

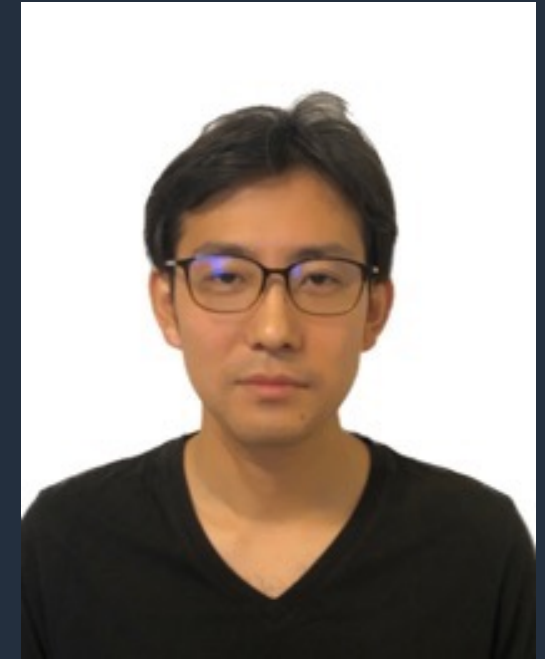


Amazon SageMaker を使って 今すぐ機械学習をはじめる方法

Amazon Web Services Japan
Solutions Architect
Masahiro Imai

自己紹介

Masahiro Imai (今井 真宏)



- シニアソリューションアーキテクト
 - スマートシティ/ヘルスケア/MLスペシャリスト
 - 今までの経験
 - MLテック企業CTO
 - Finテック企業CTO
 - 慶應義塾大学 Visited Researcher
- 好きなAWSサービス
 - Amazon SageMaker, Amazon Lookout for Vision

Tens of thousands of customers using Amazon SageMaker



50+
capabilities added
since last year

AWS の機械学習スタック

機械学習機能において最も幅広く完全なセットを提供

AI サービス

HEALTH AI



NEW

Amazon HealthLake



Amazon Transcribe for Medical



Amazon Comprehend for Medical

INDUSTRIAL AI



NEW

AWS Panorama + Appliance



NEW

Amazon Monitron



NEW

Amazon Lookout for Equipment



NEW

Amazon Lookout for Vision

ANOMALY DETECTION



NEW

Amazon Lookout for Metrics

CODE AND DEVOPS



NEW

Amazon DevOps Guru



Amazon CodeGuru

VISION



Amazon Rekognition

SPEECH



Amazon Polly



Amazon Transcribe
+Medical

TEXT



Amazon Comprehend
+Medical



Amazon Translate



Amazon Textract

SEARCH



Amazon Kendra

CHATBOTS



Amazon Lex

PERSONALIZATION



Amazon Personalize

FORECASTING



Amazon Forecast

FRAUD



Amazon Fraud Detector

CONTACT CENTERS



Contact Lens

Voice ID

For Amazon Connect

ML サービス



Amazon SageMaker

Label data

NEW

Aggregate & prepare data

NEW

Store & share features

Auto ML

Spark/R

NEW

Detect bias

Visualize in notebooks

Pick algorithm

Train models

Tune parameters

NEW

Debug & profile

Deploy in production

Manage & monitor

NEW

CI/CD

Human review

SAGEMAKER STUDIO IDE

NEW: SageMaker JumpStart

NEW: Model management for edge devices

フレームワーク & インフラストラクチャ



Deep Learning AMIs & Containers

GPUs & CPUs

Elastic Inference

Trainium

Inferentia

FPGA

Amazon SageMaker



Amazon SageMaker overview

Amazon SageMaker

PREPARE

SageMaker Ground Truth

Label training data for machine learning

SageMaker Data Wrangler **NEW**

Aggregate and prepare data for machine learning

SageMaker Processing

Built-in Python, BYO R/Spark

SageMaker Feature Store **NEW**

Store, update, retrieve, and share features

SageMaker Clarify **NEW**

Detect bias and understand model predictions

BUILD

SageMaker Studio Notebooks

Jupyter notebooks with elastic compute and sharing

Built-in and Bring your-own Algorithms

Dozens of optimized algorithms or bring your own

Local Mode

Test and prototype on your local machine

SageMaker Autopilot

Automatically create machine learning models with full visibility

SageMaker JumpStart **NEW**

Pre-built solutions for common use cases

TRAIN & TUNE

Managed Training

Distributed infrastructure management

SageMaker Experiments

Capture, organize, and compare every step

Automatic Model Tuning

Hyperparameter optimization

Distributed Training **NEW**

Training for large datasets and models

SageMaker Debugger **NEW**

Debug and profile training runs

Managed Spot Training

Reduce training cost by 90%

DEPLOY & MANAGE

Managed Deployment

Fully managed, ultra low latency, high throughput

Kubernetes & Kubeflow Integration

Simplify Kubernetes-based machine learning

Multi-Model Endpoints

Reduce cost by hosting multiple models per instance

SageMaker Model Monitor

Maintain accuracy of deployed models

SageMaker Edge Manager **NEW**

Manage and monitor models on edge devices

SageMaker Pipelines **NEW**

Workflow orchestration and automation

SageMaker Studio

Integrated development environment (IDE) for ML

データ準備/保管

開発環境

学習環境

推論環境

MLOps

エッジ推論

Amazon SageMaker 主な利点

最も完全なエンドツー
エンドの ML サービス



ML 開発を加速する

ML 開発ライフサイクル全体をカバーする20以上のツール



データサイエンティストの生産性を高める

世界初の統合開発環境 (IDE)



コスト削減

ML 用に最適化された統合機能を備えたカスタム統合コードを作成するコストを排除



Amazon SageMaker 機械学習フルマネージドサービス



機械学習をすぐに開始できる開発環境



開発と処理、学習、推論提供環境の分離



ユースケースに応じて利用できる柔軟性



機械学習をすぐに開始できる開発環境



開発と学習、推論提供環境の分離



ユースケースに応じて利用できる柔軟性

Amazon SageMaker Studio

DEMO

SageMaker Studio Notebook

機械学習の開発を
迅速に開始
ML統合開発環境

The screenshot displays the Amazon SageMaker Studio interface. The main window shows a Jupyter Notebook with the following content:

- Have the predictor variable in the first column
- Not have a header row

But first, let's convert our categorical features into numeric features.

```
[ ]: model_data = pd.get_dummies(churn)
      model_data = pd.concat([model_data['Churn?_True'], model_data.drop(['Churn?_True'], axis=1)], axis=1)
```

And now let's split the data into training, validation, and test sets. This will help prevent us from overfitting the model, and allow us to test the models accuracy on data it hasn't already seen.

```
[ ]: train_data, validation_data, test_data = np.split(model_data.sample(frac=1), [int(model_data.shape[0]*0.7), int(model_data.shape[0]*0.85)])
      train_data.to_csv('train.csv', header=False, index=False)
      validation_data.to_csv('validation.csv', header=False, index=False)
```

Now we'll upload these files to S3.

```
[ ]: boto3.Session().resource('s3').Bucket(bucket).Object(os.path.join(prefix, 'train.csv')).upload(train_data.to_csv(index=False))
      boto3.Session().resource('s3').Bucket(bucket).Object(os.path.join(prefix, 'validation.csv')).upload(validation_data.to_csv(index=False))
```

On the right side, there are two panels:

- Trial Component Chart:** A line graph showing training loss over 6 periods. The y-axis is labeled 'train:loss_last' and ranges from 0.0 to 0.4. The x-axis is labeled 'period' and ranges from 0 to 6. Multiple lines represent different trials, showing a general downward trend in loss over time.
- Trial Component List:** A table listing trial components. It shows 10 rows selected, all with a status of 'Completed' and a type of 'Training job'.

Status	Experiment	Type
✓ Completed	customer-churn-predi...	Training job
✓ Completed	customer-churn-predi...	Training job
✓ Completed	customer-churn-predi...	Training job
✓ Completed	customer-churn-predi...	Training job

<https://aws.amazon.com/jp/sagemaker/studio/>

DEMO

SageMaker JumpStart

機械学習アプリケーションを
簡単かつ迅速に
マーケットに投入



15以上の一般的なMLユースケースの構築済みソリューション

すぐに使用できるソリューションを活用、または特定のビジネス上の問題に合わせてカスタマイズします



150を超えるオープンソースモデルの展開時間を短縮

人気のあるModel ZooのMLモデルとアルゴリズムをワンクリックでデプロイ可能に



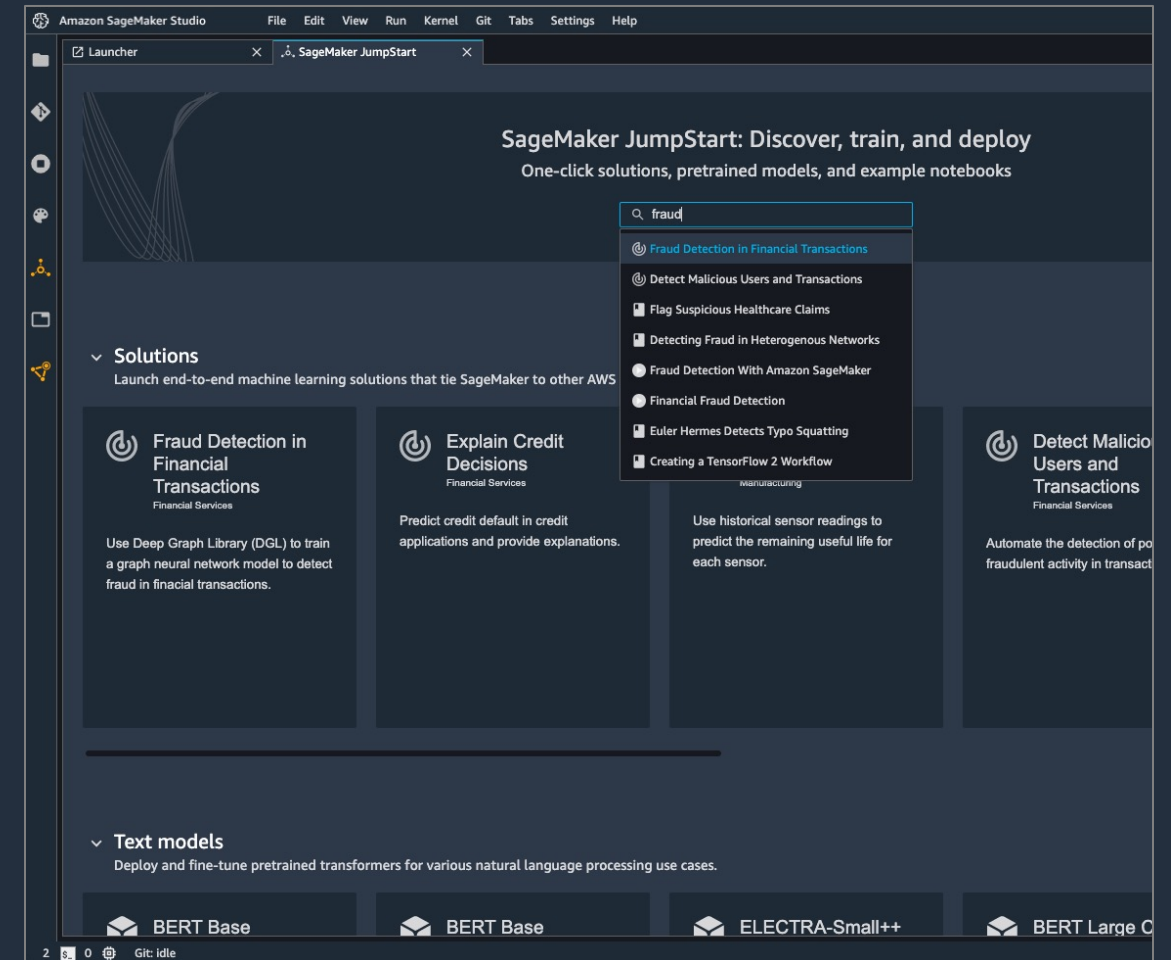
数回クリックで始めましょう

事前に構築されたソリューション、Model ZooのMLモデルとアルゴリズムを使用して、MLアプリケーションを簡単にマーケットに投入し、コンテンツを開始します

Amazon SageMaker JumpStart

機械学習アプリケーションのテンプレートを利用して、迅速にマーケットにローンチ

- 15以上の一般的なMLユースケースに対し、カスタム可能な構築済みソリューションを提供
- 150を超えるOSSモデルを展開時間を短縮、ワンクリックでデプロイ可能に
- 事前に構築されたソリューション、Model ZooのMLモデルとアルゴリズムを使用して、MLアプリケーションを簡単にマーケットに投入し、コンテンツを開始



<https://aws.amazon.com/jp/sagemaker/getting-started/>

Amazon SageMaker JumpStart 組み込みソリューション

予知保全

[Predictive maintenance for manufacturing >](#)

[Predictive maintenance for vehicle fleets >](#)

需要予測

[Demand forecasting with deep learning >](#)

不正検出

[Detect malicious users and transactions >](#)

[Fraud detection in financial transactions using deep graph library >](#)

信用リスクの予測

[Explain credit decisions >](#)

+11.11% ≥
+10.28%
+10.22%
+8.89%
+8.66%
+8.47%

ドキュメントからデータを抽出・分析

[Document summarization, entity, and relationship extraction >](#)

[Handwriting recognition >](#)

[Filling in missing values in tabular records >](#)

[Differential privacy for sentiment classification >](#)

コンピュータビジョン

[Product defect detection in images >](#)

自動運転

[Visual perception with active learning >](#)

パーソナライズレコメンデーション

[Entity resolution in identity graphs >](#)

[Purchase modeling >](#)

チャーン予測

[Churn prediction with text >](#)

ソリューションの詳細はこちら:

<https://aws.amazon.com/sagemaker/getting-started/>

組み込みアルゴリズムの詳細はこちら:

https://docs.aws.amazon.com/ja_jp/sagemaker/latest/dg/algos.html

DEMO

Amazon SageMaker Pipelines

完全に自動化された
ML ワークフローを
スケーラブルに構築



ML 開発の加速

数回のクリックで、完全に自動化された ML ワークフローを作成でき、数ヶ月のコーディングを数時間に削減可能



何百もの生成されたモデルを自動で追跡

生成されたモデルを自動で追跡し、構造化された監査証跡を維持することで手動プロセスを排除



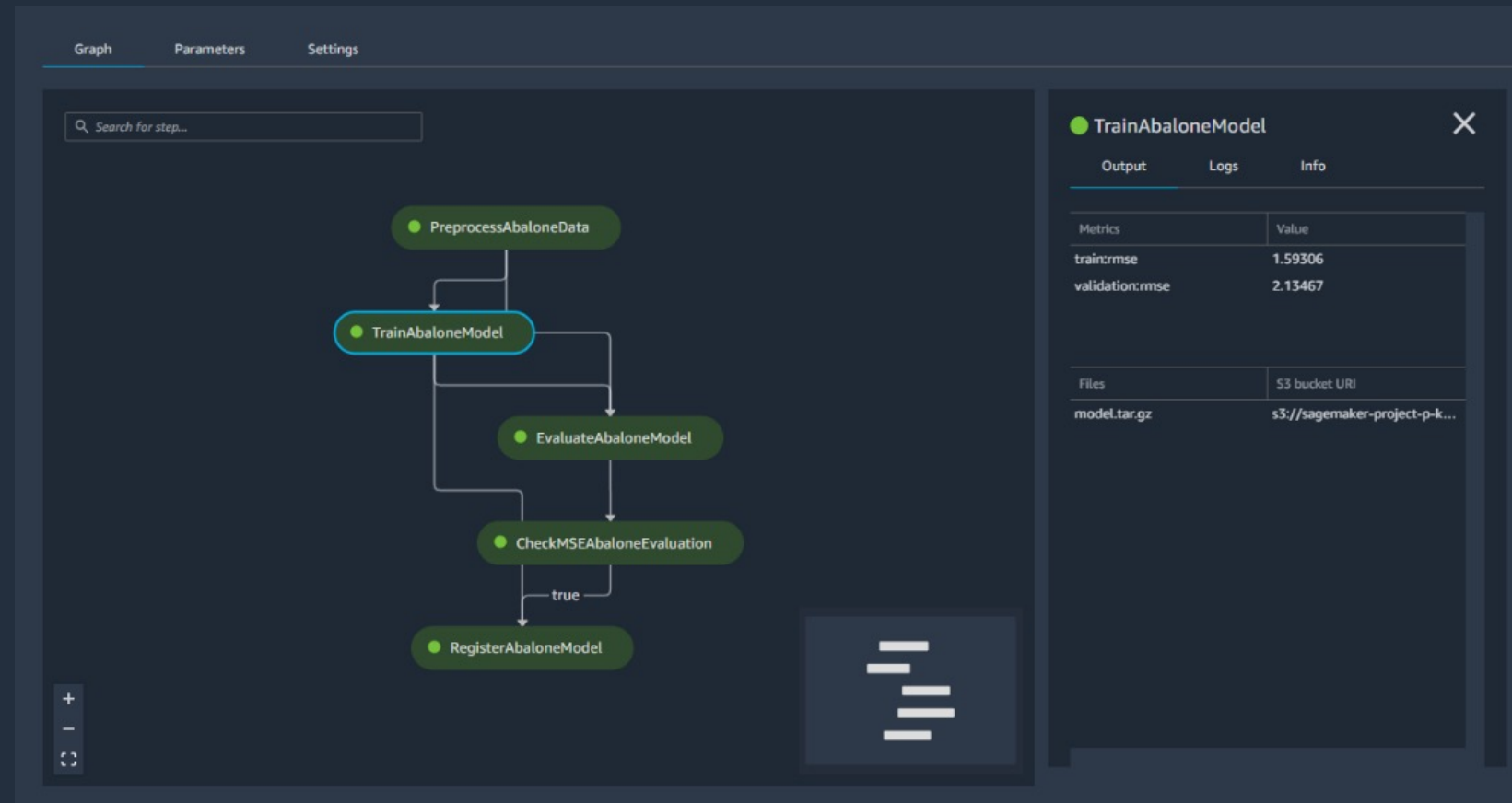
本番環境で動く ML モデルを数千にスケール

ビルトインテンプレートを利用して CI / CD パイプラインをセットアップし、ワークフローを自動化することで ML モデルをスケーラブルにデプロイ可能

Amazon SageMaker Pipelines と Project

機械学習プロジェクトのパイプラインと成果物の階層管理

- 機械学習ワークロードの CI/CD を実現
- 機械学習ワークフローのデータロードや学習処理、モデルの最適化などの一連の処理ステップを任意のタイミングや所定の時間に実行できる
- 本番環境にデプロイするのに最適なモデルを選択したり、モデルの自動追跡が可能。
- 各ステップの処理結果はSageMaker Experimentsで記録され、モデルの出来映えや学習パラメータなどを視覚化できる
- 定義されたワークフローは再利用や共有が容易なので、モデルの再生成や改善の際にも有益



<https://aws.amazon.com/sagemaker/pipelines/>

DEMO

SageMaker Data Wrangler

機械学習のためのデータを準備する最速で簡単な方法



データを素早く選択してクエリを実行

複数のソースからのデータをサポート



組み込みのデータ変換機能でデータを簡単に変換

組み込みのデータ変換を使用して、機械学習のために生データを特徴量に変換



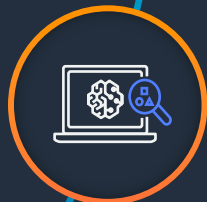
データ変換をカスタマイズ

PySpark、SQL、Pandasで独自のカスタム変換を行う柔軟性



データを視覚的に理解

外れ値や極端な値を素早く検出 - コードを書かなくてもできます



MLモデルの精度を素早く推定

MLモデルの精度を妨げる可能性のあるデータ準備ワークフローの潜在的な問題を診断



ワンクリックでデータ準備のワークフローをデプロイ

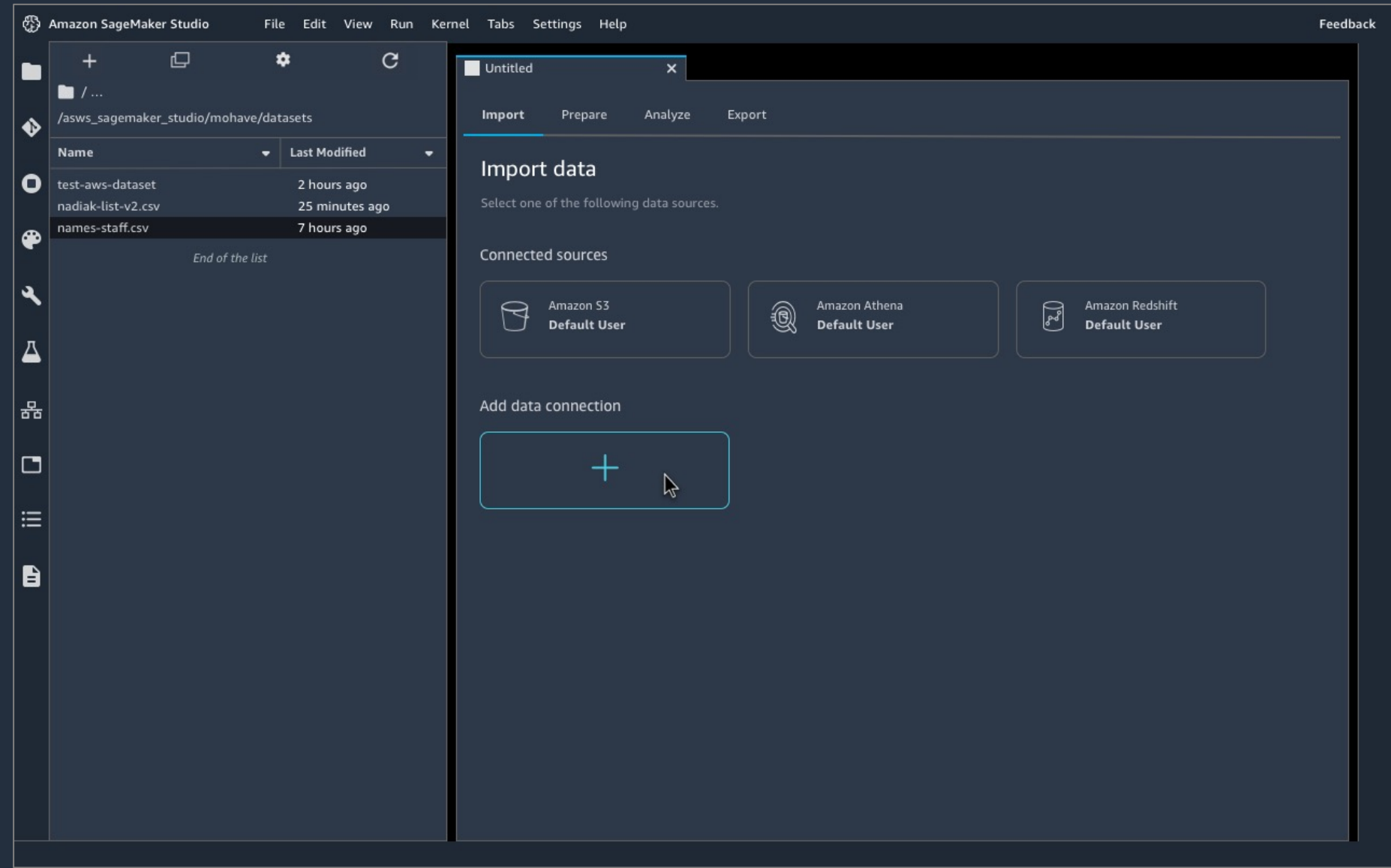
データ準備のワークフローのすべてのステップを単一の視覚的なインターフェースで管理し、ワークフローを本番環境に迅速に運用可能にします

Amazon SageMaker Data Wrangler

データを素早く選択してクエリを実行する

データの前処理を簡潔に実行し、機械学習ワークフローを実現

- Amazon Athena, Amazon Redshift, AWS Lake Formation, Amazon S3からデータをインポート
- Amazon S3やSageMaker Feature Storeから特徴量をエクスポート
- データソースに対してクエリを記述、ETLの処理ファイルを自動生成
- CSVファイル、Parquetファイル、データベーステーブルなどのインポート形式に対応



<https://aws.amazon.com/jp/sagemaker/data-wrangler/>

SageMaker Feature Store

リアルタイム推論・バッチ推論のための特徴量を安全に保存・発見・共有



バッチでもストリーミングでも特徴量を投入可能

特徴量を保存するために高い write スループットを兼ね備えている



オンラインとオフラインの特徴量

リアルタイム推論に利用可能なオンライン特徴量と、モデル学習とバッチ推論に利用される蓄積データのためのオフライン特徴量を扱える



特徴量のメタデータとデータカタログ

特徴量のためのメタデータを保存し、簡単に問い合わせで特徴量データを取り出すための自動のデータカタログ化を活用できる



特徴量の発見と再利用

特徴量を発見するための検索の仕組みを提供している



セキュリティとアクセス制御

特徴量のデータとメタデータへのアクセス制御と、保存データの暗号化 (encryption at rest), 閉域網での利用を可能にする VPC および Private Link

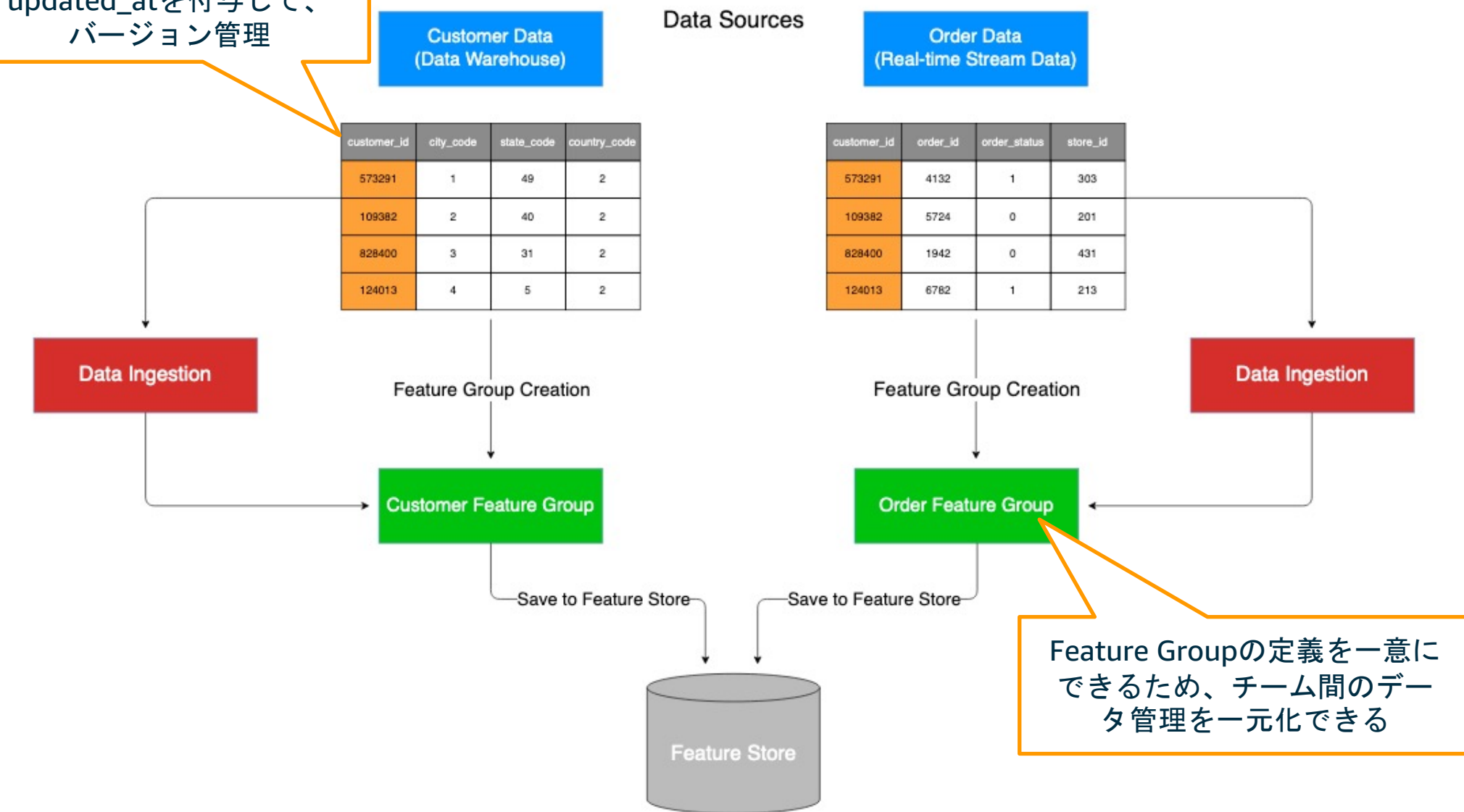


完全にマネージドなサービス

オンライン特徴量は低レイテンシの保存領域にキャッシュされる。学習と推論時のデータ不一致による性能劣化を避けるため、オンラインとオフラインの本尊領域は一貫性を保持する

学習と推論のための SageMaker Feature Store

データに一意的IDと updated_at を付与して、バージョン管理



2種類のストアでユースケースに応じた特徴量ストアを利用

Online feature store

- 主にリアルタイム推論へ利用される
- 特徴量の最新のコピーのみ保持
- 高い write スループット
- オンライン予測のための低ミリ秒のレイテンシ

Offline feature store

- 主にバッチ推論とモデル学習のために利用される
- 特徴量の履歴のための保存領域を提供
- 高い write スループット
- 15分以内に書き込み後の読み込み整合性 (read after write consistency) が得られる

<https://aws.amazon.com/jp/sagemaker/feature-store/>



Amazon SageMaker 機械学習フルマネージドサービス



機械学習をすぐに開始できる開発環境

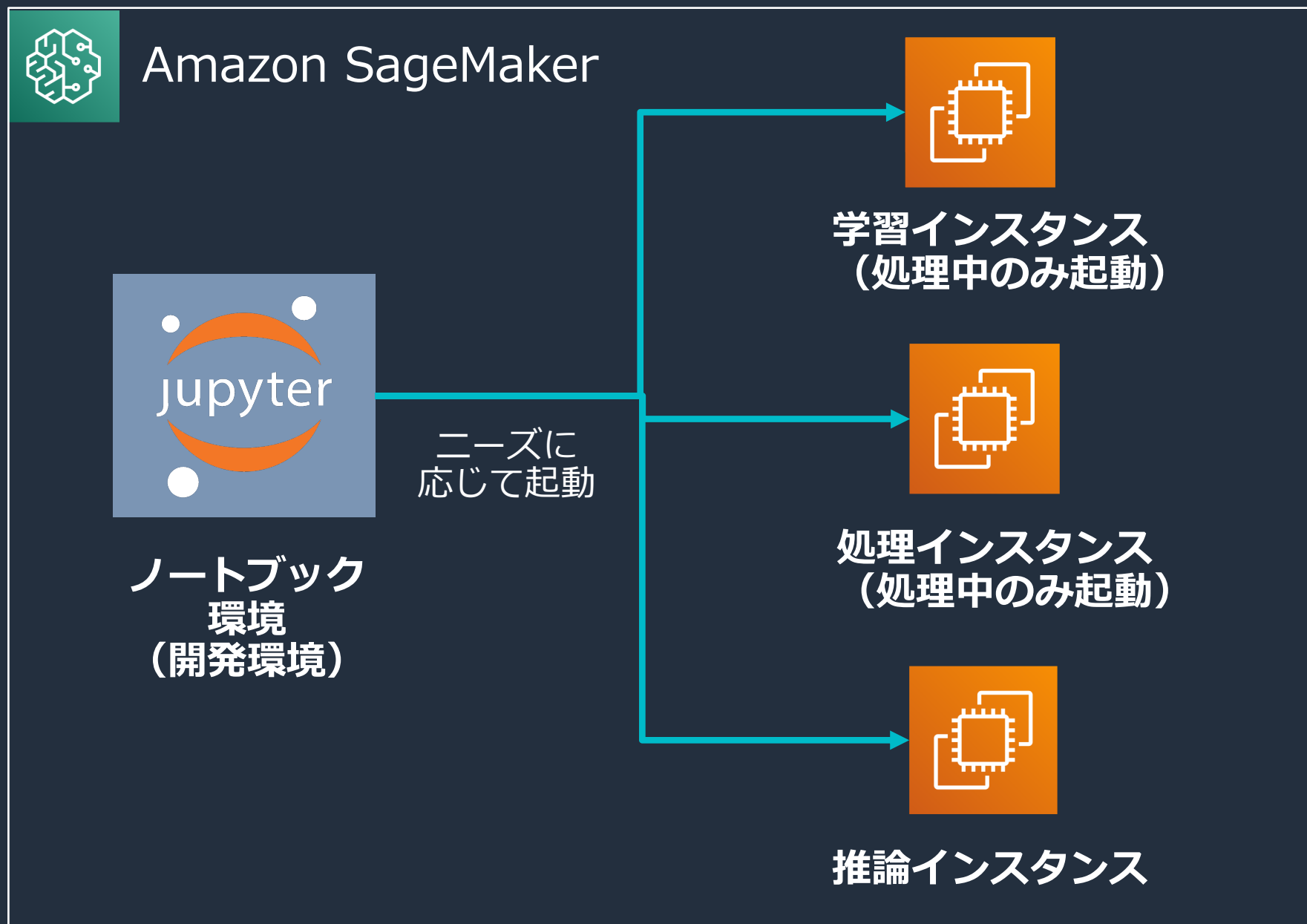


開発と学習、推論提供環境の分離

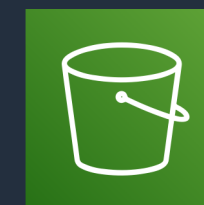


ユースケースに応じて利用できる柔軟性

開発と学習、推論提供環境の分離



学習、処理 (Processing)、推論で環境が分離 (疎結合) され、ニーズに合わせた拡張が可能



学習データ
モデル

S3



学習・推論
コンテナ

ECR

TensorFlow / mxnet

PYTORCH

<https://aws.amazon.com/sagemaker/>

The Jupyter Trademark is registered with the U.S. Patent & Trademark Office.

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



Amazon SageMaker 機械学習フルマネージドサービス



機械学習をすぐに開始できる開発環境



開発と学習、推論提供環境の分離



ユースケースに応じて利用できる柔軟性

ユースケースに応じて利用できる柔軟な環境

1. データの前処理、開発環境としてのみ利用

データ準備/保管

開発環境

学習環境

推論環境

MLOps

エッジ推論

2. 機械学習のトレーニングに利用する環境として利用

データ準備/保管

開発環境

学習環境

推論環境

MLOps

エッジ推論

3. 学習済みモデルの推論環境として利用

データ準備/保管

開発環境

学習環境

推論環境

MLOps

エッジ推論



Amazon SageMaker 機械学習フルマネージドサービス

まとめ

- Amazon SageMaker は、機械学習のワークフローを効率化
- 開発者・研究者が必要とする開発環境を迅速に提供
- 必要な計算リソースを柔軟に利用
- ユースケースに応じてご利用サービスを選択

Thank you

