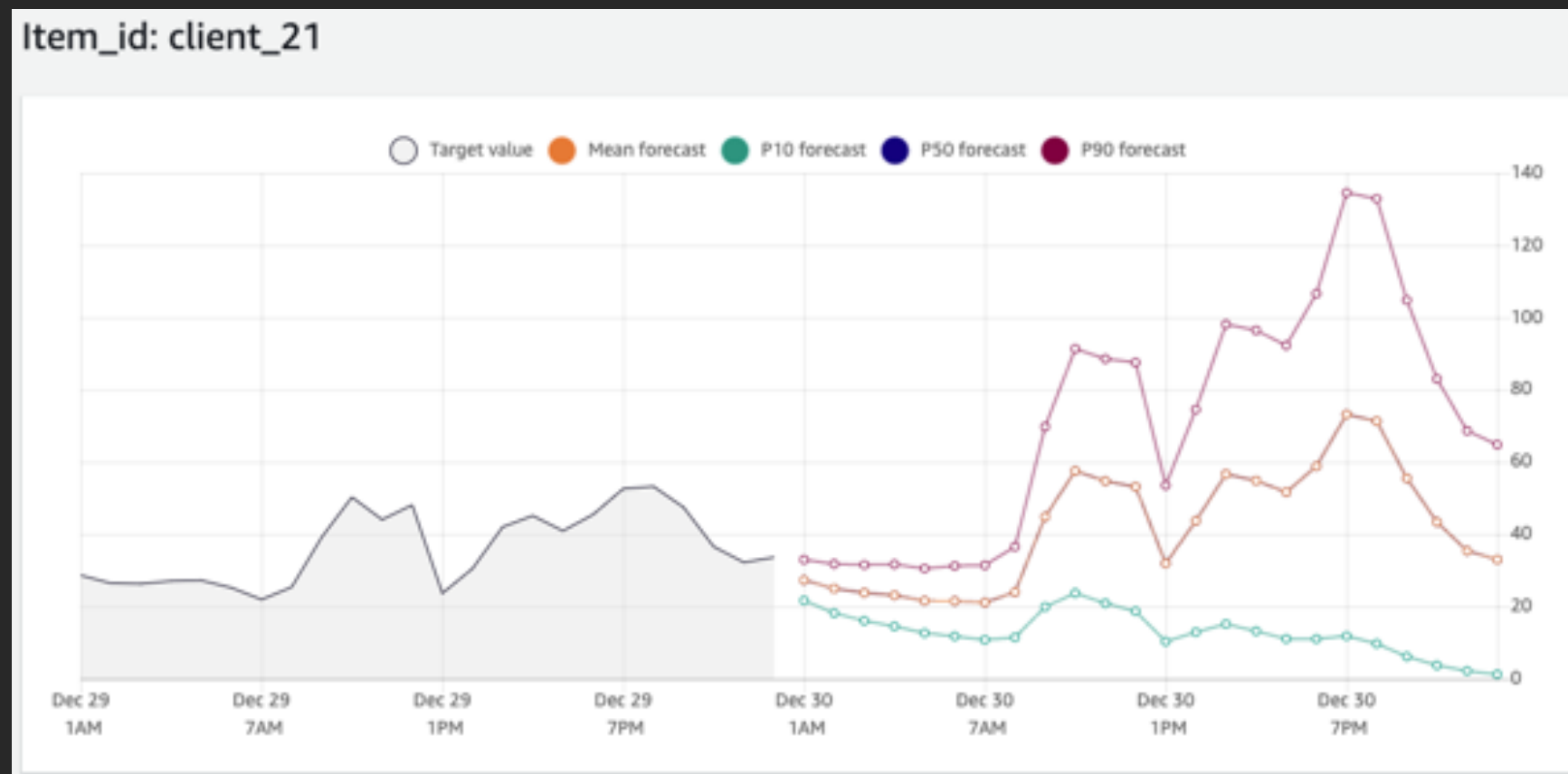
- 使用可能なすべてのアルゴリズムでモデルの学習
- 複数の評価指標の算出

もっとも結果が良かったモデルを使用可能



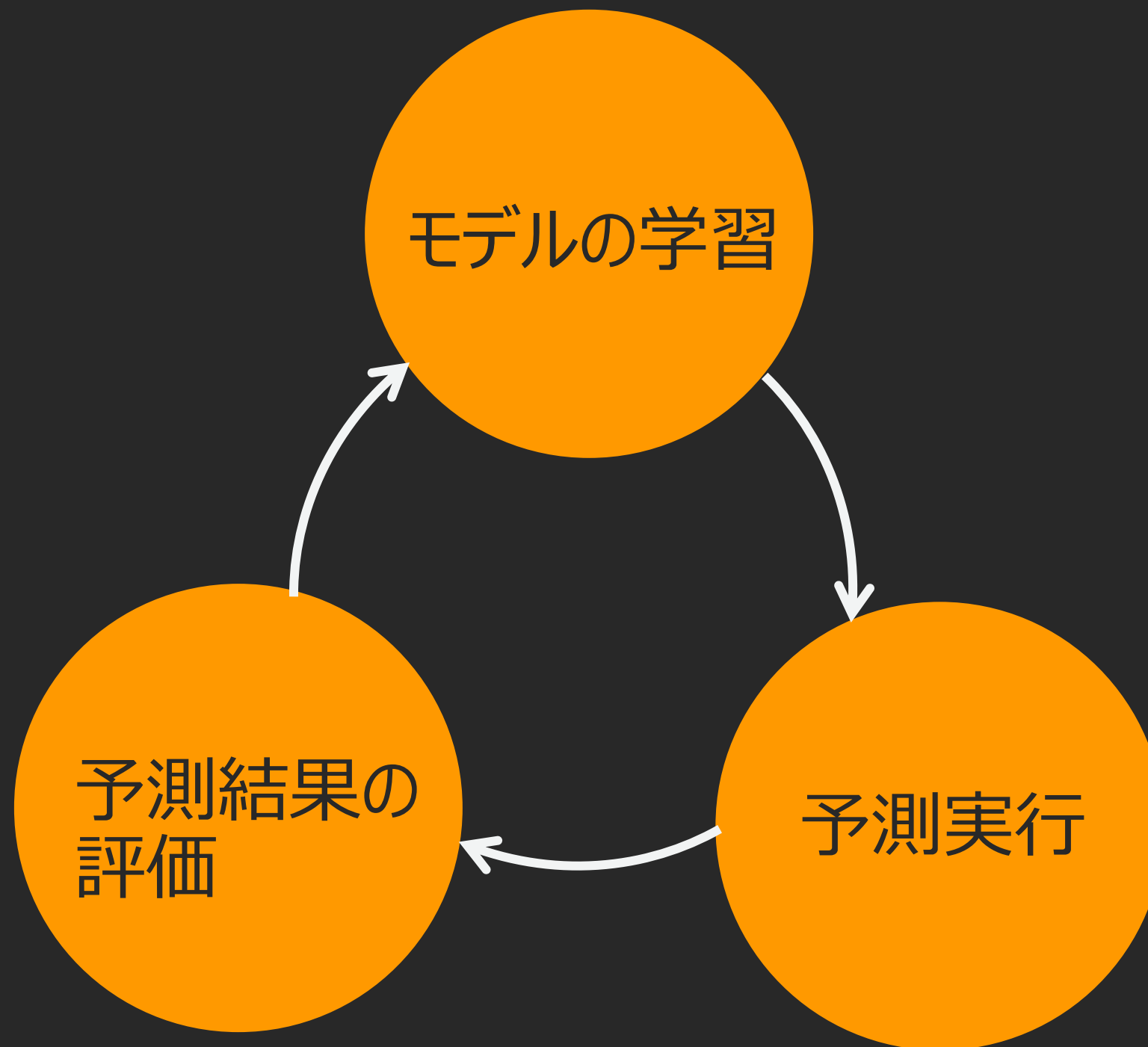
予測データの取得

- 予測は任意の分位数（デフォルトは P10, P50, P90）で出力
- API を使って予測データの取得
- 予測データを CSV にエクスポート



Amazon Forecast 活用の Tips

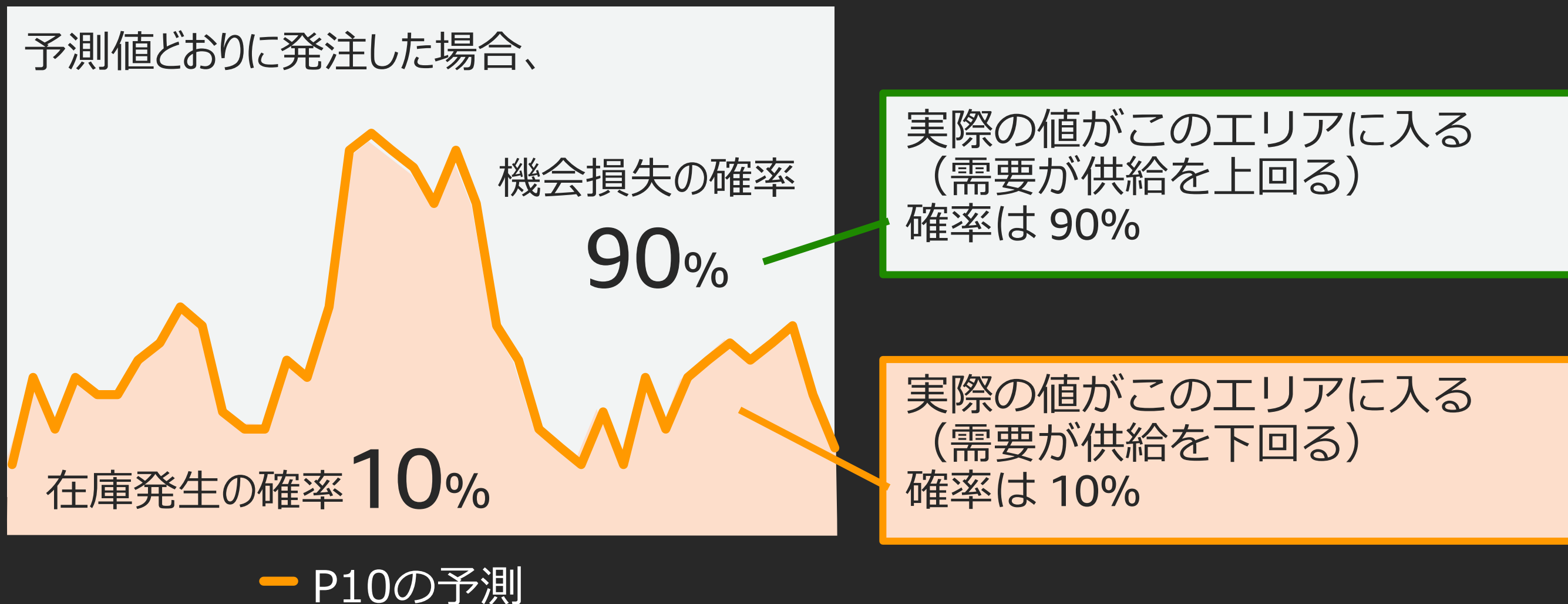
予測精度向上のサイクル



予測分位数の選び方の Tips

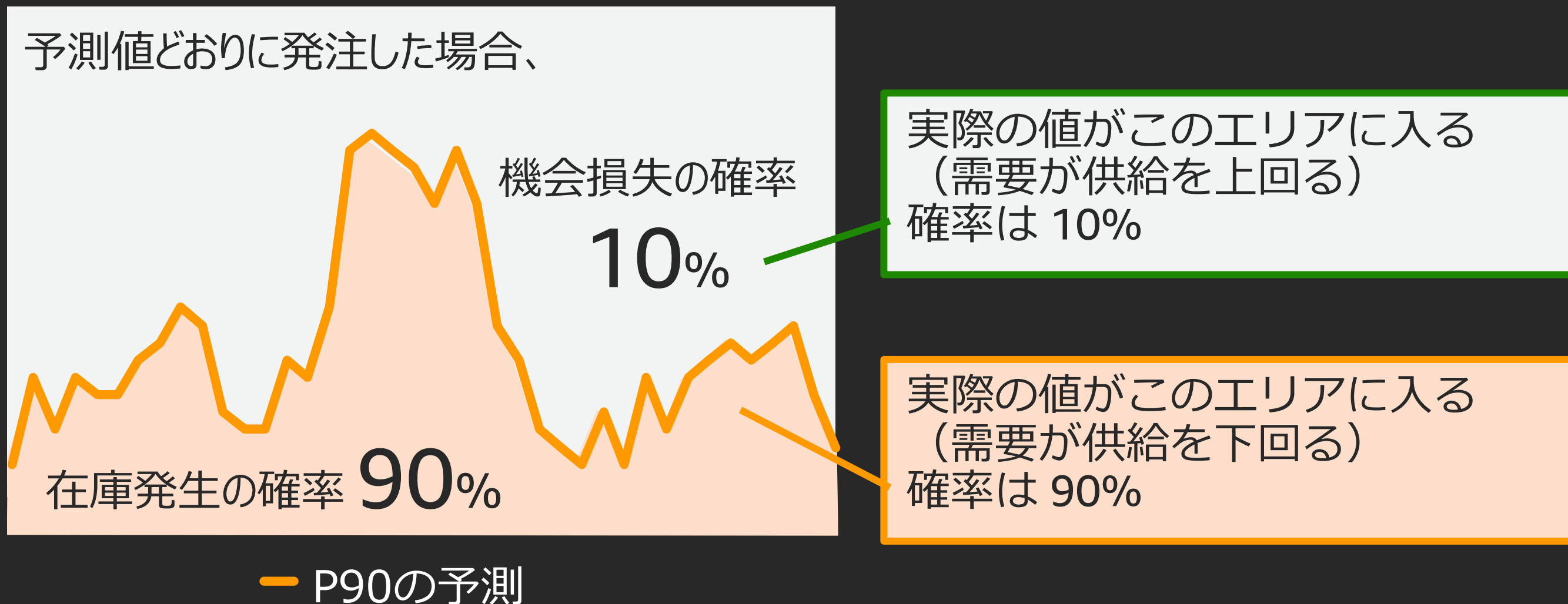
分位数 (P10)

できるだけ **在庫を抱えたくない** 商品におすすめ。



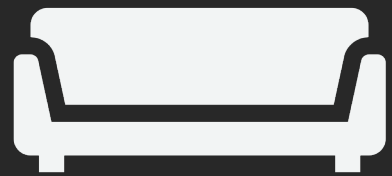
分位数 (P90)

できるだけ **機会損失したくない** 商品におすすめ。

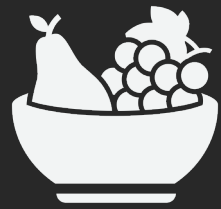


予測分位数の使い分けの例

商品の特征に合わせて最適なものを選択



サイズが大きくて
スペースを取る商品



生鮮食品など
売りきりたい商品



多くの人を求める
目玉商品



在庫になる確率 10%

P10 を使用



機会損失の確率 10%

P90 を使用

モデルの評価における Tips

Amazon Forecast で算出される評価指標

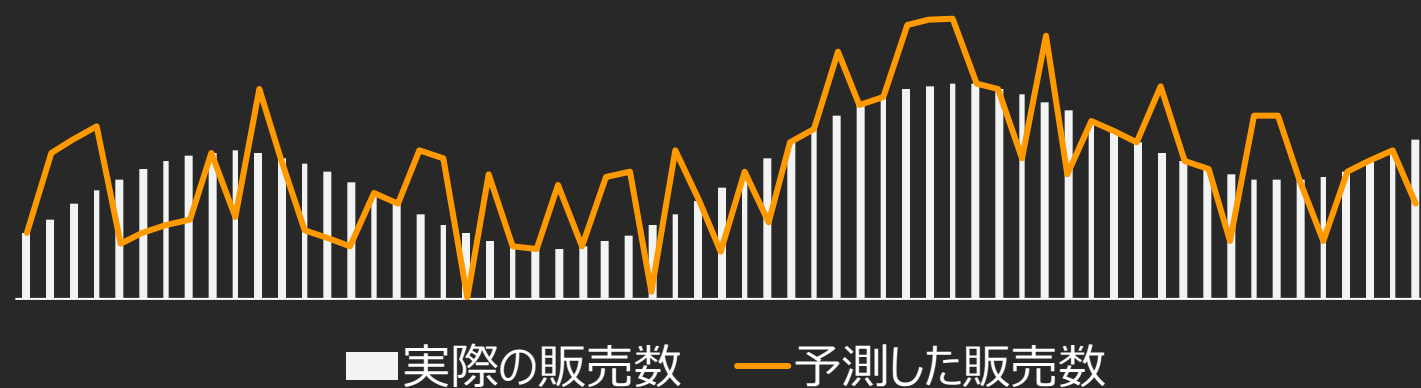
RMSE

MAPE

wQuantileLoss

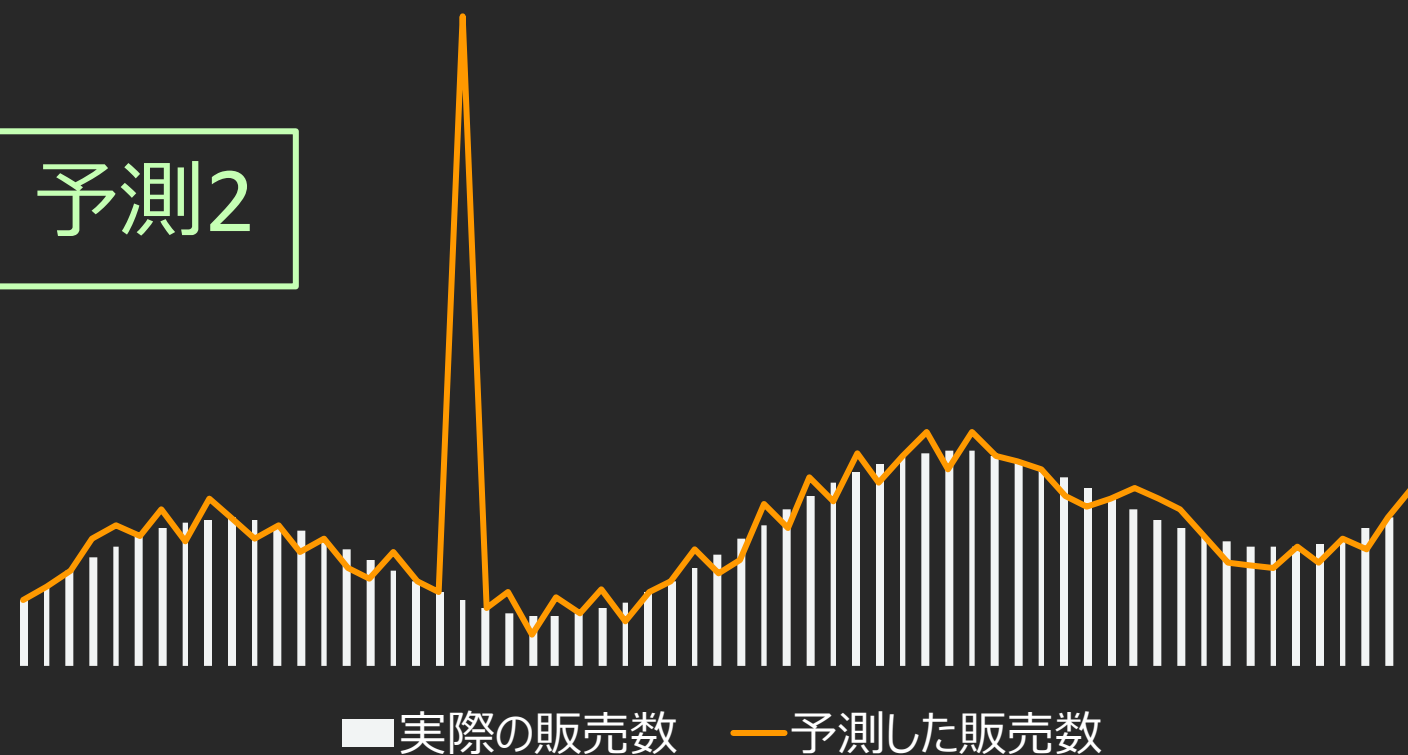
Amazon Forecast で算出される評価指標 (RMSE)

予測1



RMSE: 0.80

予測2



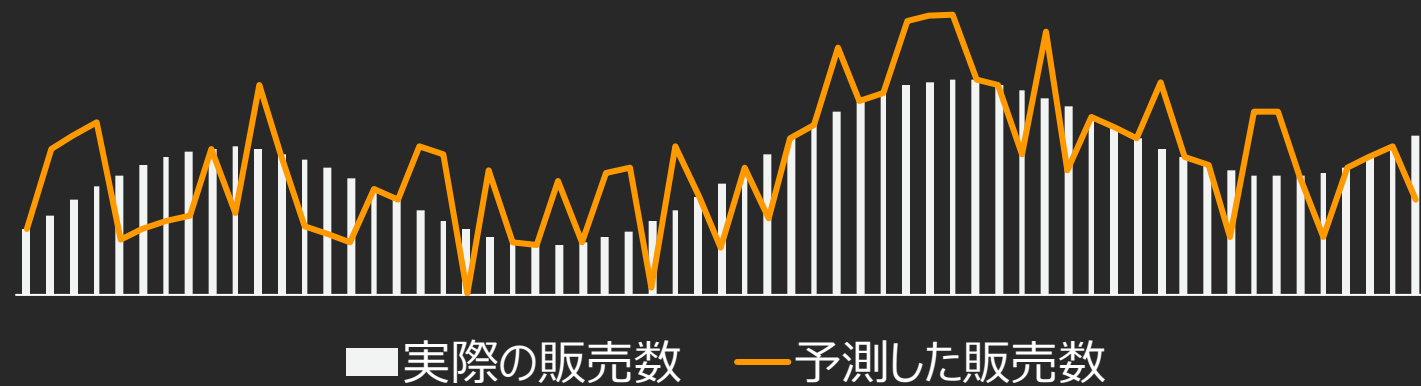
RMSE: 1.21

全体的な予測のズレは予測1の方が大きいですが、大きな外れ値のある予測2の方が RMSE の値は大きくなる。

➡ 外れ値の影響を重く見たい場合に RMSE を参照する。

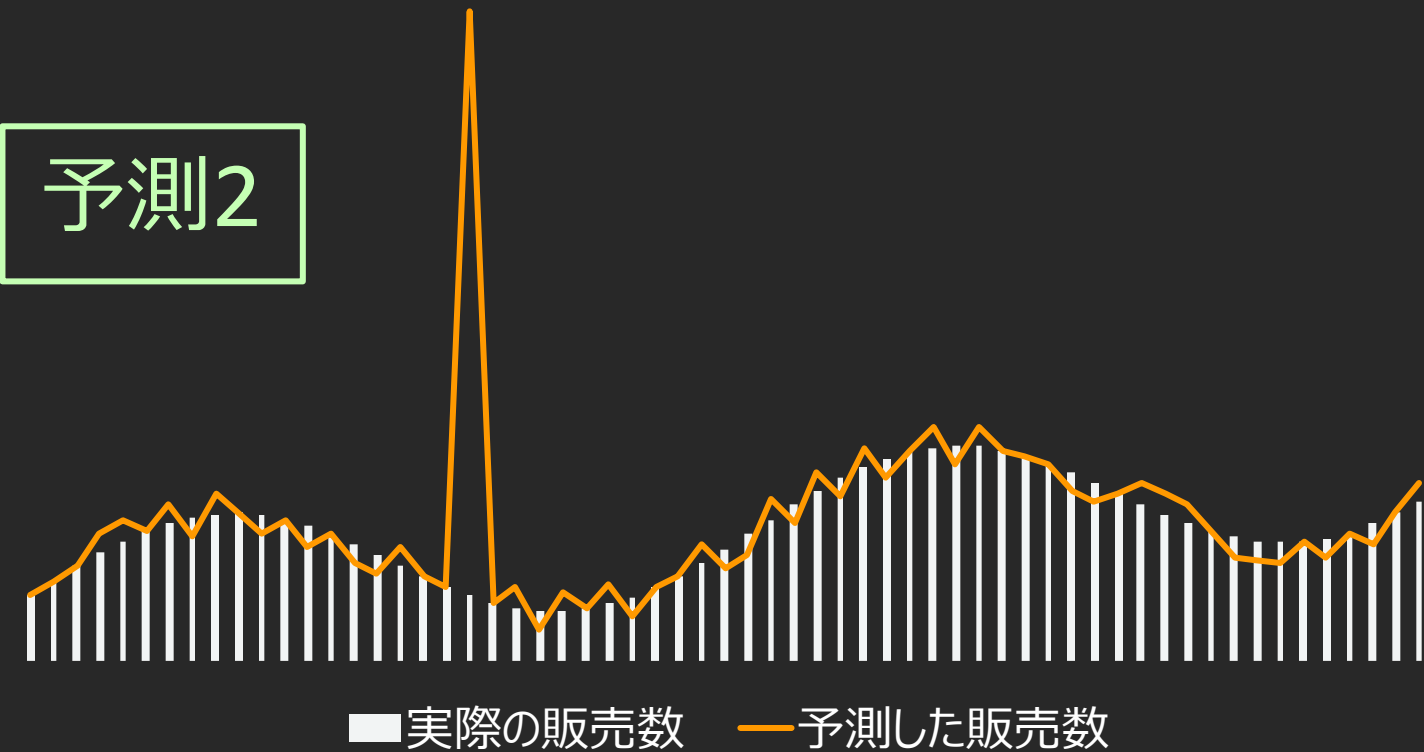
Amazon Forecast で算出される評価指標 (MAPE)

予測1



MAPE: 0.32

予測2



MAPE: 0.17

MAPE の値は、予測1は全体的に 3.2割程度実測値からずれており、予測2は全体的に 1.7割程度実測値からずれていることを示す。

➡ 平均的なモデルの性能を評価する場合に MAPE を参照する。

Amazon Forecast で算出される評価指標 (wQuantileLoss)

MAPE と同様に予測値と実測値の差を割合として算出する指標。

P10 の予測を使用する際は $wQL[0.1]$

P90 の予測を使用する際は $wQL[0.9]$

の値を使用して予測性能を評価する。

予測結果の詳細な評価

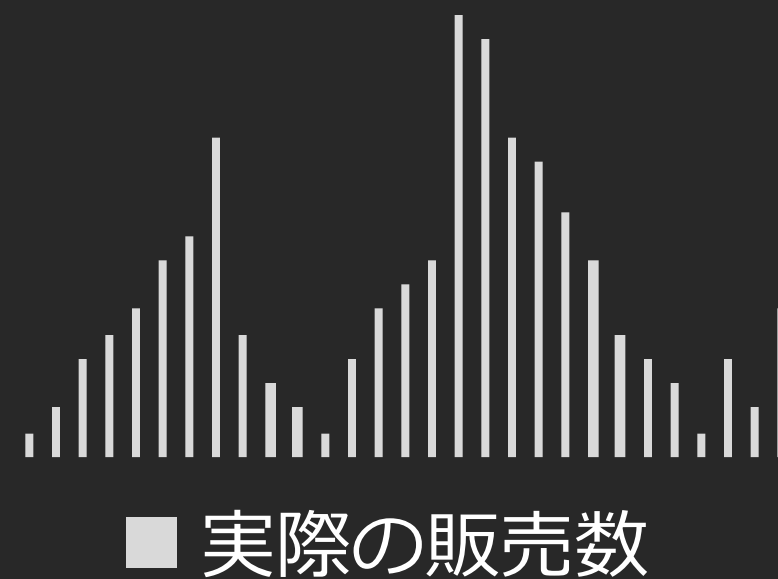
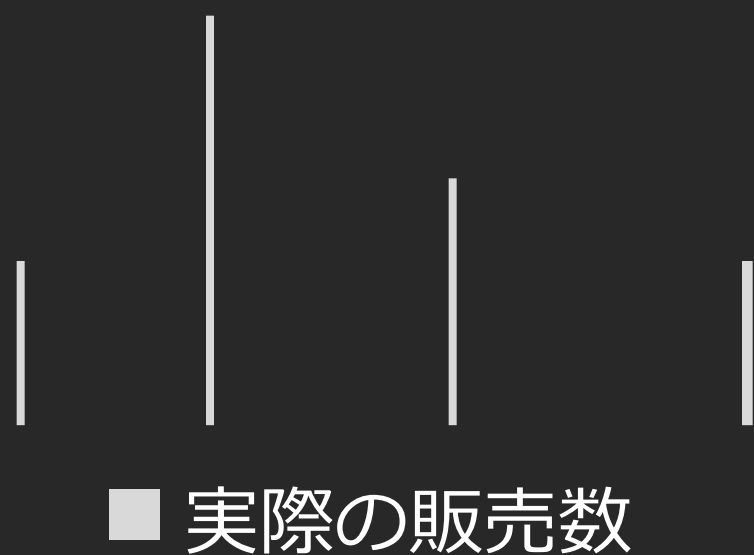
ビジネス価値を生むことができるか確認

商品別に評価指標を算出して比較

- 商品ごとに予測と実際の値の差が異なるか？
- 重要な商品の予測はどの程度正しいか？
- 誤差が大きい商品と小さい商品の違いは何か？
 - 販売数、ジャンル、傾向などに違いはないか？
 - 誤差が大きい商品の値に連動する関連データはあるか？
 - 誤差が大きい商品と小さい商品でモデルを分けるべきか？

使用するデータにおける Tips

予測がむずかしいのはどっち？



予測がむずかしいデータ

数日間ゼロが続くようなデータは予測が非常に困難。
ゼロが続くデータを使っていないかをまず確認。

対策例

- なんらかの方法でデータを統合してゼロを減らす
 - 日単位ではなく週単位にデータをまとめる
 - 類似カテゴリのデータを加算
- 関連時系列データを追加する
 - プロモーション実施日

予測がむずかしいデータ

数日間ゼロが続くようなデータは予測が非常に困難。
ゼロが続くデータを使っていないかをまず確認。

対策例

- なんらかの方法でデータを統合してゼロを減らす
 - 日単位ではなく週単位にデータをまとめる
 - 類似カテゴリのデータを加算
- 関連時系列データを追加する
 - プロモーション実施日

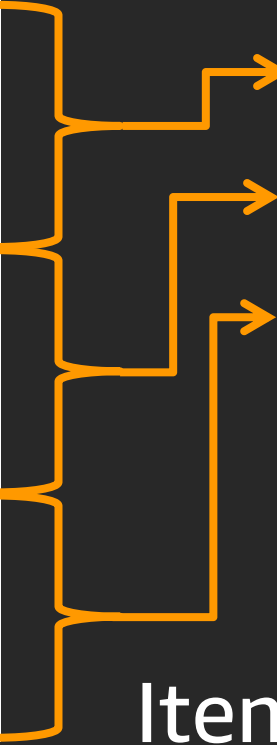
データ統合の例 – データを週単位に

タイムスタンプ	アイテムID	販売数
2020/6/21	Item_1	0
2020/6/22	Item_1	0
2020/6/23	Item_1	2
2020/6/24	Item_1	0
2020/6/25	Item_1	0
2020/6/26	Item_1	0
2020/6/27	Item_1	0
2020/6/28	Item_1	1
2020/6/29	Item_1	0

タイムスタンプ	アイテムID	販売数
2020/6/21	Item_1	2
2020/6/28	Item_1	1
2020/7/5	Item_1	1

1週間分の販売数を積算

データ統合の例 – 類似アイテムのデータを統合

タイムスタンプ	アイテムID	販売数	
2020/6/21	Item_1	0	
2020/6/21	Item_2	1	
2020/6/22	Item_1	2	
2020/6/22	Item_2	0	
2020/6/23	Item_1	0	
2020/6/23	Item_2	1	

タイムスタンプ	アイテムID	販売数
2020/6/21	Item_1	1
2020/6/22	Item_1	2
2020/6/23	Item_1	1

Item_1 と Item_2 の販売数を加算

予測精度を上げるには関連データが重要

関連時系列：予測したいデータに連動するデータ

- 祝日に連動するデータか？
- プロモーションを実施しているか？
- 店舗の近くで開催されるイベントが影響するか？

項目メタデータ：商品の特徴を示すデータ




- 商品のジャンル
- 商品の価格
- 商品の生産地

ドメイン知識を駆使して、
関連データを見つける。

運用フェーズにおける Tips

モデルの再学習なしに最新データで予測

モデルを再学習（Predictor の作成）なしに最新の予測を行う。

-  学習に使用したデータ
-  追加で予測に使用するデータ
-  予測したデータ

1 ~ N 日目のデータ

N + 1

1 ~ N 日目のデータ

N + 1

N + 2

1 ~ N 日目のデータ

N + 1 ~ N + 2




N + 3

1 ~ N + 3 日目のデータ

N + 4

モデルの再学習なしに最新データで予測

モデルを再学習（Predictor の作成）なしに最新の予測を行う。

-  学習に使用したデータ
-  追加で予測に使用するデータ
-  予測したデータ

1 ~ N 日目のデータ

N + 1

1 ~ N 日目のデータ

N + 1

N + 2

1 ~ N 日目のデータ

N + 1 ~ N + 2




N + 3

1 ~ N + 3 日目のデータ

N + 4

モデルの再学習なしに最新データで予測

モデルを再学習（Predictor の作成）なしに最新の予測を行う。

-  学習に使用したデータ
-  追加で予測に使用するデータ
-  予測したデータ

1 ~ N 日目のデータ

N + 1

1 ~ N 日目のデータ

N + 1

N + 2

1 ~ N 日目のデータ

N + 1 ~ N + 2




N + 3

1 ~ N + 3 日目のデータ

N + 4

モデルの再学習なしに最新データで予測

モデルを再学習（Predictor の作成）なしに最新の予測を行う。

-  学習に使用したデータ
-  追加で予測に使用するデータ
-  予測したデータ

1 ~ N 日目のデータ

N + 1

1 ~ N 日目のデータ

N + 1

N + 2

1 ~ N 日目のデータ

N + 1 ~ N + 2




N + 3

1 ~ N + 3 日目のデータ

N + 4

モデルの再学習なしに最新データで予測

モデルを再学習（Predictor の作成）なしに最新の予測を行う。

-  学習に使用したデータ
-  追加で予測に使用するデータ
-  予測したデータ

1 ~ N 日目のデータ

N + 1

1 ~ N 日目のデータ

N + 1

N + 2

1 ~ N 日目のデータ

N + 1 ~ N + 2




N + 3

1 ~ N + 3 日目のデータ

N + 4

モデルの再学習なしに最新データで予測

モデルを再学習（Predictor の作成）なしに最新の予測を行う。

-  学習に使用したデータ
-  追加で予測に使用するデータ
-  予測したデータ

1 ~ N 日目のデータ

N + 1

1 ~ N 日目のデータ

N + 1

N + 2

1 ~ N 日目のデータ

N + 1 ~ N + 2

N + 3

1 ~ N + 3 日目のデータ

N + 4

Amazon Forecast の料金

<前提>

2000 個の商品を 50 店舗で販売。

データセットの合計サイズは 5GB で、モデルの学習には 20 時間かかった。

予測結果は P10, P50, P90 の 3つの分位数 を取得。

合計 \$185.24	料金表	料金の試算
データ量	1GB あたり \$0.088	$\$0.088 \times 5 \text{ GB} = \0.44
モデルの学習 (Predictor 作成)	1時間あたり \$0.24	$\$0.24 \times 20 = \4.80
予測の実行 (Forecast 作成)	1000 予測あたり \$0.60	2000 商品 x 50 店舗 = 100k 予測数 $\$0.60 \times 100\text{k 予測数} / 1000 \times 3 \text{ 分位数}$ = \$180

Amazon Forecast の料金

<前提>

2000 個の商品を 50 店舗で販売。

データセットの合計サイズは 5GB で、モデルの学習には 20 時間かかった。

予測結果は P10, P50, P90 の 3つの分位数 を取得。

	料金表	料金の試算
合計 \$185.24		
データ量	学習頻度を減らす	088 x 5 GB = \$0.44 分位数を最小限の数にする
モデルの学習 (Predictor 作成)	1時間あたり \$0.24	\$0.24 x 20 = \$4.80
予測の実行 (Forecast 作成)	1000 予測あたり \$0.60	2000 商品 x 50 店舗 = 100k 予測数 \$0.60 x 100k 予測数 / 1000 x 3 分位数 = \$180

Amazon Forecast まとめ

Amazon Forecast まとめ

- ビジネス価値を生むために予測する
- Amazon Forecast を使うことで、
機械学習の深い知識がなくても時系列データ予測が可能
- 予測精度を向上させるには、使用するデータの工夫が必要
- 運用フェーズでは、モデルの学習と予測の頻度を最適化する

参考 URL

- AWS Black Belt Online Seminar: Amazon Forecast
<https://aws.amazon.com/jp/blogs/news/webinar-bb-amazon-forecast-2020/>
- AWS ブログ：学習済み Predictor を再利用して最新の予測を更新する
<https://aws.amazon.com/jp/blogs/news/amazon-forecast-rolling-forecast/>
- Amazon Forecast & Personalize ハンズオンセミナー開催報告
<https://aws.amazon.com/jp/blogs/startup/event-report-amazon-forecast-and-personalize/>

関連セッション

- **AWS-02**

突然「機械学習プロジェクトをやってくれ！」
と言われて困ってませんか？ビジネスに適用するための
機械学習システム開発プロジェクトの進め方

- **AWS-42**

AWSを使った機械学習の始め方

AWS トレーニングと認定で機械学習を学ぶ

AWS のエキスパートが作成した教材で機械学習スキルの習得とスキルチェック



ビジネス意思決定、データ基盤エンジニア、データサイエンティスト、開発者のための機械学習コース



65以上ある無料のオンデマンドデジタルコースやAWS のエキスパートによるハンズオンラボでの演習

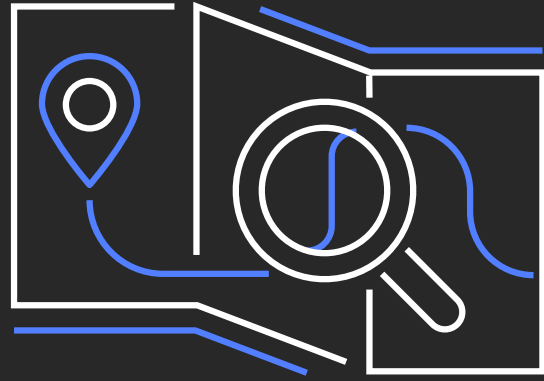


AWS 認定 機械学習の取得 – 機械学習モデルを構築、調整、デプロイするスキルを確かめる

ML 学習コースはこちら <https://aws.training/ML>

Thank you!

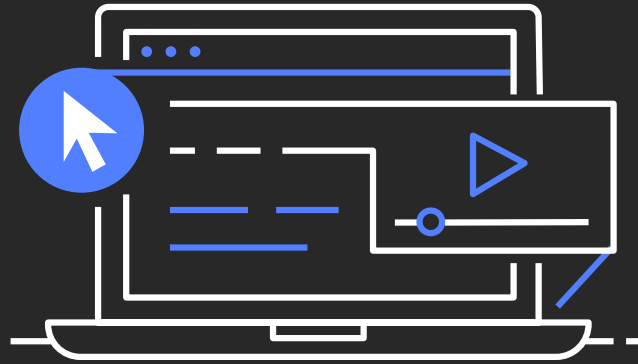
AWS トレーニングと認定



クラウド人材の育成

AWS トレーニングを活用し、
ビジネスを牽引する人材の育成
と組織作りを促進する

[AWS トレーニング活用事例 »](#)



自習コンテンツの活用

ウェビナーやのデジタルトレ
ーニングを受講して、個人のスキ
ルアップを目指す

[AWS デジタルトレーニング »](#)



AWS 認定取得を目指す

認定取得を目指して知識を底上
げし、AWS の経験とスキルを
証明する

[AWS 認定の詳細 »](#)

学習パスをお探しの方に

日本語版ランプアップガイドを公開しました。AWS ウェブページ、無料のデ
ジタルトレーニング、クラスルームコース、動画、ホワイトペーパー、認定等
を含んだ、9 種の役割別学習ガイドをご覧ください。 [詳細を見る »](#)

aws.amazon.com/training