



アップデート紹介とちょっぴり DIVE DEEP する AWS の時間 - AI/ML

Amazon SageMaker Canvas で Business User と Data Scientist の シームレスコラボを実現！

Xiaoli Shen

Solutions Architect

Amazon Web Services Japan

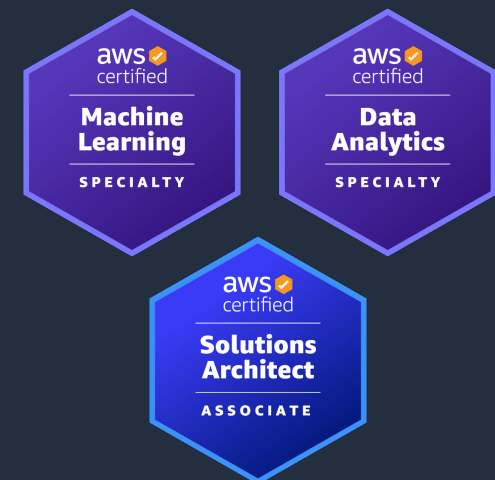
Xiaoli Shen (シン シオリ)

アマゾン ウェブ サービス ジャパン合同会社 技術統括本部
ソリューションアーキテクト

- AWS を活用してビジネス価値を創出するお客様にアーキテクティング面の技術支援を提供
- Low Code No Code (LCNC) ML と MLOps にフォーカス
- 前職は Sr. Software Engineer
- 好きな AWS サービス



Amazon SageMaker



Agenda

- 機械学習の民主化と Low Code No Code (LCNC) ML の価値
- Amazon SageMaker の LCNC ML オフアー
- Amazon SageMaker Canvas のコラボレーション機能
- Business User と Data Scientist コラボのデモ

話さないこと

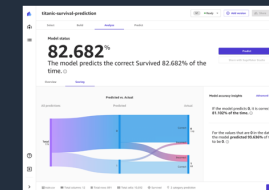
- Amazon SageMaker Canvas 利用開始のご紹介

<https://pages.awscloud.com/JAPAN-event-OE-Hands-on-for-Beginners-SageMaker-Canvas-2023-reg-event.html>

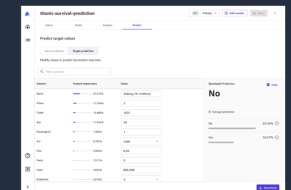
【AWS Hands-on for Beginners】 Amazon SageMaker Canvas ～ ノーコードで機械学習を始めよう ～



データ探索、準備



モデル構築、学習、評価



予測

広がる機械学習の世界



支出の増加

2024年までに世界の人工知能
への支出は1,100億ドルに達す
るだろう

—IDC



試験運用から本番運用へ

2024年末までに、75%の企
業がAIの試験運用から本番
運用へと移行するだろう

—Gartner



AI トランスフォーメーション

57%が、今後3年間でAIが組
織を変革すると回答

—Deloitte

機械学習へのチャレンジにおける課題

- 1 機械学習の実用化には時間がかかり、深い専門知識とコーディングのスキルが必要
- 2 機械学習のニーズは増える一方で、Data Science チームはフル稼働
- 3 Business User やアナリストは、機械学習の知識不足で短時間での習得は困難
- 4 ステックホルダー間でのコラボレーションにセキュアで統合的なワークスペースやツールが求められている

もし機械学習が民主化されると？

ビジネスの主要な成果を達成



業務改善
効率化
≡
コスト削減



多くの情報に
基づいた決断
≡
売上の向上

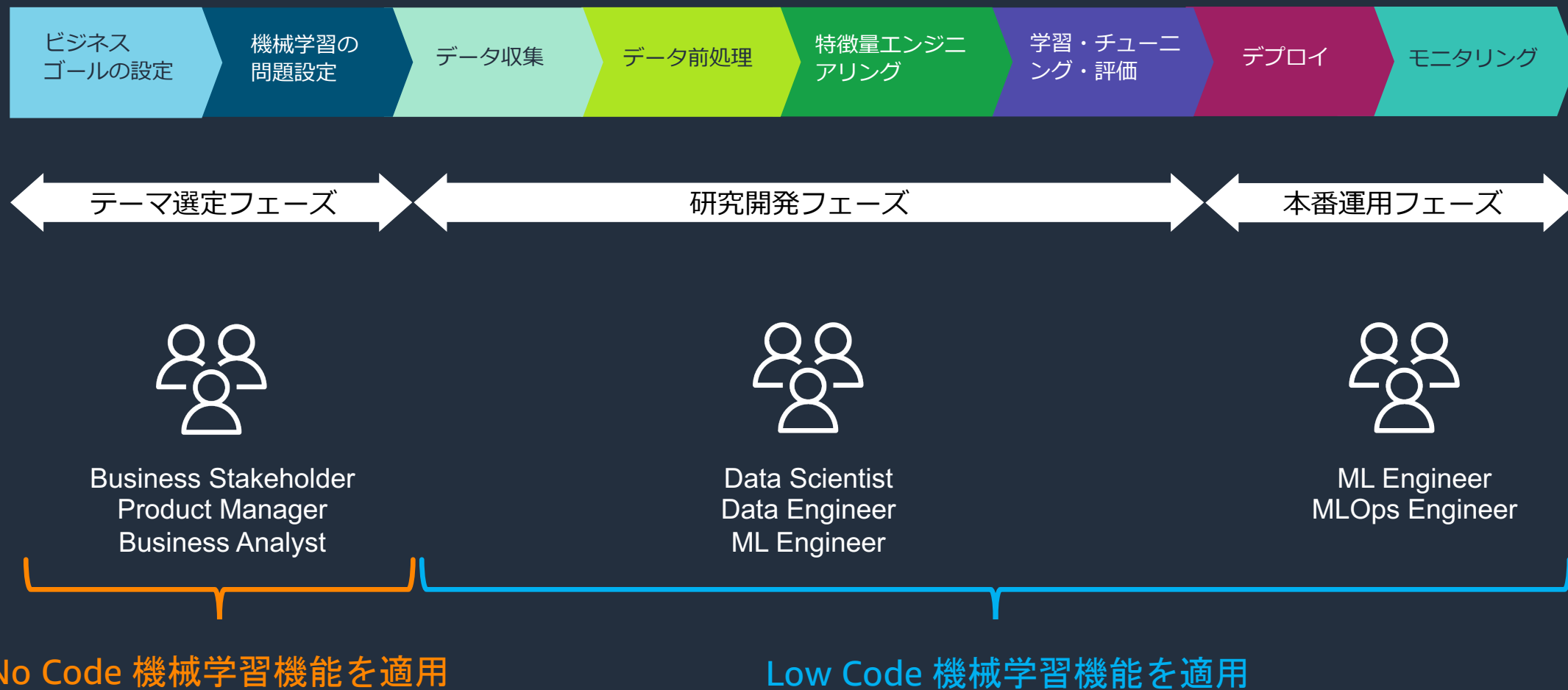


革新的なアイデアを開発し
新しいビジネスの創造
≡
新製品の市場投入による
ビジネス拡大

機械学習ライフサイクルのフェーズとロール



Low Code No Code (LCNC) ML の価値



AWS の LCNC ML オフナー

ビジネス
要件



データ準備 &
特徴量作成



モデル構築
学習・チューニング



モデルデプロイ
推論・モニタリング

Business User



Data Scientist
ML Engineer

Amazon SageMaker Canvas

機械学習の深い知識がなくても利用できるデータアナリスト向けのモデル構築ソリューション

Canvas

コードを書かずにドラッグ & ドロップで簡単に使える — ノーコードで、GUIでモデルを作成し予測を得られるAutoMLサービス

Amazon SageMaker Studio

ウェブベースの単一ビジュアルインターフェイスを提供し、データサイエンスチームの生産性を最大 10 倍向上

Data Wrangler

機械学習用のデータを集約して準備するのにかかる時間を数週間から数分に短縮

Autopilot

機械学習モデル構築のための重労働を排除、表形式のデータセットを提供、予測対象の列を選択するだけで、自動的にさまざまなソリューションを探索、最適なモデルを探索

JumpStart

人気の高いモデルのコレクション（「モデルズー」）および一般的なユースケースを解決するエンドツーエンドのソリューション、機械学習ワークフローを高速化

多数の
Deployment
オプション

コラボレーション

Amazon SageMaker Canvas

MLモデルを構築し正確な予測結果を取得

ノーコード: コードを書かずにドラッグ&ドロップで簡単に使える



機械学習(以下**ML**)へ素早くアクセス、データ統合



AutoMLビルトインでモデル構築と予測結果を取得



MLモデルを共有し、データサイエンティストチームとコラボレーション



従量課金モデルで、**TCO**削減

SageMaker Canvas で Business User と Data Scientists のコラボ

• シナリオ ① モデルコラボレーション

- Business User から SageMaker Canvas で作ったモデルを Data Scientist が利用する SageMaker Studio に共有し、フィードバックやアップデートをもらう

• シナリオ ② SageMaker Canvas に BYOM (独自のモデルを持参)

- Data Scientist から構築済みのモデルを Business User が利用する SageMaker Canvas に共有し、簡単に予測を行う

シナリオ ① モデルコラボレーション - デモ



前提条件：

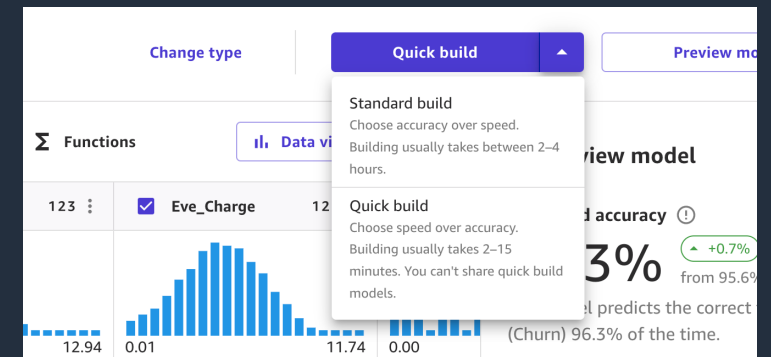
- Canvas ユーザーと Studio ユーザーが同じ Amazon SageMaker ドメインに属している
 - SageMaker ドメインには関連付けられた Amazon Elastic File System (EFS) ボリューム、承認されたユーザーのリスト、ならびにさまざまなセキュリティ、アプリケーション、ポリシー、および Amazon Virtual Private Cloud (VPC) の設定が含まれている
 - 各ユーザーは、ノートブック、Git リポジトリ、データファイル用に EFS 内の個人用およびプライベートホームディレクトリを利用できる
- SageMaker Canvas の Standard Build で構築したモデルのみが共有できる

Quick Build

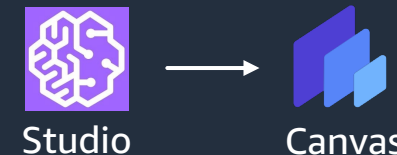
- 2~20分かかります
- 精度より速度を重視しています
- 最大5万行のデータをサポートします
- Canvas からログアウトするとビルドが中断される

Standard Build

- 2~4時間かかります
- 速度より最高の精度を重視しています
- Canvas からログアウトしてもビルドが中断されない



シナリオ ② BYOM (独自のモデルを持参) – デモ



前提条件

- Studio ユーザーと Canvas ユーザーが同じ Amazon SageMaker ドメインに属している
- SageMaker Autopilot、もしくは SageMaker JumpStart でビルド済み、または外部で作成されて SageMaker Model Registry に登録済みのモデルのみがインポート可能

Autopilot

機械学習モデル構築のための重労働を排除、表形式のデータセットを提供、予測対象の列を選択するだけで、自動的にさまざまなソリューションを探索、最適なモデルを探索

JumpStart

人気の高いモデルのコレクション (「モデルズー」) および一般的なユースケースを解決するエンドツーエンドのソリューション、機械学習ワークフローを高速化

Model Registry

機械学習モデルをカタログ化できるようにする専用サービス
モデルをモデルパッケージグループという論理グループにカタログ化し、モデルの増分バージョンをモデルパッケージバージョン (主にモデルアーティファクト) とメタデータとして保存

制約

- 表形式のデータでトレーニングされたモデルのみが共有可能
- 時系列モデルは共有不可

まとめ

- 機械学習の民主化と Low Code No Code (LCNC) ML の価値
- Amazon SageMaker の LCNC ML オフアー
- Amazon SageMaker Canvas のコラボレーション機能
 - モデルコラボレーション
 - BYOM (独自のモデルを持参)
- Business User と Data Scientist コラボのデモ



Thank you!