



【Showcase】 データベース技術者のための機械学習入門

捨田利 沙羅 / 出口 龍之介

Cloud Solution Engineer
Oracle Corporation Japan
May. 20th, 2022



oradev22



Agenda



14:15 【データの準備】 Autonomous Database を使ってみよう (70分)

Autonomous Database 概要説明

Showcase : Autonomous Database 作成 ・ DBの各種操作

15:30 【モデル構築・評価/デプロイ】 機械学習を体験してみよう – AutoML UI (45分)

AutoML UI 概要説明

Showcase : AutoML UIを使用した予測モデルの作成 ・ アプリの作成

* セッションやデモの進行状況によって多少終了時間が前後する場合がございます。

機械学習とは



機械学習（きかいがくしゅう、[英](#): Machine Learning）とは、経験からの学習により自動で改善するコンピュータアルゴリズムもしくはその研究領域で、人工知能の一種であるとみなされている。

「訓練データ」もしくは「学習データ」と呼ばれるデータを使って学習し、学習結果を使って何らかのタスクをこなす。

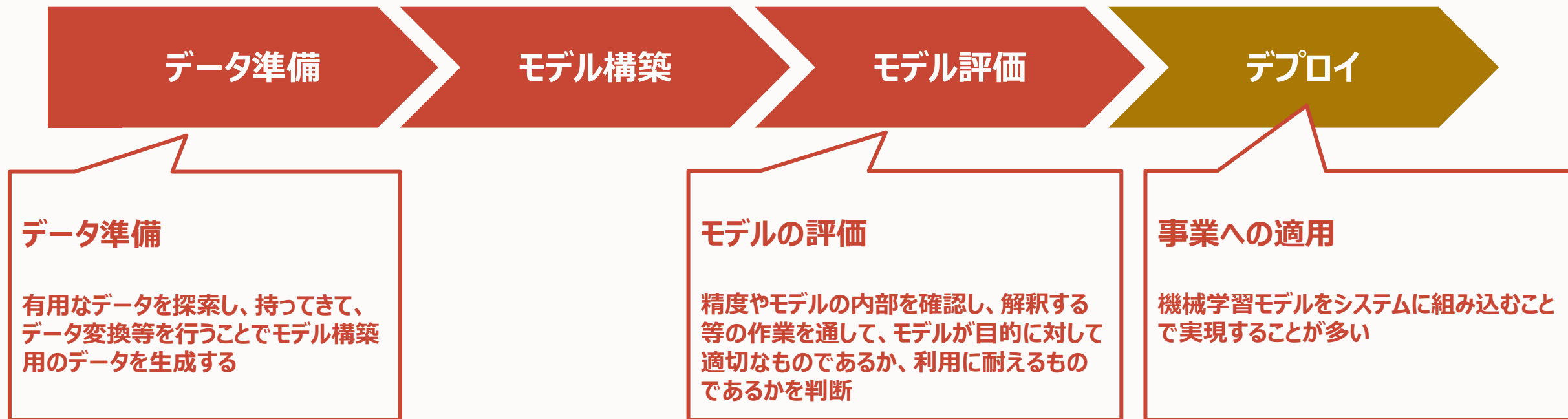
例えば過去のスパムメールを訓練データとして用いて学習し、スパムフィルタリングというタスクをこなす、といった事が可能となる。

<https://ja.wikipedia.org/wiki/%E6%A9%9F%E6%A2%B0%E5%AD%A6%E7%BF%92>

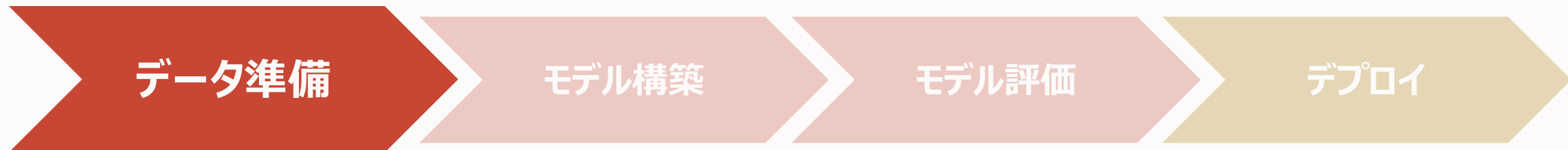
Wikipedia:機械学習 より引用

機能的には、
データから学習されたプログラムを生成、利用する

機械学習のプロセス



機械学習のプロセス



データ準備

有用なデータを探索し、持ってきて、
データ変換等を行うことでモデル構
築用のデータを生成する



分析、データ理解： データ連携・集計
データ変換処理： データ型を揃える 等
データ抽出： 必要なデータのみを抽出

機械学習のプロセス



データ準備

有用なデータを探索し、持ってきて、
データ変換等を行うことでモデル構
築用のデータを生成する



分析、データ理解： データ連携・集計
データ変換処理： データ型を揃える 等
データ抽出： 必要なデータのみを抽出

80%の時間がデータの準備/検索/管理に費やされている

→ *Autonomous Database*を使って作業の負担を軽減可能



【データの準備】

Autonomous Database を使ってみよう

- Autonomous Database 概要説明
- Showcase : Autonomous Database 作成 • DBの各種操作

Oracle Database デプロイメント選択肢

様々なシステム要件に対応

On Premise

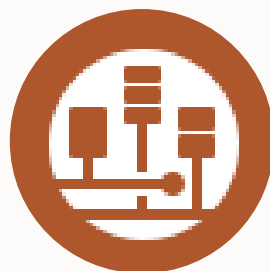


汎用サーバ上に
Databaseを
インストール



Exadataに
Databaseを構成

Oracle Cloud Infrastructure



IaaS上に
Databaseを
インストール



DBCS -
BM/VM



Exadata Cloud
Service
Or
Exadata Cloud
at Customer



**Autonomous
Database**
(Shared
or
Dedicated)

←
Most Manual

Most Autonomous →

シングル・データ・プラットフォームがデータドリブン実践の鍵

従来

基幹システム



SaaS



新規アプリ



データタイプ、機能別のデータサイロ

- 散在するデータ、複雑なデータ配置
- 複雑なデータ活用、価値提供の鈍化
- 多大なセキュリティリスク、運用保守コスト

理想

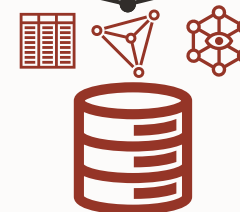
基幹システム



SaaS



新規アプリ



集約された、シングル・データ・プラットフォーム = コンバージドデータベース

- 集約され、一貫性のあるデータ
- 市場変化に追従する、鮮度のよいデータ活用
- 一元化されたセキュリティ、シンプルな運用保守

Oracleが提供するシングル・データ・プラットフォーム

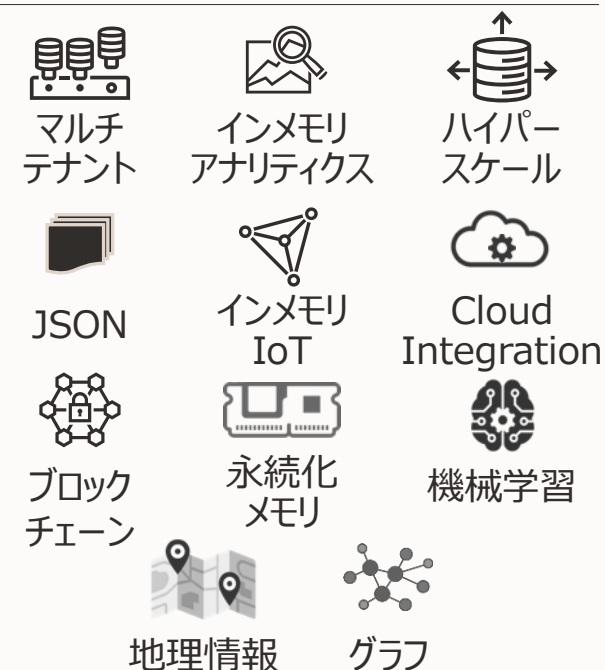
企業内の様々なデータ



シングル・データ・プラットフォーム = コンバージドデータベース



コンバージド機能



Oracle Autonomous Database

次世代データベース – 全てを自動化

エンタープライズ・クラス



Oracle Exadataを基盤
性能、スケーラビリティ、
可用性、セキュリティ

自動運転



低コストを実現する
真のクラウド伸縮性と
運用自動化

最高の生産性



迅速なデータ・ドリブン・
アプリケーション開発を実現する
統合されたLow-code開発環境

Autonomous Databaseの完全な柔軟性

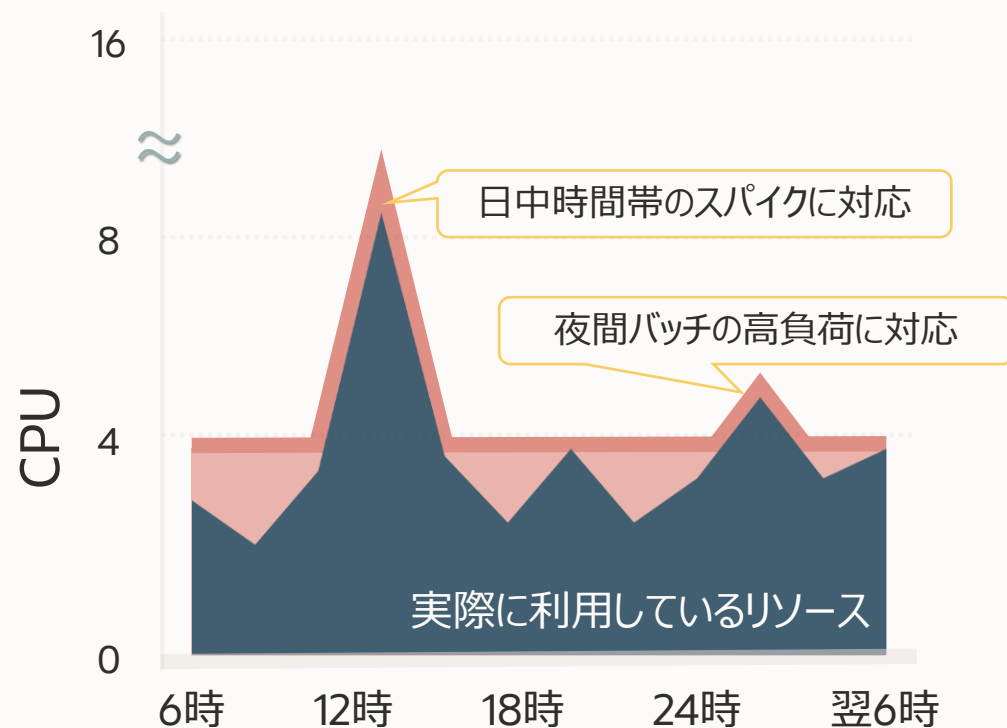
1 CPU単位 + 無停止 + 自動でスケーリング

負荷状況を判断し、**自動的に**リソースの拡張・縮小を**無停止で**実施します

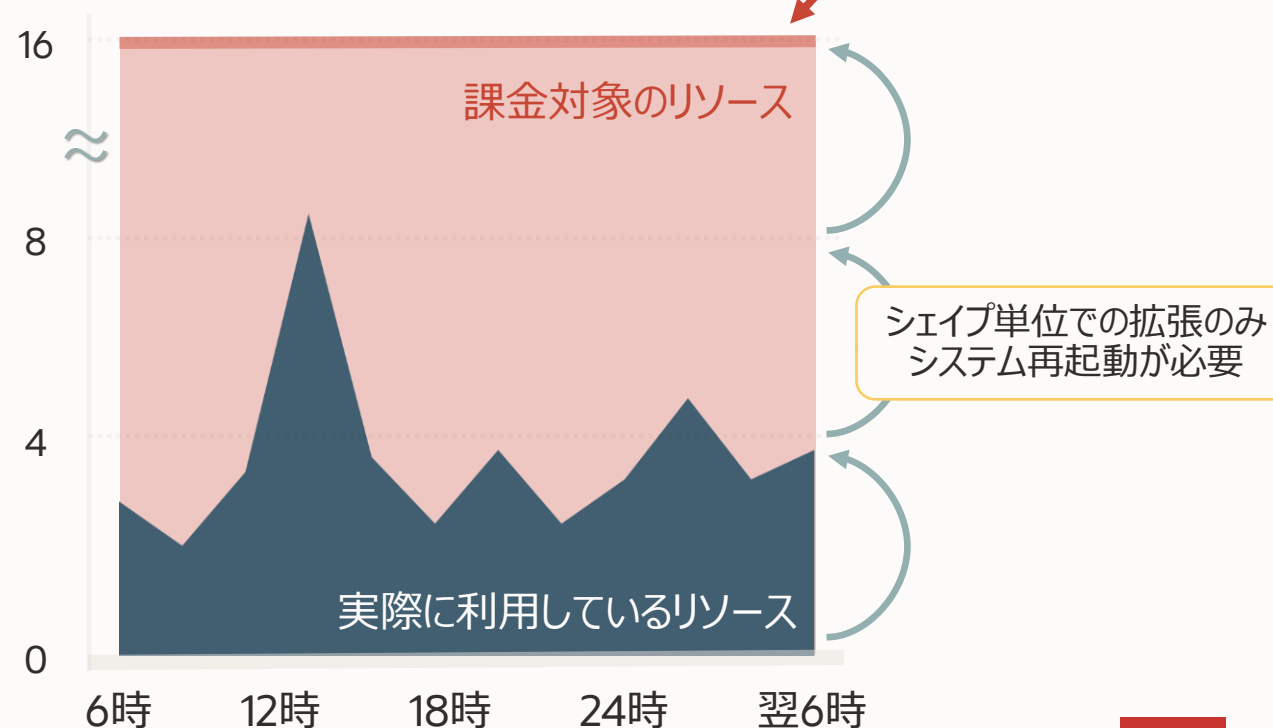
1 CPU単位で拡張・縮小を行い、**秒単位**で課金されます

柔軟性が欠如しているため、クラウドにも関わらず、ピークに合わせたサイジングが必要

Autonomous Database



他社データベース・サービス



Autonomous Databaseがもたらす運用自動化の効果

お客様は機械学習が実現する自動化の効果を実感

運用に人手をかけない

“Autonomous Data Warehouse Cloudにより Hertzは運用コストを削減し、その分を我々の顧客サービスやリワードの拡充に振り向けることが可能になります”



人的ミスがなく安心

“パフォーマンスの一貫性は多くのビジネスユーザーにとって重要です。Autonomous Data Warehouse Cloudによって人的ミスによるクエリー・パフォーマンスへの影響を排除することが可能になります。”



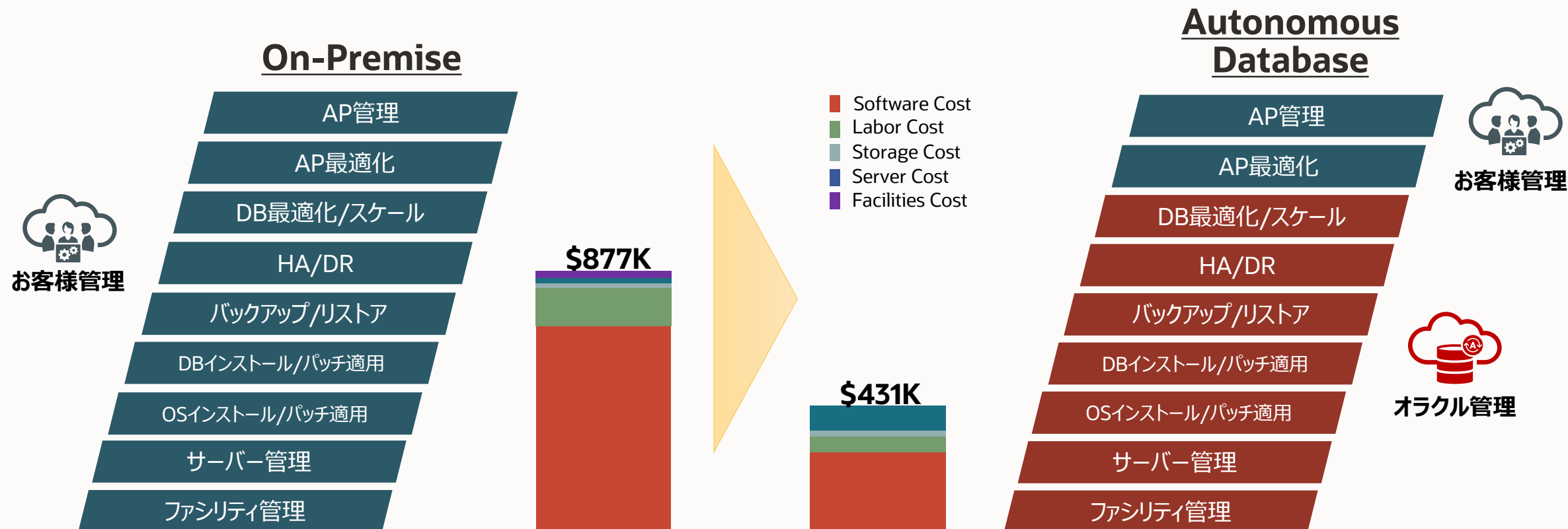
チューニング不要

“人間が生きるに当たって時間がすべてです。Autonomous Data Warehouse Cloudが自分自身でチューニングを行いダウンタイムがないことは我々にとっても我々の患者にとっても重要なことです。”



Autonomous Database

完全な自動化によるデータ管理の効率化/コストの削減



管理コストを最大80%、TCOを3年間で最大50%削減

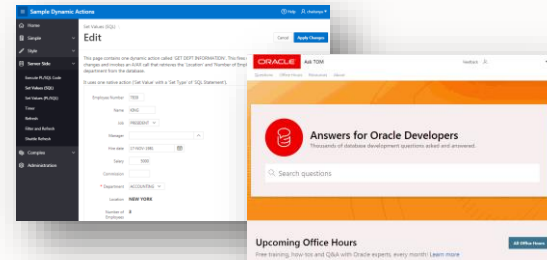


アプリケーション開発を強力にサポートするツール群 インスタンスを作成すれば、すぐに利用可能

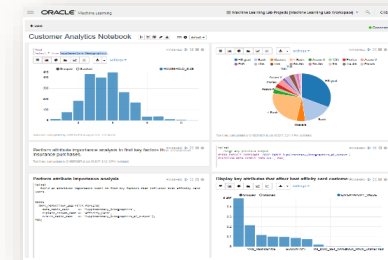
データ操作 Database Actions



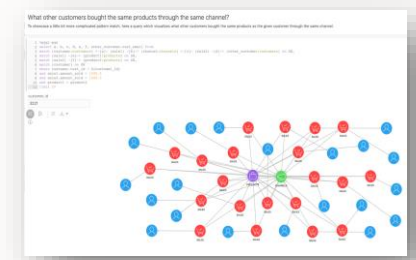
Webアプリ開発 Application Express : APEX



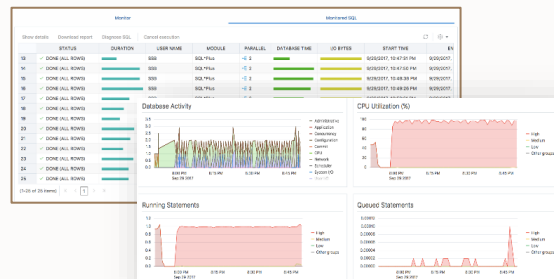
機械学習アナリティクス Notebook / AutoML UI



グラフDB開発 Graph Studio



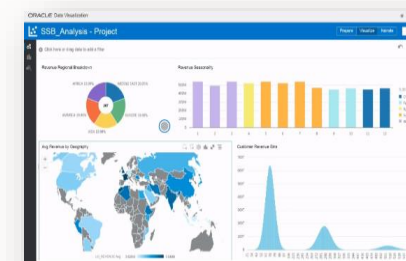
Service Console SQL Monitor / Performance Hub



API開発 Oracle Rest Data Services



ビジネスインテリジェンス Oracle Analytics Desktop



Autonomous Database 価格体系

フルマネージドのExadataを安く利用可能

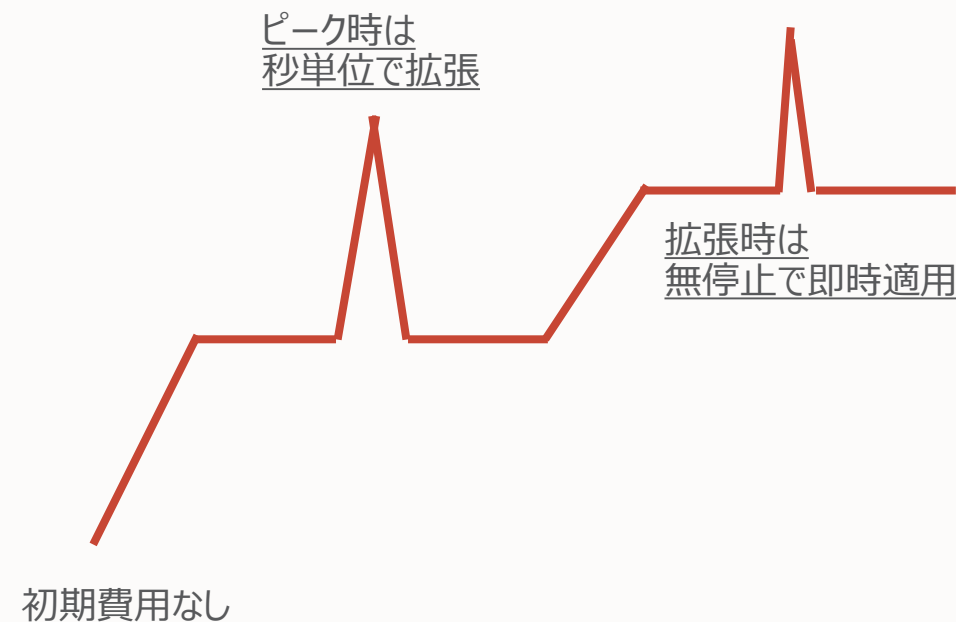
* 料金：2022/1 現在

CPU、ストレージのサイズを個々に設定

- OCPU：1 OCPU単位で秒課金
 - 1 OCPUを1時間利用した場合：約161円
- ストレージ：1 TB単位で秒課金
 - 1TB を 1ヵ月利用した場合：約14208円

従量課金制

- 最低1分からスタート可能
- 1分を超えた利用については秒単位で課金
 - 一時的なスケールアップに対するコストを抑制
 - テストのための一時利用を必要最低限のコストで対応
- お手持ちのOracle Database ライセンスの持ち込みも可能（BYOL）



Autonomous Database: 見積もり例



1 OCPU(2vCPU)/ストレージ 1TB/Backup 60日間分含む
24時間31日間(744時間)

	サービス名	数量	単価	単位	費用
1	Oracle Autonomous Data Warehouse Oracle Autonomous Transaction Processing	1	¥161,292	1時間/OCPU	¥120,001 (744時間)
2	Oracle Autonomous Database - Exadata Storage	1	¥14,208	1ヶ月/TB	¥14,208
	合計				¥134,209

* その他の費用 : Data Transfer (Outbound) 等
* As of June 2021

見積りサイト

https://cloud.oracle.com/ja_JP/cost-estimator



Globalで100社を超える顧客事例、日本国内で20社を超える顧客事例



Autonomous Databaseの革新を体験いただいたお客様の声

お客様	用途	比較対象	結果	コメント
Propre Japan	不動産情報ビッグデータ基盤	インメモリDB on AWS EC2	83秒→0.7秒 (99%短縮)	データを移行しただけでチューニングをしていない状態で高い性能が得られたことは衝撃的でした。稼働開始後の運用においても、チューニングが必要ないこと実感しています。運用管理が自動化されたことで新サービスの検討や開発に注力できるようになりました
オカムラ	全社データ分析基盤	オンプレミス 他社DWH	レスポンスタイム 最大60分の1 コスト36%削減	従来のデータベース環境と比較して少ないリソースで高いパフォーマンスを実現しています。具体的には、 バッチ処理で平均1.5倍に性能が改善したほか、BIのレスポンスタイムを最大60分の1に短縮しています。また年額費用は約36%削減しており、コスト最適化も果たしています。 レスポンスの高速化により従業員のレポート作成時の作業効率向上にもつながっています。
ファンコミュニケーションズ	アフィリエイトサービスのデータ管理基盤	AWS RDS for Oracle	コスト約5分の1	高性能とAIによる運用の自動化を実現したOracle Autonomous Database Cloudは、 まさに究極のデータベース と言っても過言ではありません。今、振り返ると、 その存在になぜもっと早く気付けなかったのかと悔しくなりませんね
ベネフィット・ワン	会員データ分析	Excel	2-3時間 -> 15分	従来 2～3 時間を費やしていた集計業務を わずか 15 分に短縮 会員の流入動向などの把握が月次または週次から 日次ベース へ
ピー・ビーシステムズ(ダリア)	販売管理データの分析基盤	AWS RDS for Oracle	95%短縮 コスト20%削減	パフォーマンスはもちろんのことコストも大きく向上 することを確認できましたので即ADWを採用しました。ADWのパフォーマンスと可能性には大きな期待をしております

Oracle Database と Oracle Machine Learning で新しい顧客を獲得

Forth Corporation, Thailandとその子会社であるForth Smart Corp.は、プリペイド携帯電話の充電、銀行送金、公共料金の支払い、モバイルゲームの支払いなどの金融サービスを提供する12万台以上のキオスクを運営しています。

Forth Smart社は、データベース内のOracle Machine Learningを使用して、**MLモデルの構築、評価、展開**にオラクルの既存の投資を活用しながら、事業範囲を拡大し、新しい機会を模索しています。

Oracle Machine Learningを使用することで、Forth Smartは90%もの精度でターゲットを絞ることができ、**コンバージョン数が3倍**になったことを確認しました。





【データの準備】

Autonomous Database を使ってみよう

- Autonomous Database 概要説明
- Showcase : Autonomous Database 作成 ・ DBの各種操作

Autonomous Database ワークロード/デプロイメント方式



ワークロード

**Autonomous
Data Warehouse
(ADW)**

(データマート/DWH)

**Autonomous
Transaction
Processing (ATP)**

(OLTP/混在ワークロード)

**Autonomous
JSON Database
(AJD)**

(JSONメイン)

**APEX Application
Development**

(APEXメイン)

デプロイメント

**Shared
Exadata Infrastructure**

(共有環境：Pluggable Database)

**Dedicated
Exadata Infrastructure**

(専有環境：OCI or C@C*)

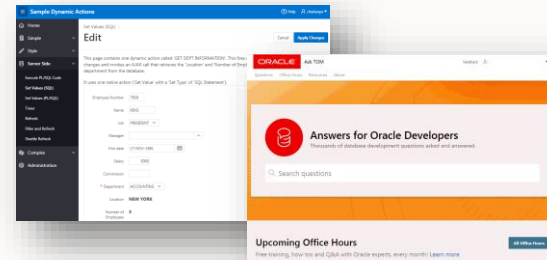


アプリケーション開発を強力にサポートするツール群 インスタンスを作成すれば、すぐに利用可能

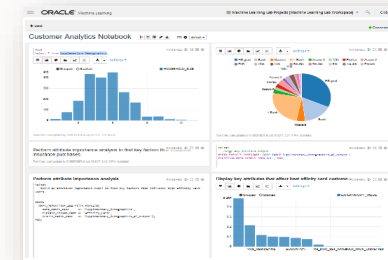
データ操作 Database Actions



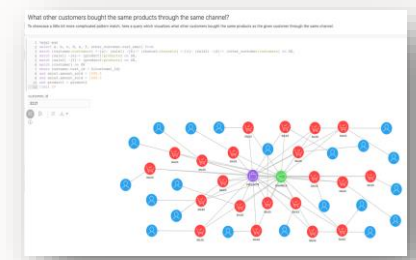
Webアプリ開発 Application Express : APEX



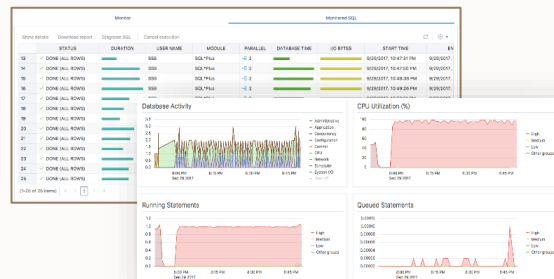
機械学習アナリティクス Notebook / AutoML UI



グラフDB開発 Graph Studio



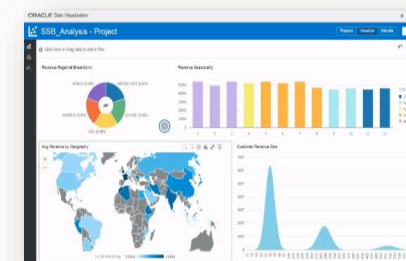
Service Console SQL Monitor / Performance Hub



API開発 Oracle Rest Data Services



ビジネスインテリジェンス Oracle Analytics Desktop



Database Actions

データベース・アクション

Oracle Databaseの各種操作に慣れていない方におすすめ

- より高度なクライアントツールを導入する前の、事前ステップとしても活用可能

SQL Developer Webの後継ツールとして、GUIで様々な操作が可能

- 開発：SQL実行、**RESTデータサービスの操作**、データモデル図の生成、**JSONデータの操作**
- データ・ツール：**データのロード（一括、自動）**、**データ・インサイト**、**カタログ**
- 管理：データベースユーザのパスワード管理、領域管理 など



参考資料

- 「Speakerdeck Database Actions」で検索ください



【モデル構築・評価/デプロイ】 機械学習を体験してみよう – AutoML UI

- AutoML UI 概要説明
- Showcase：AutoML UI を使用した予測モデルの作成・アプリの作成

Oracle Machine Learning – “AutoML UI” とは

ポイント：予測することで正しい対応を可能にする

<What>

過去のデータから未来を予測したいときに、機械学習を使って、予測モデルを作成し活用することができるOracle Databaseの機能。クリックだけ(GUI)で予測モデルを作成することができる製品。

<Differentiation>

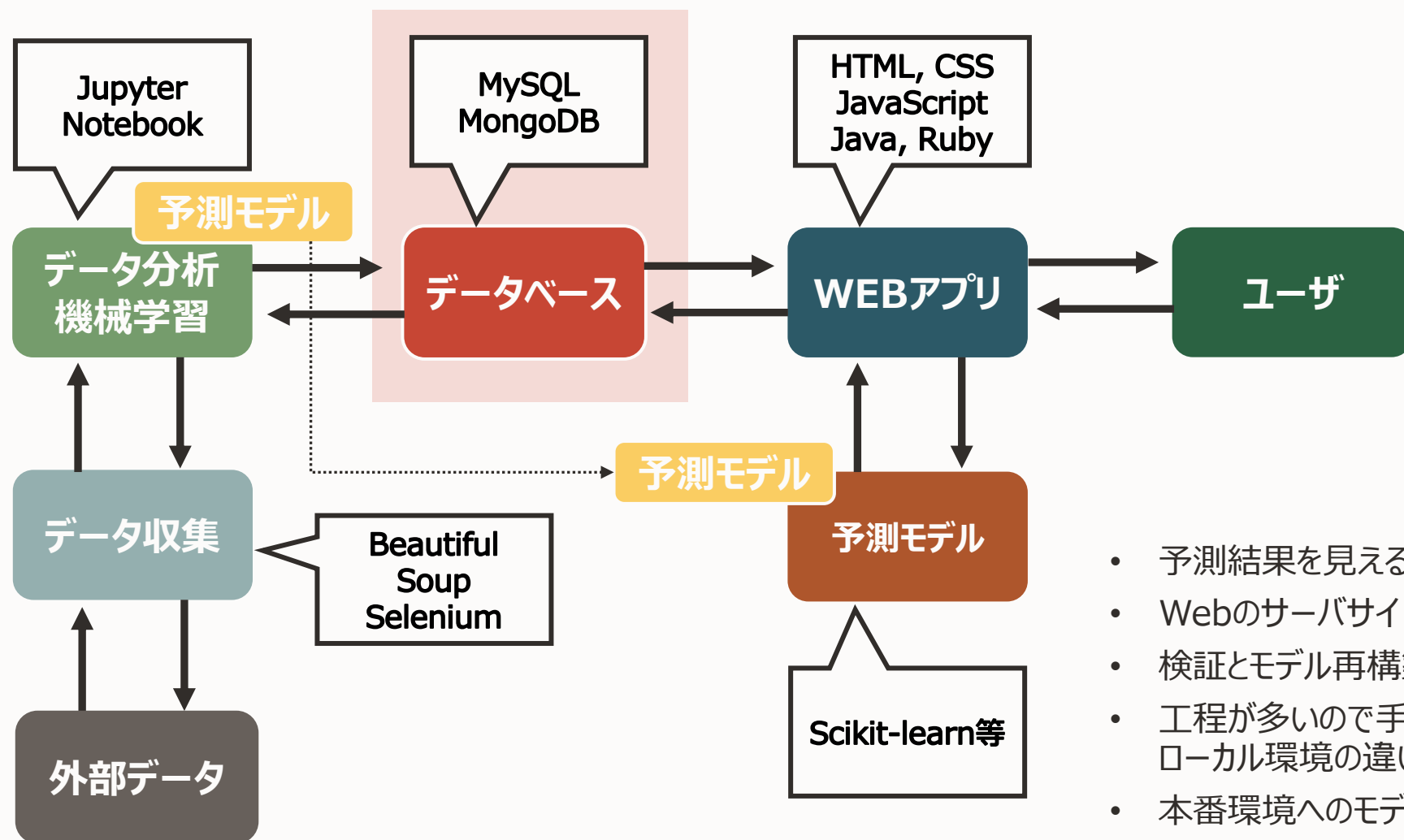
Autonomous Databaseを作成するだけで、Autonomous Databaseの中に、機械学習に必要な環境が一気に揃う。



<Case Study>

- ・ 在庫に関する悩みを抱えているお客様が、来月の在庫を予測
- ・ 購買行動を増やすために、顧客分析と機械学習で個人個人に効果的なクーポンを予測
- ・ コンバージョン率を上げるために、個人個人に効果的な広告を予測 etc.....

機械学習を組み込んだシステム構成図の1例

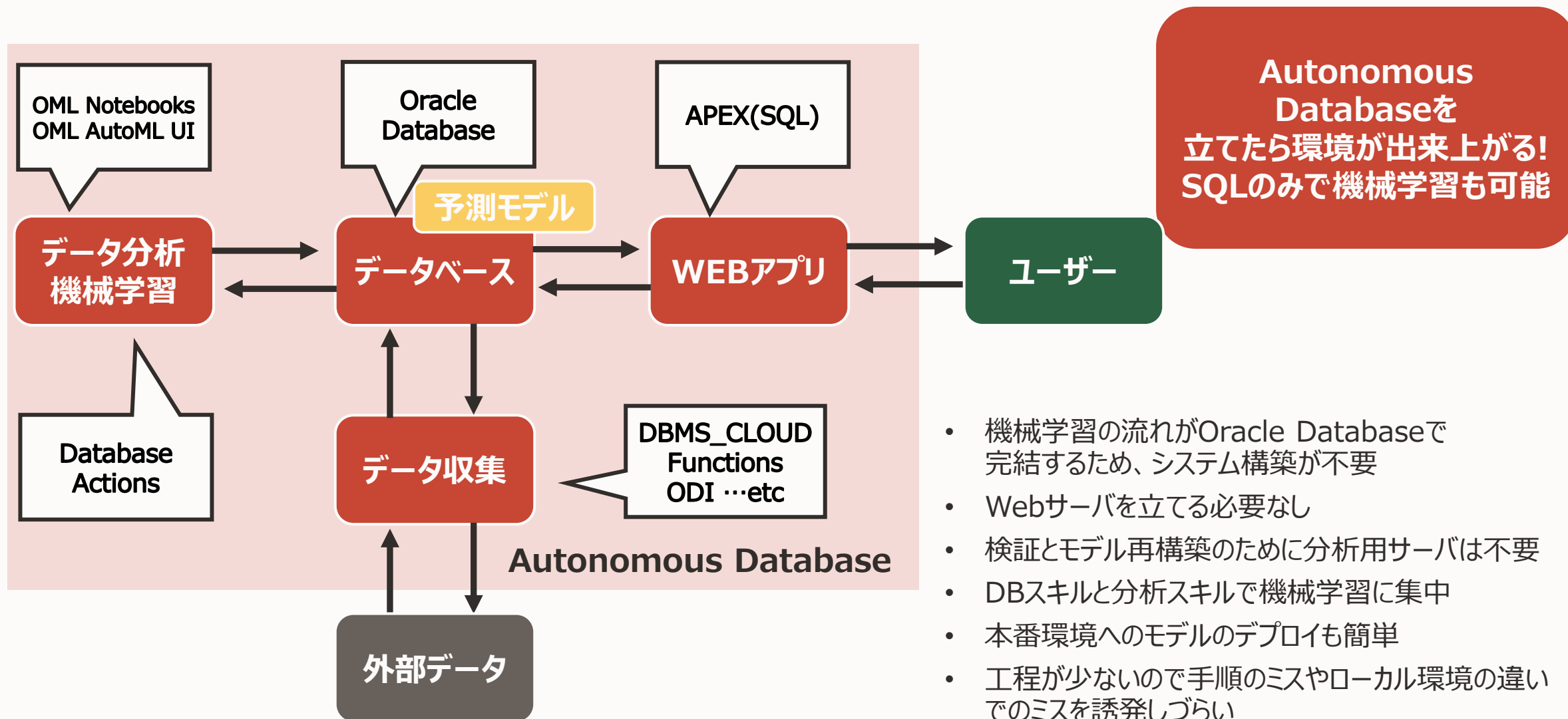


環境構築の工程が多く
手間がかかる
必要なスキルが多く
手順のミスが発生することも

- 予測結果を見える化させるためシステム構築が必要
- Webのサーバサイドのスキルが必要なるケースも
- 検証とモデル再構築のために分析用サーバが必要
- 工程が多いので手順のミスやローカル環境の違いで失敗することも
- 本番環境へのモデルのデプロイには様々な作業が必要



機械学習を組み込んだオラクルのシステム構成図の 1 例





【モデル構築・評価/デプロイ】 機械学習を体験してみよう – AutoML UI

- AutoML UI 概要説明
- Showcase：AutoML UI を使用した予測モデルの作成・アプリの作成

Oracle Machine Learning (OML) ショーケースの流れ

本日のデモでは、**kaggle**のチュートリアルとして有名な「**wine quality**」(**ワインの品質判定**) を用います。
アルコール(alcohol)や酸性度(fixed acidity)などのワインの中の成分から品質を評価するものです。

<今の状態>

ワインの成分に関する下記の2つのcsvファイルをDatabaseにロード済み

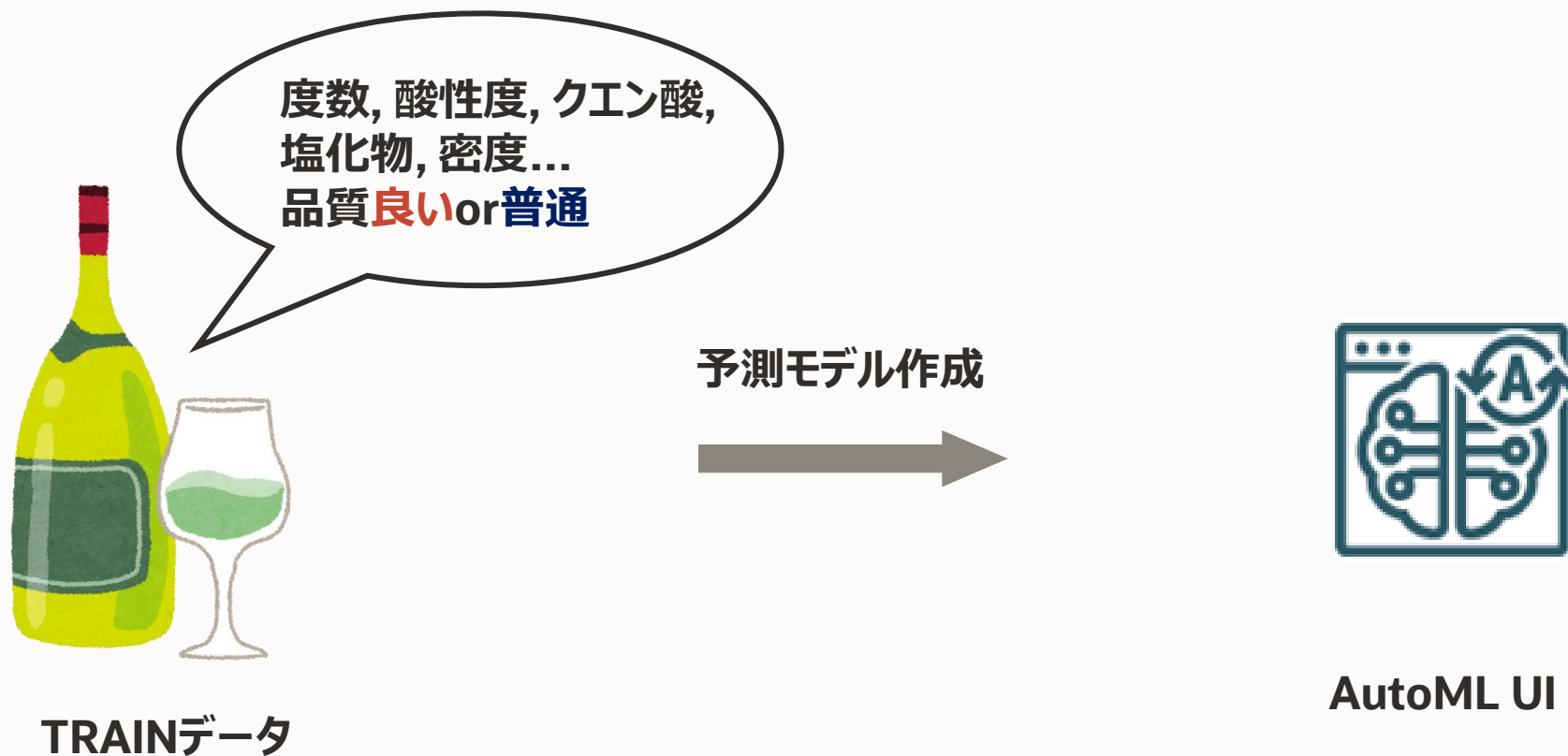
- **TRAIN表**(機械学習モデルを作成するための学習用データセット)
- **TEST表**(機械学習モデルが実際に予測を行う未知のデータセット)

<ハンズオンの流れ>

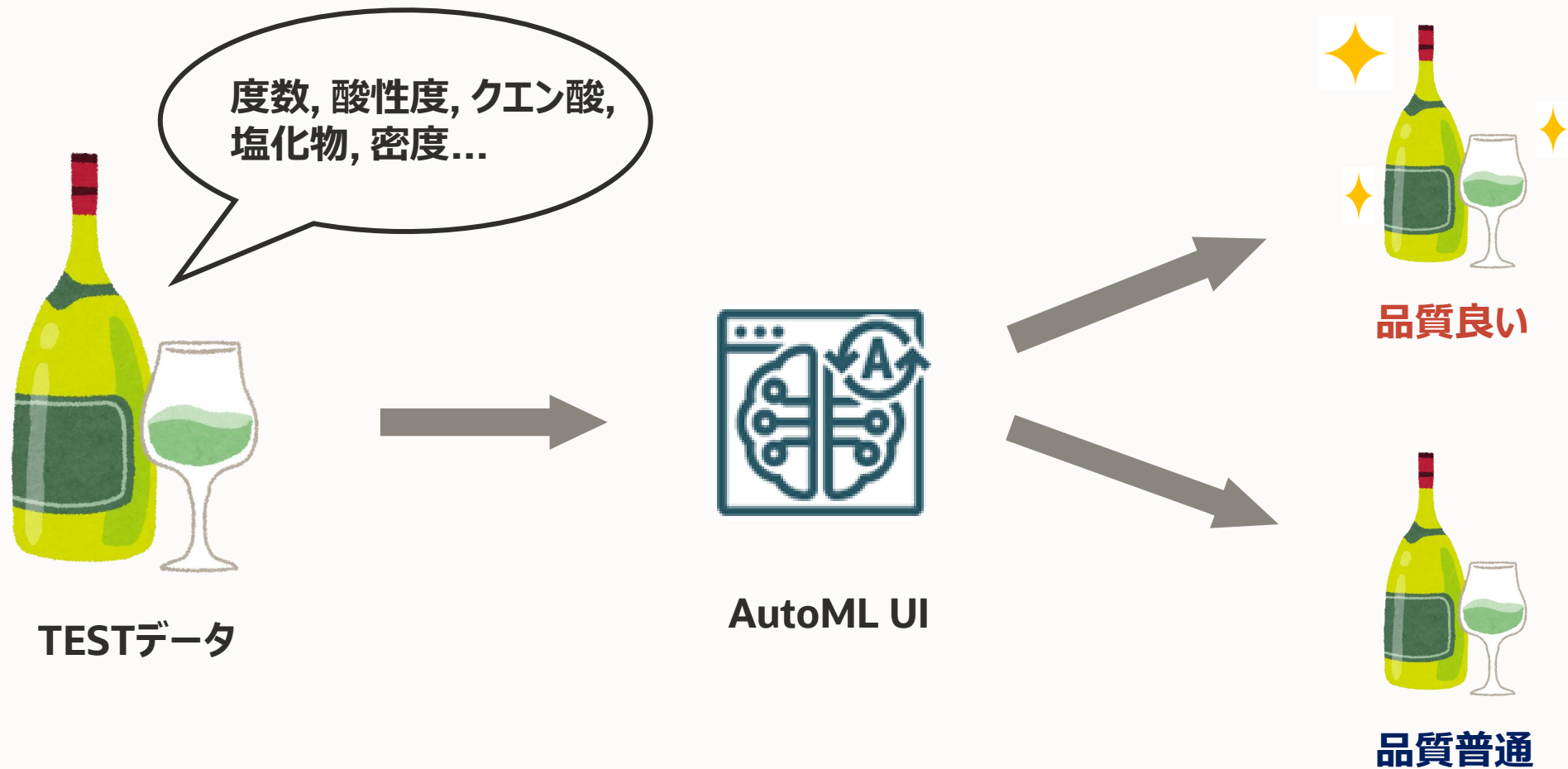
1. AutoML UIで自動で品質予測モデルを作成
2. 予測モデルで品質予測を行うViewを作成(OML Notebook活用)
3. APEXで品質予測のアプリケーションを構築



Oracle Machine Learning (OML) ショーケースの流れ



Oracle Machine Learning (OML) ショーケースの流れ



Oracle Machine Learning (OML) ショーケースの流れ

本日のデモでは、**kaggle**のチュートリアルとして有名な「**wine quality**」(**ワインの品質判定**) を用います。
アルコール(alcohol)や酸性度(fixed acidity)などのワインの中の成分から品質を評価するものです。

<今の状態>

ワインの成分に関する下記の2つのcsvファイルをDatabaseにロード済み

- **TRAIN表**(機械学習モデルを作成するための学習用データセット)
- **TEST表**(機械学習モデルが実際に予測を行う未知のデータセット)

<ハンズオンの流れ>

1. AutoML UIで自動で品質予測モデルを作成
2. 予測モデルで品質予測を行うViewを作成(OML Notebook活用)
3. APEXで品質予測のアプリケーションを構築



お客様の声

ポイント：手軽さ、データベース内機械学習、費用感

手軽さに関して

- 「AutoML UIで初学者の登竜門とされる0.85を超える精度が出るのは驚いた」
- 「本来売上があるはずだった値をきちんと予測してくれている」
- 「AutoML UIのUIは分かりやすく取っ付きやすくて良い」

データベース内機械学習に関して

- 「IN Databaseのコンセプトは、刺さるお客さんには物凄く刺さる製品」
- 「データベースで完結しているため環境構築やデプロイも楽にできるのは、助かる」

費用感に関して

- 「Autonomous Databaseを活用している/活用予定場合、OCPU数だけの課金で簡単に機械学習を始められる」
- 「機械学習の競合と比較したときに、Oracle Databaseを既に使っているためOML活用は安く済むから良い」



まとめ



Autonomous Database

新時代のデータベース・サービス



完全な
マネージド
サービス

実績のあるOracle Database / Exadata が基盤

高性能・高可用性・高セキュリティが実装済み

AI/機械学習を利用した完全自動運用
様々なツールが無償で同梱、DBで完結

完全な
柔軟性

1CPU単位でCPUを無停止で増減可能
ワークロードに応じた自動増減も可能
CPU/ストレージは1秒単位で課金

完全な
マルチモデル

あらゆるワークロード(OLTP/分析/混在)
あらゆるデータタイプ(構造化/JSON/グラフ等)
1つのデータベースで対応可能

本日実施いただいた内容



- サービス概要の理解
- データベース・インスタンスの作成、接続
- データベースへのデータ投入
- CSVファイルからWebアプリケーションを作成
- 機械学習ツールの利用体験



簡単すぐに様々なデータベース運用、およびアプリ開発ができる

参考資料

Oracle Cloud Infrastructure 技術情報

OCI 活用資料集



- OCIのサービス概要/詳細資料を公開
- 動画によるデモなど

OCI上の主要なサービスの概要・機能をまとめた資料集

OCI チュートリアル



- 実際のOCI上での操作をステップバイステップで解説
- 入門から応用レベルまで豊富

リソースの作成やその活用方法など実際の操作手順を体得

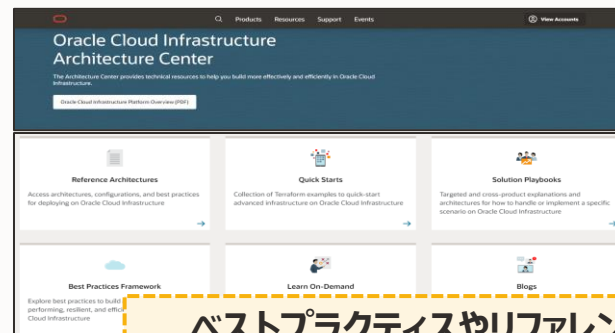
Oracle LiveLabs



- 100種類以上のワークショップ
- 具体的な操作性を完全無償で体感

テナントを用意せずとも実際の環境を触って使用感が確認可能

OCI Architecture Center



- リファレンスアーキテクチャ
- Solution PlayBooks
- Terraform サンプルスクリプト

ベストプラクティスやリファレンスアーキテクチャが参照可能





Thank you