Vertex AI で実現する 大規模データの MLOps

Google Cloud AI/ML スペシャリスト 牧 允皓

Table of Contents

Vertex AI で実現する MLOps	01
Serverless Spark Interactive Development	02
BigQuery ML ∠ Vertex AI	03
Applied ML Summit Tabular Workflows	04

01

Vertex AI で実現する MLOps



Applications

Vision and Video Conversation

Language

Structured Data

Custom machine learning

Workbench

AutoML

NAS

Prediction

ML Metadata

Data Labeling

Training

Explainable Al

Feature Store

Model Monitoring

Experiments

Vizier (Optimization)

Pipelines

Matching Engine

Vertex Al Workbench の一般提供を発表

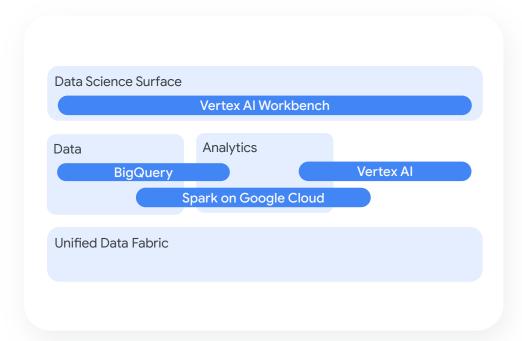
ツールが一箇所に集約された環境です

データに関する高速なワークフロー

アナリティクスと Veretx AI サービスが視覚的かつコードベースで統合されました

管理者が制御できるフルマネージドな環境

ポリシーの徹底やユーザ管理ができる、サーバレス/スケーラブルでエンタープライズ向けの JupyterLab 環境です



Vertex Al Workbench

データサイエンスのための統合 インターフェース



フルマネージドのコンピュートと管理者の制御

フルマネージド、スケーラブルでエンタープライズ向けのインフラであり、ポリシーの徹底やユーザ管理が可能です



データ処理を行う素早いワークフロー

ノートブック継続的モデル学習やパイプラインのデプロイに重要なツールです



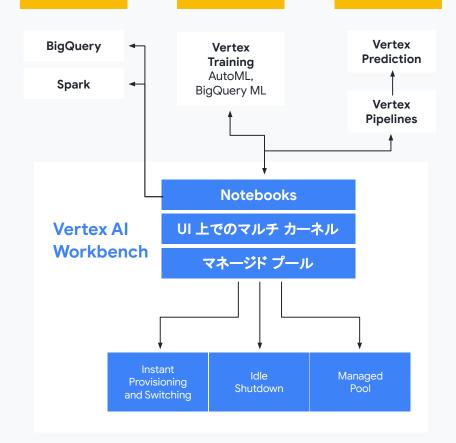
すぐに使えるインテグレーション

アナリティクスやVertex AI のサービスとシームレスな画面からコードベースで接続できます

分析

トレーニング

デプロイ



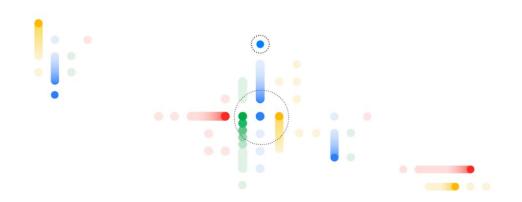
Vertex AI の アップデート

- Serverless Spark Interactive Development
- BigQuery ML と Vertex Al
- Tabular Workflows

02

Serverless Spark Interactive Development

Introducing Serverless Spark for interactive development



Abhishek Kashyap Sr. Product Manager, Google Cloud

April 8, 2022

At Google Cloud, we're committed to helping you build an open and integrated data platform that meets your specific business needs. We believe in order to build the world you want, you need to be able to use the tools you want, on a powerful and unified platform.

To that end, we're making Apache Spark even more powerful on Google Cloud, and creating a simplified experience across the platform. Data engineers, data analysts, and data scientists like Spark for its high speed data querying, analysis, and transformation with large data sets; but they want to use it from their interface of choice, without the need for custom integrations. And if you've used Spark, you know someone has to manage tuning, provisioning, and security, and make sure you're using what you pay for. Our customers have made it clear- they want that someone to be us. They want auto-tuned capabilities, easy provisioning, and an efficient cost basis without having to manually reconfigure every time load fluctuates.

Earlier this year, we made Serverless Spark generally available, creating the industry's

Google Cloud の ブログでも発表

Serverless Spark 上で Interactive に開発する機能がプライベート プレビューになりました。4月8日のブログで詳細が公開されており、本日は機能の紹介しデモを行います

出典:

https://cloud.google.com/blog/products/data-analytics/making-serverless-spark-even-more-powerful

Vertex Al Workbench 上の Spark ワークフロー

1クリックで Spark

データサイエンティストは Vertex Al Workbench から Spark を操作できます

ビルトインのセキュリティと認証

Vertex AI のセキュリティやユーザアクセスが 自動的に Spark に適用されます





Vertex Al Workbench

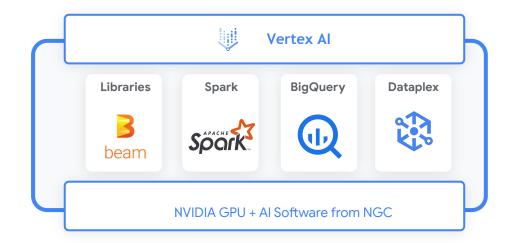
単一インターフェースでデータ処理、分析、機械学習ワークフローを カバーする次世代のNotebook

探索と分析を容易に

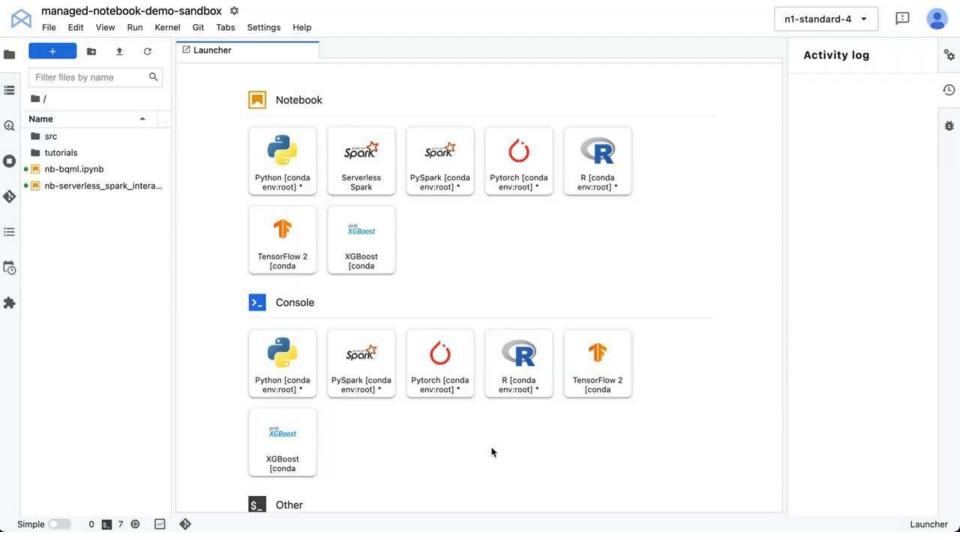
データへのアクセスがシンプルになり、ノートブックか母igQuery、Dataproc、Spark、Vertex AIを使って機械学習が扱えます

高速なプロトタイピングと機械学習モデル開発

Vertex AI トレーニングを使った強力な演算環境で、実験やプロトタイプ作成が実現でき、さらに大規模なデータも扱えるます



Demo



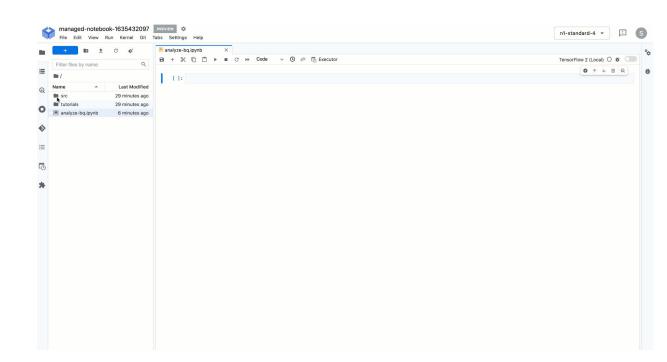
03

BigQuery ML & Vertex Al

Vertex Al Workbench から直接 BigQuery のデータに接続

Workbench ノートブック環境を使用し、探索的データ分析、可視化、そして特徴量エンジニアリングを行うことが可能です

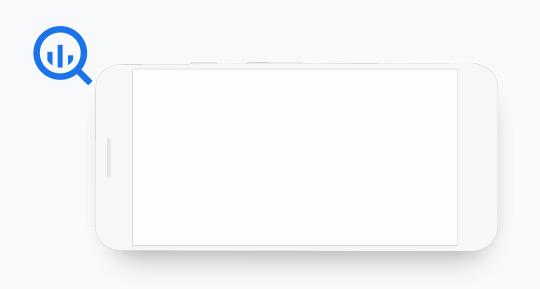
SQL クエリで BigQuery の データに直接アクセスした り、Python の Pandas Dataframe としてダウン ロードすることも可能です



BigQuery ML / AutoML / Vertex AI の比較

学習ツール	Data Analyst	ML Developer	Data Scientist	適したシナリオ	サービング	
Vertex Training			~	・以下のBigQuery ML、AutoMLに該当しない場合 ・既にオンプレや他のクラウド環境で学習が実行されていて、プラットフォームを横断しても一貫性が必要な場合は 適しています		
AutoML in Vertex Al			~	 ・課題がAutoMLで提供されるいずれかの機能に適している場合は、画面操作で実装できます ・Google Cloud から提供されている自然言語やビデオモデル、またエッジダウンロードに対応した視覚やテーブルモデルに適しています 	Vertex Prediction や BigQuery ML batch prediction	
BigQuery ML	~	•	~	 全てのデータがBigQuery に格納されている場合 SQL が最も使い慣れているユーザに向いています 公開されているBigQuery ML のモデル が解決しようとしている問題に適している場合 		

BigQuery ML によって 直接 AI を利用できます



- 1 SQL インターフェースと組み込まれた機能で素早くコーディングできます
- 2 BigQuery からデータを移すことなく学習できます
- 3 BigQuery のオペレーションと して**デプロイ**できます

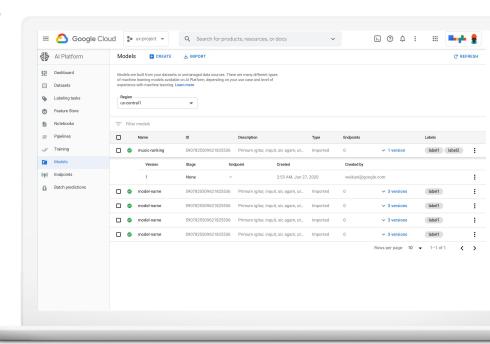
Vertex Al Model Registry を発表

Vertex AI や BigQuery ML のモデルを一 箇所で管理し統制できます

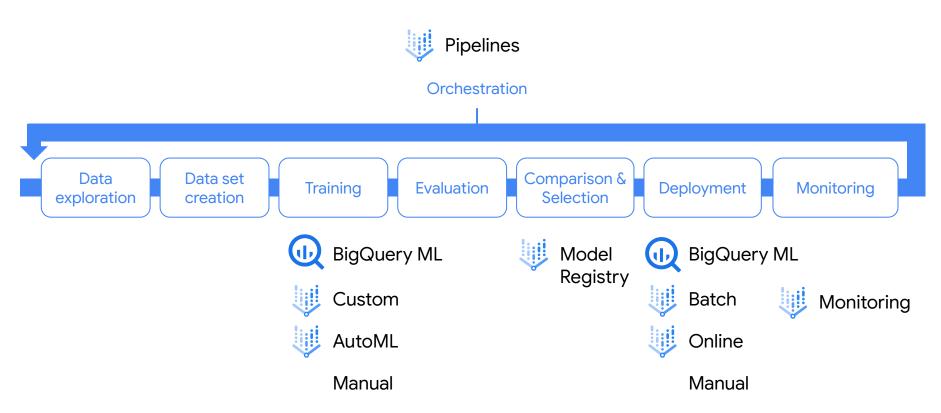
学習された機械学習モデルを登録、整理、監視、版管理できます

モデルのメタデータや デプロイするための依存関係 を保存できます

モデルを繰り返し改善するため **メトリックを保存し比 較** できるようになっています



Vertex AI による ML ジャーニー



Google Cloud の統一されたデータや AI の機能によりAI をエンタープラズで実装可能に

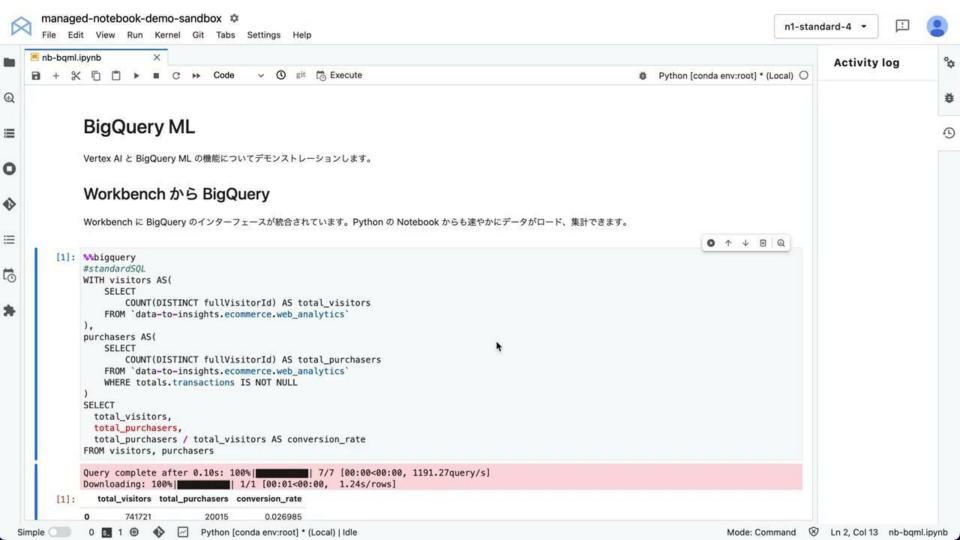
データ分析や AI 開発のための統合されたインターフェース

最小限のデータ移動で強力な AI システムの実装

AI 開発のための MLOps ツール



Demo



03

Applied ML Summit Tabular Workflows

多くのデータはテーブル型

Forecasting

Classification

Regression



需要予測

製品需要 在庫管理 コスト最適化 人材管理



マーケティング

購入頻度 離脱確率(尤度) コンバージョン促進 ライフタイム バリュー キャンペーン効果測定 製品配置



リソース最適化

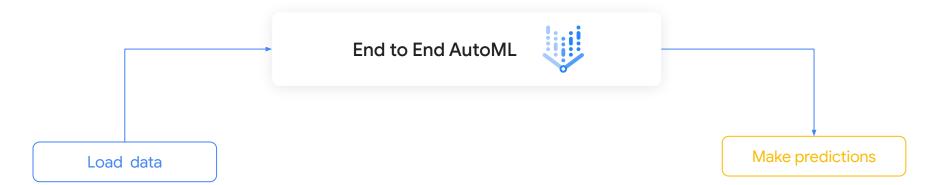
機材の故障予測 流通の需要/供給 機材の価値評価 従業員の離職予測 製品の寿命予測



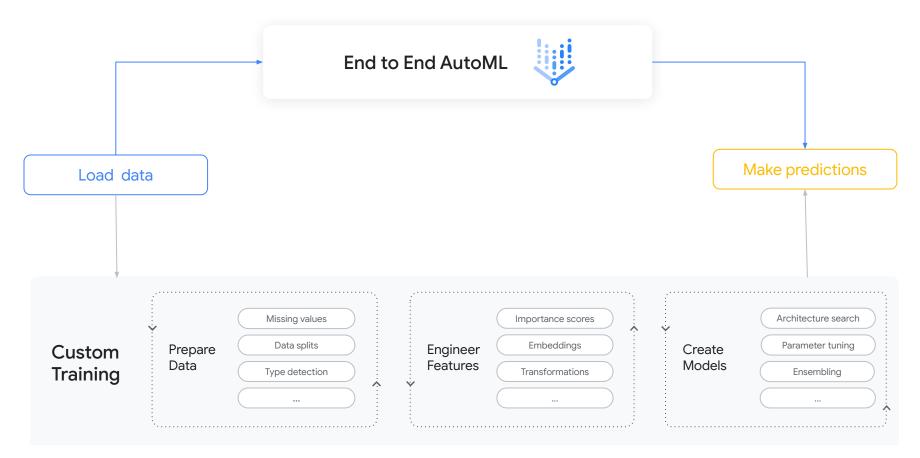
リスク管理

クレーム リスク 不正検知 デフォルト リスク 不渡

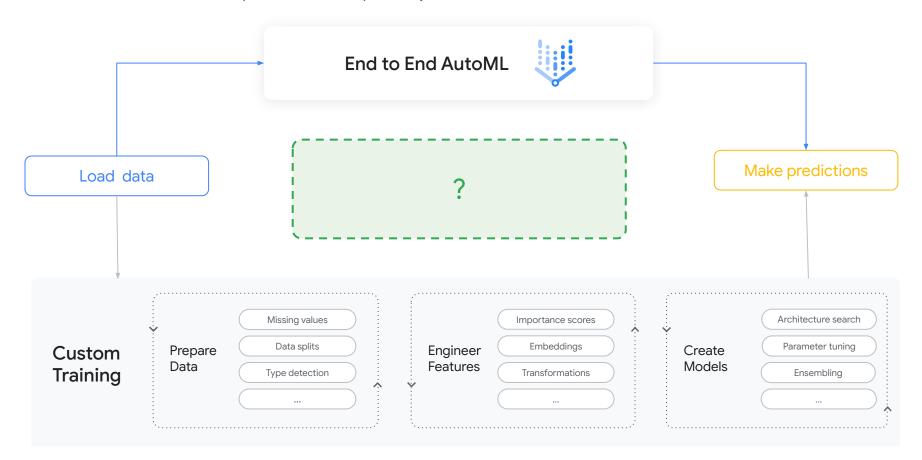
Vertex AI でのテーブルデータ



Vertex AI でのテーブルデータ



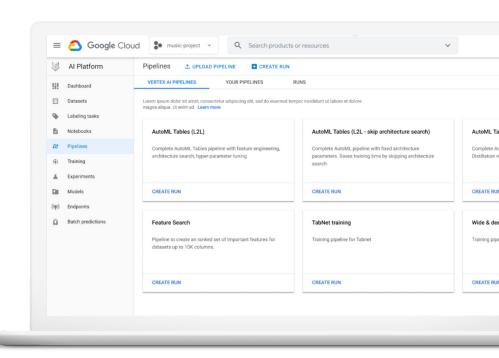
Vertex AI でのテーブルデータ



Tabular Workflows とは

テーブルデータを使ったモデル開発を支援するための、フルマネージドでスケーラブルなパイプライン、そして個別のコンポーネントが組み込まれます。

- AutoML
- Managed Algorithms
- Feature Search
- Feature Transformations



本番環境での使用を想定

組み込み

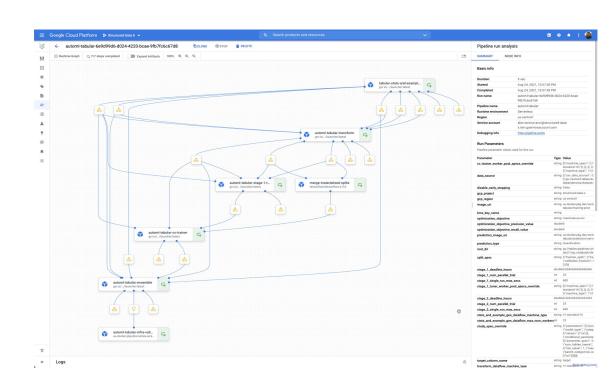
版管理やオーケストレーション、 MLOps 実装のため、Vertex AI のマ ネージドな Pipelines が使えます

透明性

各コンポーネントの入力パラメータ、 成果物、ログにアクセス可能です

柔軟性

必要に応じて、Pipelines のまま使用 したり、コピー/変更が可能です



同じコンポーネントが多くのワークフローへ対応

End To End AutoML Tables	Distilled model	TabNet model
Data Split		
Feature Transforms		
Architecture Search		Feature Search
		TabNet Trainer
Model ensemble		1 1 1 1 1 1 1
	Model Distillation	

AutoML ワークフローが ML を加速



自動化

AutoML は反復した手動の機械 学習タスクを自動化します。多くの ケース(ハイパーパラメータチュー ニングなど)でAutoML は開発時 間を節約しモデルの効率的に実 装できます

柔軟

Workflows はモジュールとしてカスタマイズ性が高く、特定の手順をスキップしたり修正することも可能で、ハードウェアのオプションも変更できます

透明性

Workflow pipelines は高い透明性をもち、各ステップ、トリガーサービスからのログ、そして成果物の詳細な情報をコンポーネントから手軽に見ることができます

組み込み

Workflowsは Feature Store、
Model Metadata、Experiment
Tracking などの MLOps ツール
も組み込まれています。管理やモニタリングのオーバーヘッドを抑え
ながら実験を増やせます。

利用可能なワークフロー

Classification and Regression

- 1. End to End AutoML Public Preview
- 2. End to End AutoML with Distillation Public Preview
- 3. TabNet Model Training Private Preview
- 4. Wide & Deep Model Training Private Preview
- 5. XGBoost Model Training Coming Soon

Forecasting

- 1. End to End AutoML Coming Soon
- Temporal Fusion Transformers
 Model Training Coming Soon

Feature Engineering

- 1. Feature Search Private Preview
- 2. Feature Transformations Private
 Preview



Thank you.