

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

上総 虎智

機械学習プロトタイピングソリューションアーキテクト  
アマゾン ウェブ サービス ジャパン 株式会社

# 自己紹介

上総 虎智

機械学習ソリューションアーキテクト

好きなAWSサービス

- Amazon SageMaker
- AWS StepFunctions Data Science SDK





# 本セッションでお話すること

1. 機械学習システム開発プロジェクトの流れ
2. 機械学習プロジェクトを成功させるためのポイント
3. アマゾン ウェブ サービス (AWS) の機械学習系サービスについて

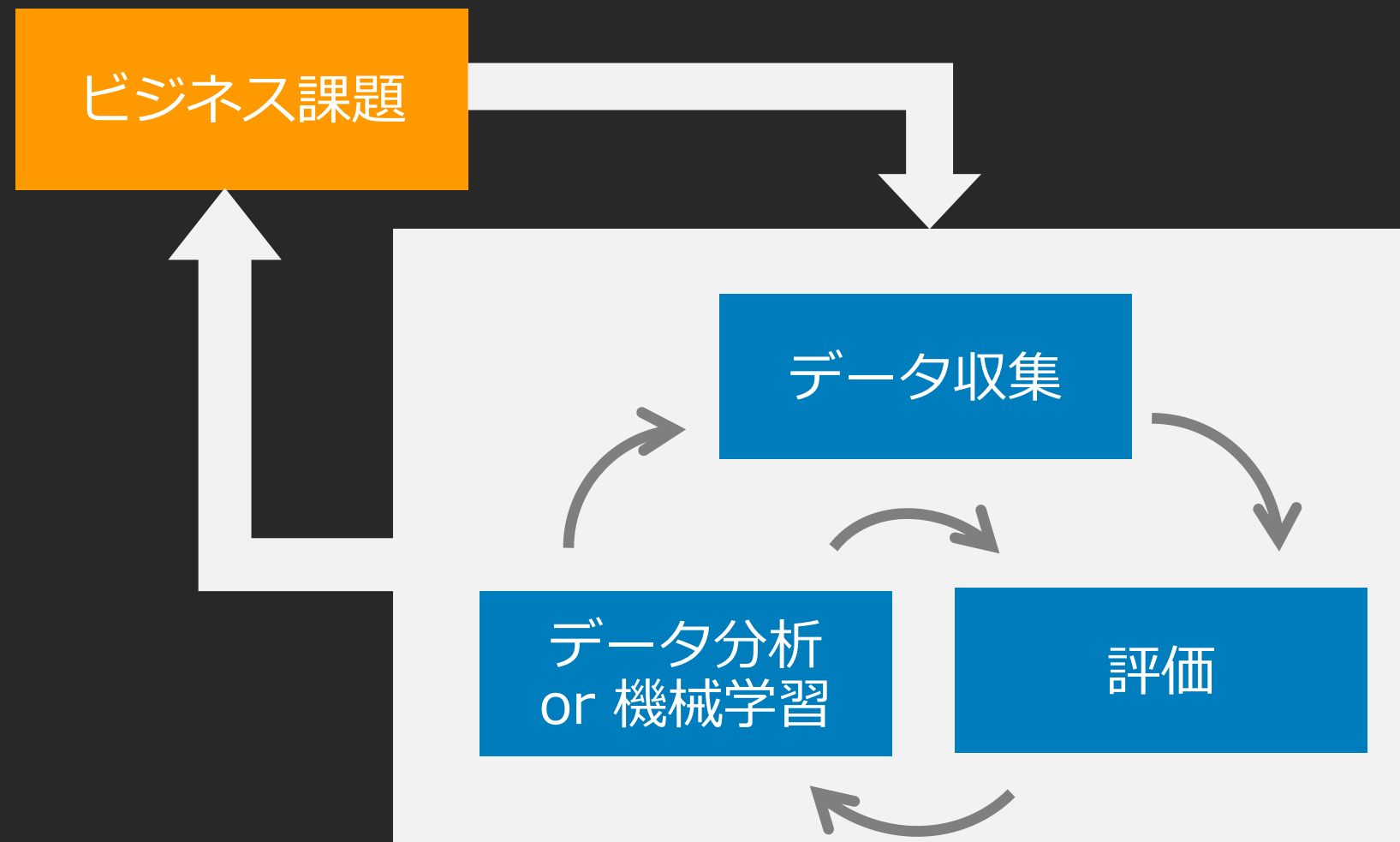
## お話ししないこと

- 機械学習そのものやアルゴリズムについて
- AWS 機械学習サービスの詳細について

# 機械学習システム開発プロジェクトの流れ

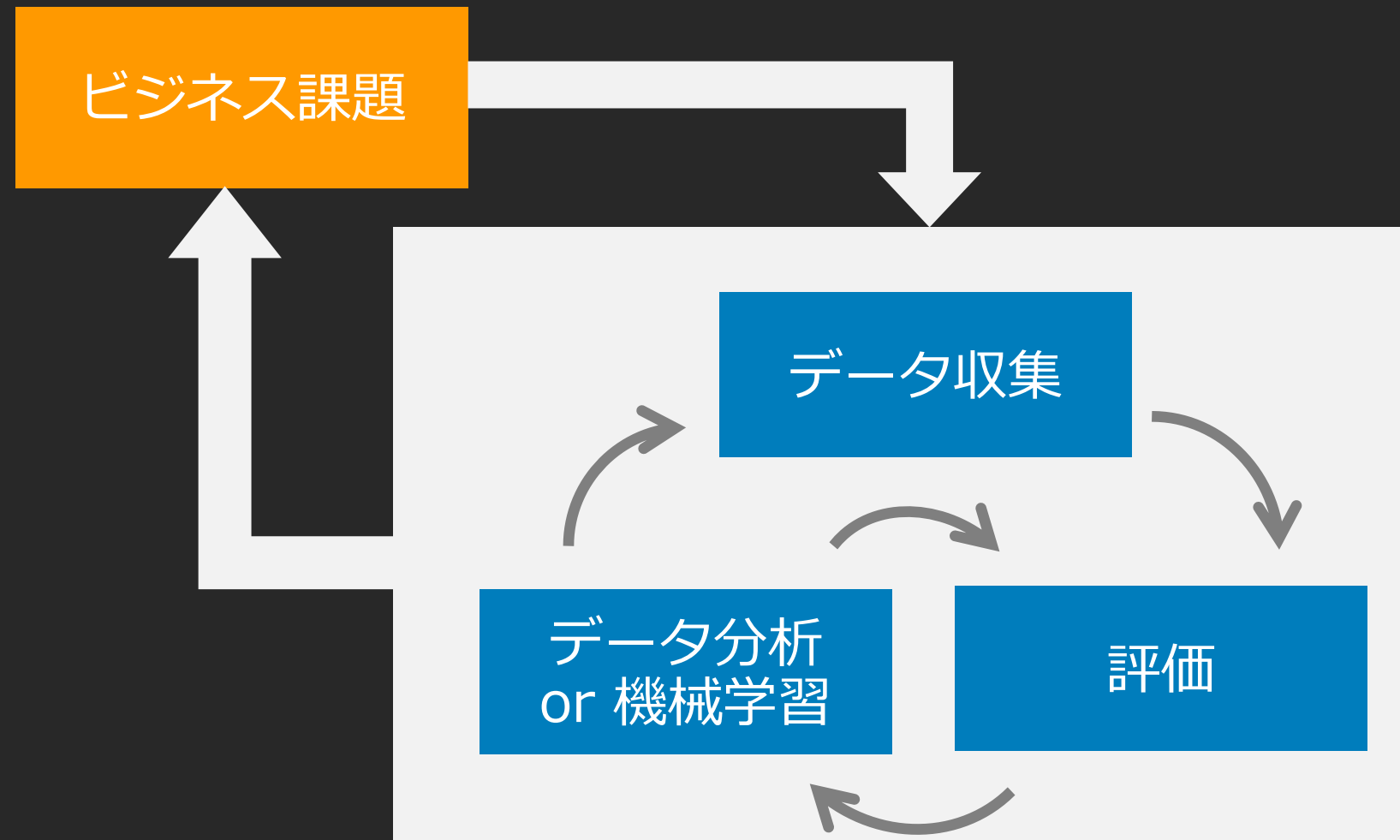
# ビジネス適用のための機械学習活用プロセス

- シーズからではなく、**ビジネス課題からスタート**する
- 試行錯誤の**ループ**を回して**評価**することを前提とする



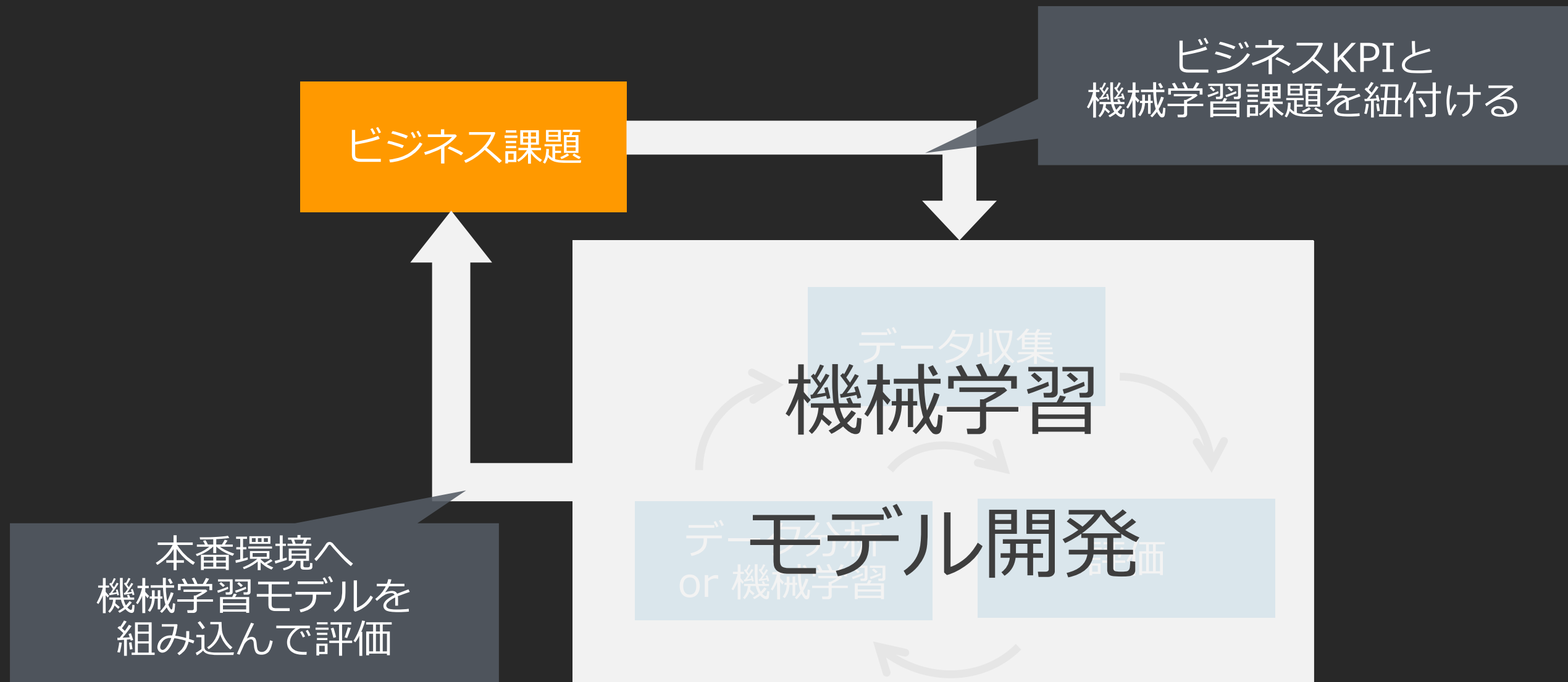
# 試行錯誤における2つのループ

- ビジネス適用からのフィードバックを得るループ
- 機械学習モデル開発のループ



# ビジネス適用からのフィードバックを得るループ

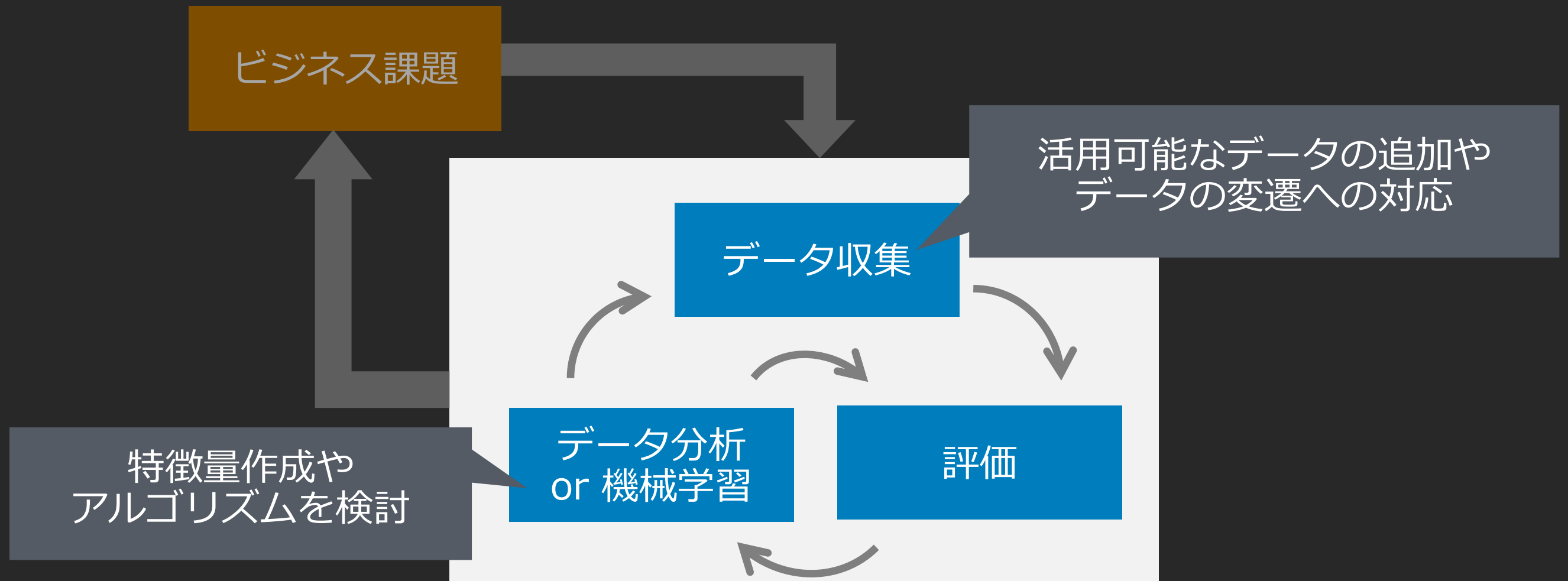
実施しているアプローチがビジネス価値を生み出すかを確認するためのための試行錯誤





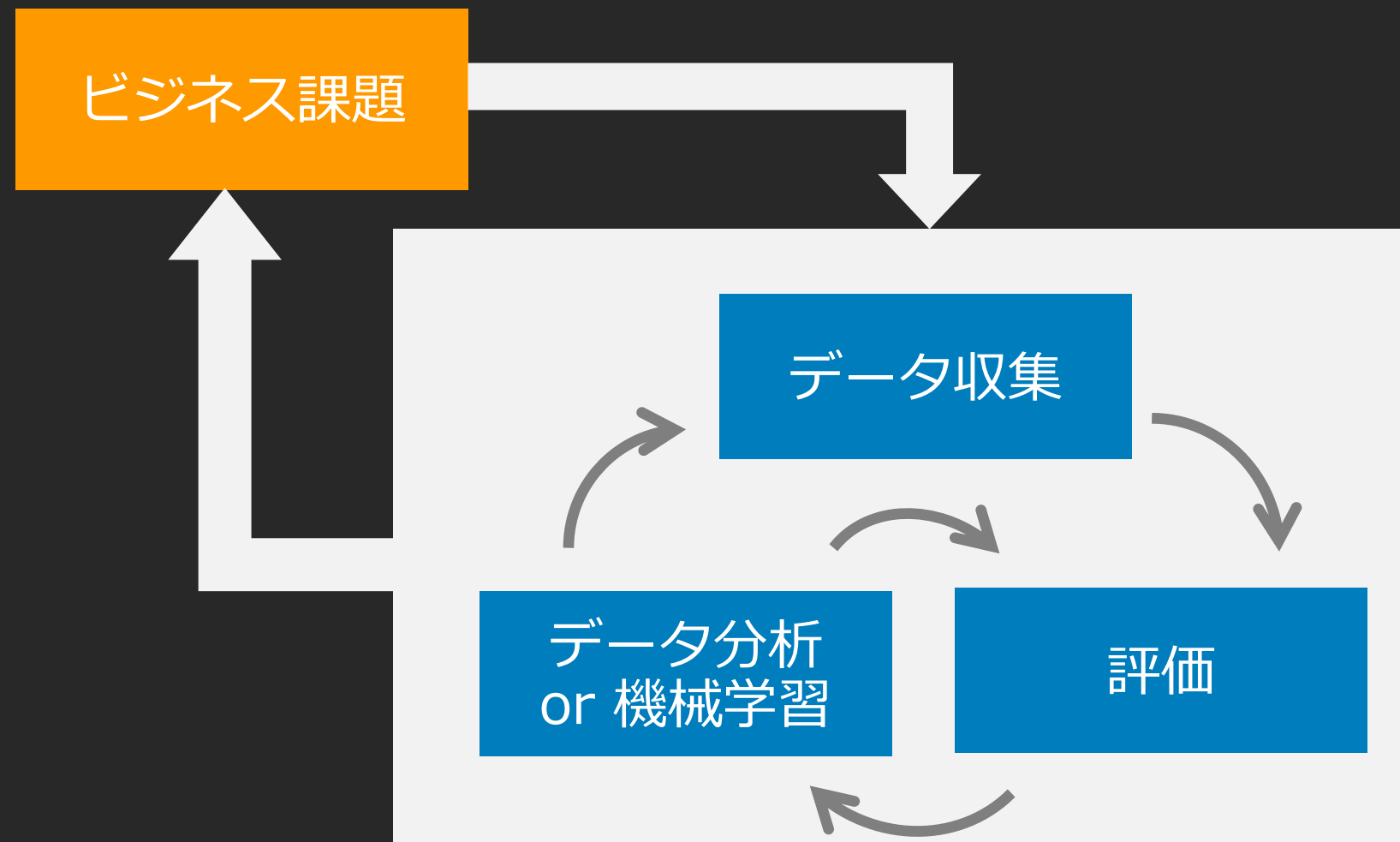
# 機械学習モデル開発のループ

取り組む機械学習課題の中でモデルの精度向上のための試行錯誤



# 試行錯誤における2つのループ

- ビジネス適用からのフィードバックを得るループ
- 機械学習モデル開発のループ



# 機械学習プロジェクトを成功させるポイント

2つのループをいかに高速に回せるか

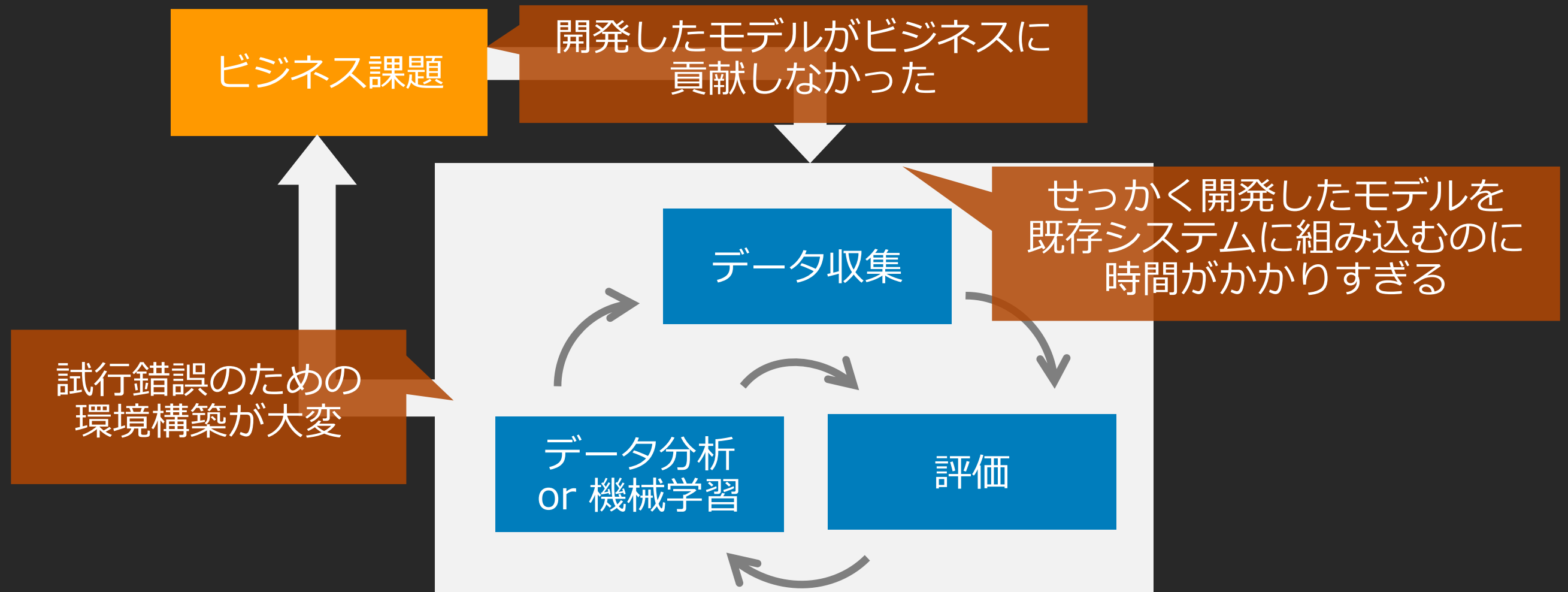
ビジネス価値につながらないものは

「なるべく作らない」を考える



# 試行錯誤を阻害する要因

試行錯誤の過程で、ビジネス上の価値を創出する以外の活動に時間を使ってしまうと、プロジェクトの成功が遠のいてしまいます。





# 機械学習開発での Undifferentiated Heavy Lifting

## 開発環境構築

- 必要なリソースの見積もりと購入の決断
- 開発チームで均一な開発環境構築
- フレームワークのインストール、バージョン管理

## 機械学習モデルの学習

- CPU/GPU など用途にあったハードウェア環境提供
- スケーラブルな分散学習構築と広帯域な通信環境

## 運用

- 推論環境の準備とモデルのホスティング
- 機械学習と異なるスキルセットが求められる



# 機械学習プロジェクトを進めるためのポイント

## 1. 最初から高度で完璧な解決を目指さない

- ✓ 開始当初は無理して独自の機械学習を開発せず**手間がかからないもの**から効果を確認
- ✓ **ビジネスへ貢献できると評価できれば徐々に高度化**

## 2. 試行錯誤の中で取り組む内容が変わることを前提に

- ✓ 変わる手段に応じて素早く試行錯誤ができるように **マネージドサービス** を活用
- ✓ 多様なサービスに対応できるようにデータを収集、保存する

# 機械学習プロジェクトを進めるためのポイント

## 1. 最初から高度で完璧な解決を目指さない

- ✓ 開始当初は無理して独自の機械学習を開発せず**手間がかからないもの**から効果を確認
- ✓ **ビジネスへ貢献できると評価できれば徐々に高度化**

## 2. 試行錯誤の中で取り組む内容が変わることを前提に

- ✓ 変わる手段に応じて素早く試行錯誤ができるように **マネージドサービス** を活用
- ✓ 多様なサービスに対応できるようにデータを収集、保存する

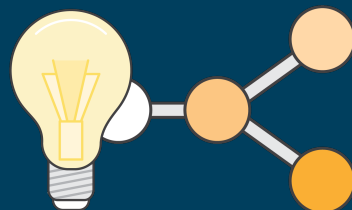
# 機械学習開発における「手間がかからない」とは？

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する

## 開発段階



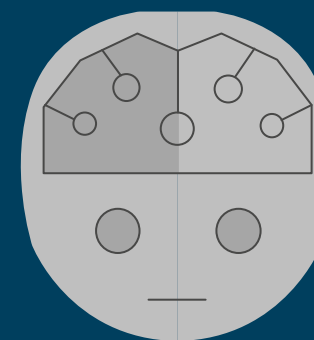
学習用データ



機械学習アルゴリズム

学習

## 活用段階



機械学習モデル



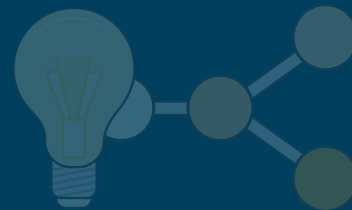
# 機械学習開発における「手間がかからない」とは？

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する

## 開発段階

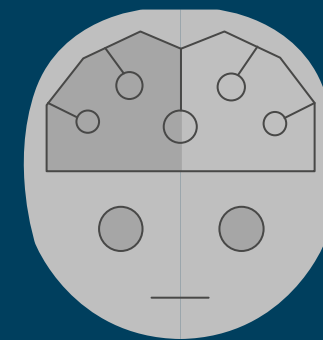


学習用データ



機械学習アルゴリズム

## 活用段階

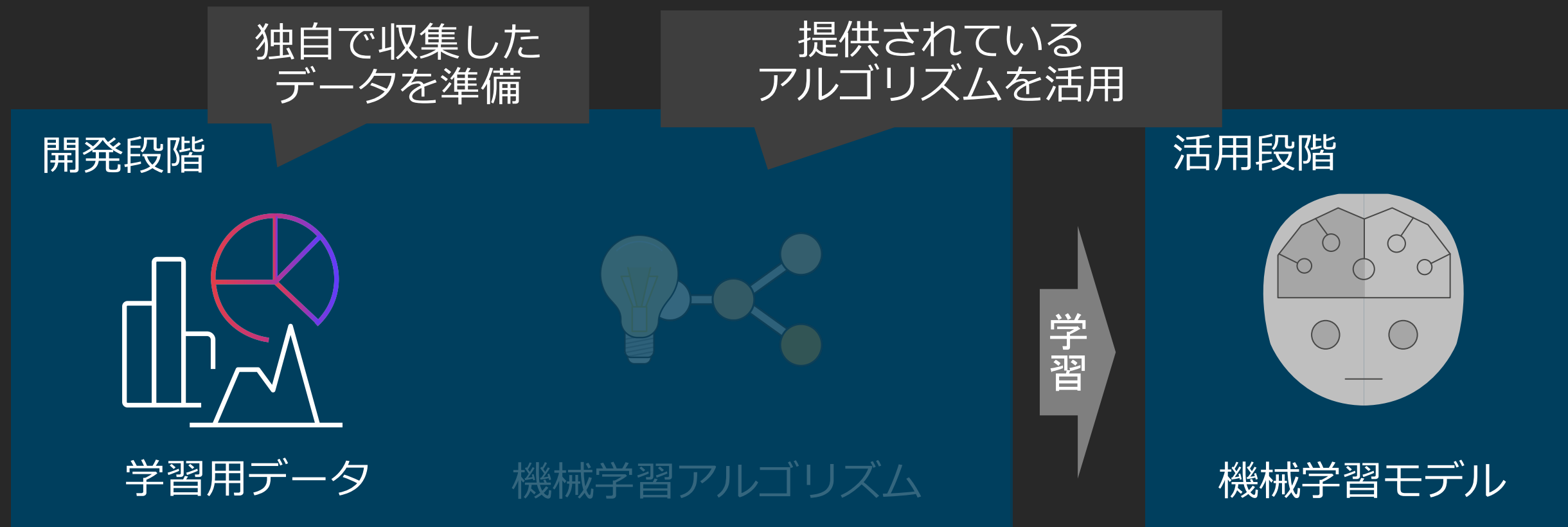


機械学習モデル

既に学習された  
モデルを活用

# 機械学習開発における「手間がかからない」とは？

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する



# 機械学習開発における「手間がかからない」とは？

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する



# 機械学習プロジェクトを進めるためのポイント

## 1. 最初から高度で完璧な解決を目指さない

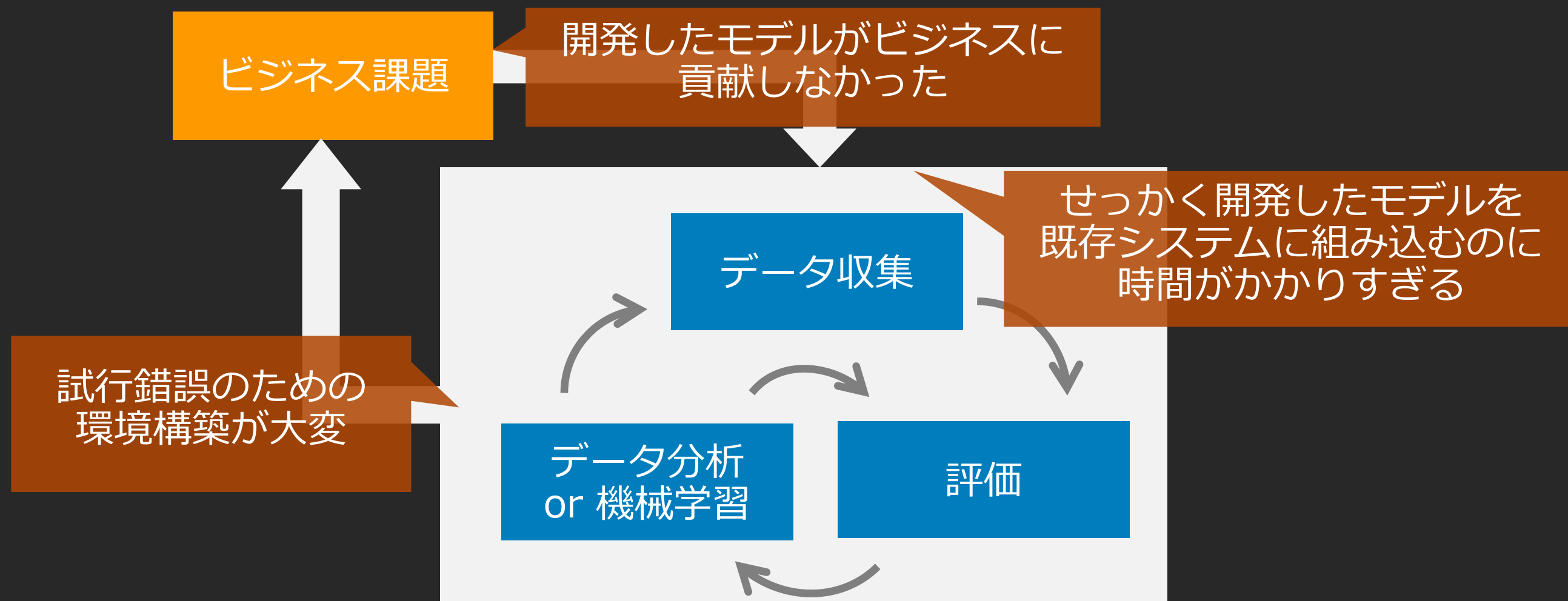
- ✓ 開始当初は無理して独自の機械学習を開発せず **手間がかからないもの**から効果を確認
- ✓ **ビジネスへ貢献できると評価できれば徐々に高度化**

## 2. 試行錯誤の中で取り組む内容が変わることを前提に

- ✓ 変わる手段に応じて素早く試行錯誤ができるように **マネージドサービス** を活用
- ✓ 多様なサービスに対応できるようにデータを収集、保存する

# 試行錯誤を阻害する要因

試行錯誤の過程で、ビジネス上の価値を創出する以外の活動に時間を使ってしまうと、プロジェクトの成功が遠のいてしまいます





# AWS の機械学習スタック: サービスの提供形態

事業に集中できるサービスモデルを提供することがクラウドの価値

AI / ML

Infrastructure

|             |             |             |             |
|-------------|-------------|-------------|-------------|
| アプリケーション作成  | アプリケーション作成  | アプリケーション作成  | アプリケーション作成  |
| デプロイ        | デプロイ        | デプロイ        | デプロイ        |
| 学習          | 学習          | 学習          | 学習          |
| モデル開発       | モデル開発       | モデル開発       | モデル開発       |
| ラベリング       | ラベリング       | ラベリング       | ラベリング       |
| FW/MW導入・パッチ | FW/MW導入・パッチ | FW/MW導入・パッチ | FW/MW導入・パッチ |
| OS導入・パッチ    | OS導入・パッチ    | OS導入・パッチ    | OS導入・パッチ    |
| HWメンテナンス    | HWメンテナンス    | HWメンテナンス    | HWメンテナンス    |
| ラッキング       | ラッキング       | ラッキング       | ラッキング       |
| 電源・ネットワーク   | 電源・ネットワーク   | 電源・ネットワーク   | 電源・ネットワーク   |

オンプレミス

ML frameworks & infrastructure

ML services

AI services

必要となるリソースや難易度が下がっていく

利用者が担当

AWS が担当














2つのループをいかに高速に回せるか

そのために「なるべく作らない」を考える


# AWS が提供する機械学習系サービス

# AWS における機械学習スタック

## AI サービス: 機械学習の深い知識なしに利用可能

|   |   |   |   |   |  |  |   |   |  |  |   |   |
|---|---|---|---|---|--|--|---|---|--|--|---|---|
| <br>Amazon Rekognition | <br>Amazon Polly | <br>Amazon Transcribe<br>+ Medical | <br>Amazon Comprehend<br>+ Medical | <br>Amazon Translate | <br>Amazon Textract | <br>Amazon Kendra | <br>Amazon Lex | <br>Amazon Personalize | <br>Amazon Forecast | <br>Amazon Fraud Detector | <br>Amazon CodeGuru<br>(Preview) | <br>Contact Lens<br>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














|   |              |                |                             |           |             |                         |          |                      |               |               |  |  |
|---|--------------|----------------|-----------------------------|-----------|-------------|-------------------------|----------|----------------------|---------------|---------------|--|--|
| <br>Amazon SageMaker | Ground Truth | ML Marketplace | Amazon SageMaker Studio IDE |           |             |                         |          | Amazon SageMaker Neo | Augmented AI  |               |  |  |
|   |              |                | Built-in algorithms         | Notebooks | Experiments | Model training & tuning | Debugger | Autopilot            | Model hosting | 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            | AWS deep learning AMIs & containers | GPUs & CPUs | Amazon Elastic Inference | AWS Inferentia | FPGA |
|  PYTORCH    |  Salesforce learn |  HOROVOD |  DeepGraphLibrary |                                     |             |                          |                |      |

# AWS における機械学習スタック

## AI サービス: 機械学習の深い知識なしに利用可能

|  |  |   |   |  |   |  |  |   |   |
|--|--|---|---|--|---|--|--|---|---|
| <b>Vision</b><br><br>Amazon Rekognition | <b>Speech</b><br> <br>Amazon Polly<br>Amazon Transcribe<br>+ Medical | <b>Text</b><br>  <br>Amazon Comprehend<br>Amazon Translate<br>Amazon Textract<br>+ Medical | <b>Search</b><br><br>Amazon Kendra | <b>Chatbots</b><br><br>Amazon Lex | <b>Personalization</b><br><br>Amazon Personalize | <b>Forecasting</b><br><br>Amazon Forecast | <b>Fraud</b><br><br>Amazon Fraud Detector | <b>Development</b><br><br>Amazon CodeGuru<br>(Preview) | <b>Contact centers</b><br><br>Contact Lens<br>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

AWS によってチューニングされた機械学習モデルを  
API で簡単に利用できるマネージドサービス群

## 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

AWS deep learning  
AMIs & containers

GPUs &  
CPUs

Amazon Elastic  
Inference

AWS Inferentia

FPGA

PYTORCH

fast  
learn

horovod

DeepGraphLibrary

# 目的に応じてご利用頂ける AI サービス

13種類の AI サービス を提供し、広い分野をカバー

静止画・動画認識



音声処理



テキスト処理



チャットボット



時系列データ予測



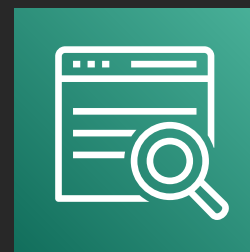
レコメンデーション



不正検知



検索



自動コードレビュー



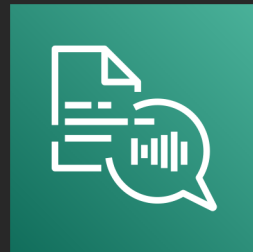
# 目的に応じてご利用頂ける AI サービス

13種類の AI サービス を提供し、広い分野をカバー

静止画・動画認識



音声処理



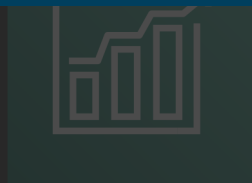
テキスト処理



チャットボット



1. 開発者はモデルの学習を行わず学習済のモデルを活用する



# 目的に応じてご利用頂ける AI サービス

13種類の AI サービス を提供し、広い分野をカバー

静止画・動画認識

音声処理

テキスト処理

チャットボット

2. 必要なデータを準備するだけでモデルが学習されて利用できる

時系列データ予測



レコメンデーション



不正検知



検索



自動コードレビュー





# AWS における機械学習スタック

AI サービス: 機械学習の深い知識なしに利用可能

Vision Speech Text Search Chatbots Personalization Forecasting Fraud Development Contact centers

機械学習の開発・学習・デプロイといった開発プロセスをサポートするマネージドサービス

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

ML Marketplace

Amazon SageMaker Studio IDE

Built-in algorithms

Notebooks

Experiments

Model training & tuning

Debugger

Autopilot

Model hosting

Model Monitor

Amazon SageMaker Neo

Augmented AI

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

AWS deep learning AMIs & containers

GPUs & CPUs

Amazon Elastic Inference

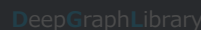
AWS Inferentia

FPGA

 PYTORCH

 fastai

 HOROVOD

 DeepGraphLibrary

# 画像認識ソリューションを作りたい時

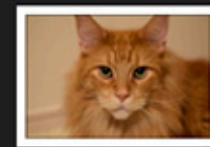
# 猫識別ソリューションを開発する場合

## 開発の背景

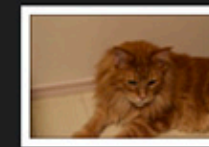
- 飼っている猫の飼育日記を残したい
- スマートフォンの中に大量の写真がある
- 猫が写っている写真だけをうまくピックアップしたい

## 課題解決の方法

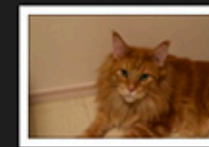
- 機械学習の活用による画像識別



DSC\_0001.jpg



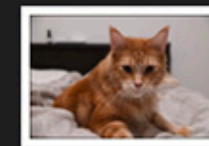
DSC\_0004.jpg



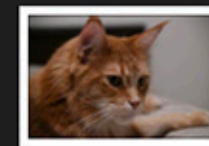
DSC\_0005.jpg



DSC\_0029.jpg



DSC\_0031.jpg



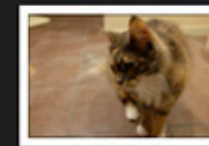
DSC\_0032.jpg



DSC\_0036.jpg



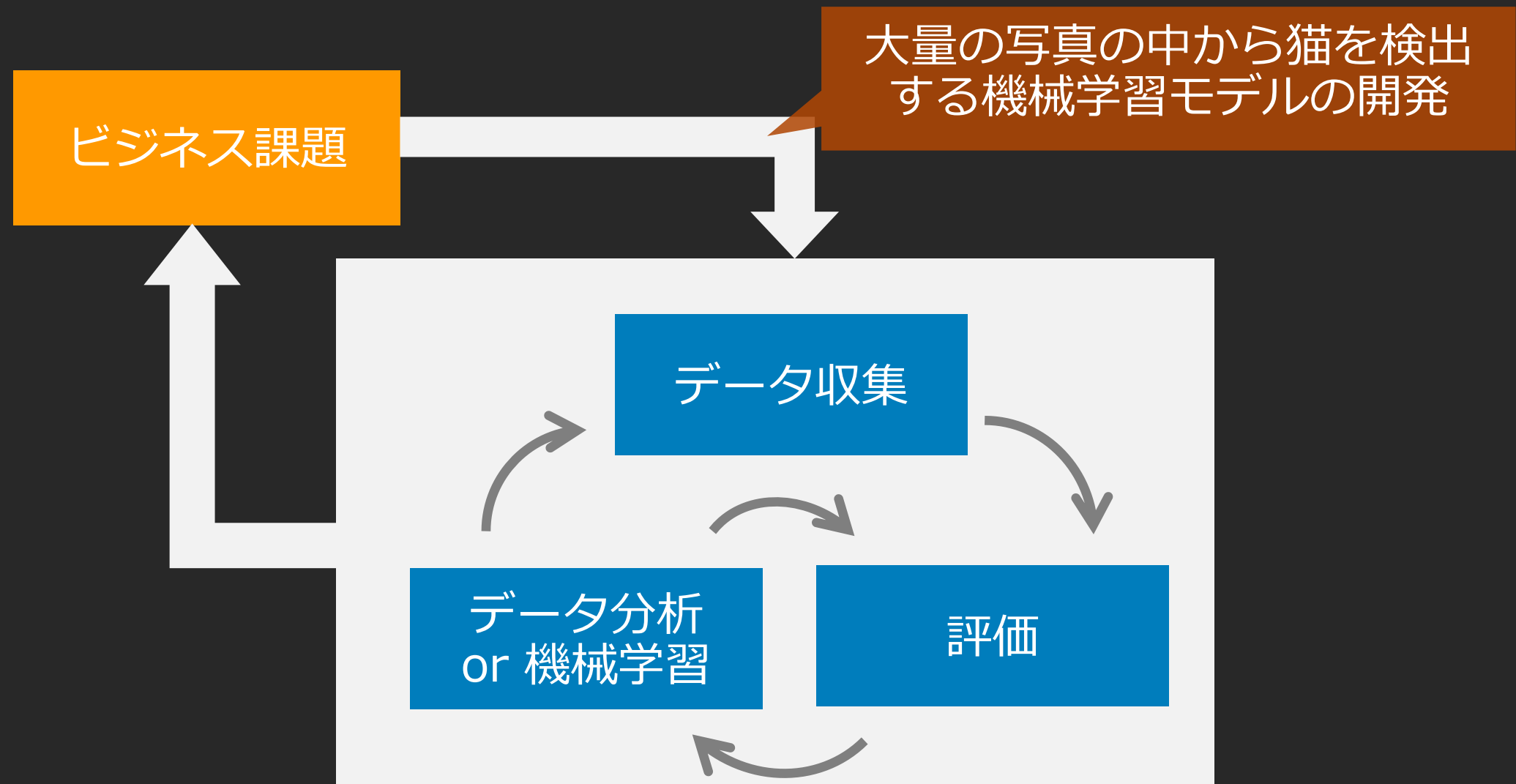
DSC\_0284.jpg



DSCF0019.jpg

# 猫識別ソリューションを開発する手順

実現したい内容に従って機械学習モデルを開発し、実際に適用してみて課題解決できているかを評価します



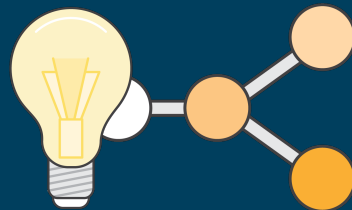
# 猫識別のための機械学習モデルを開発する手順

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する

## 開発段階



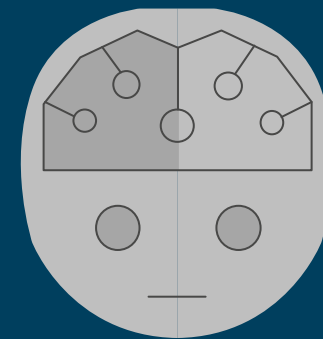
学習用データ



機械学習アルゴリズム

学習

## 活用段階



機械学習モデル

# 猫識別のための機械学習モデルを開発する手順

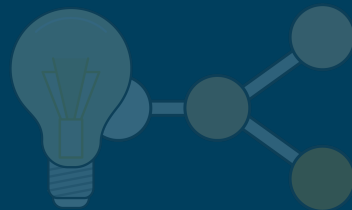
1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する

Amazon Rekognition の活用

開発段階



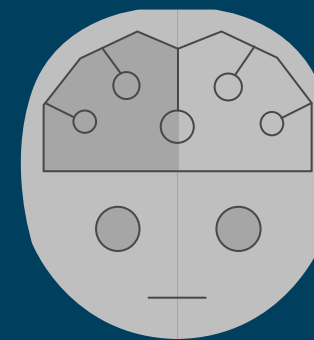
学習用データ



機械学習アルゴリズム

学習

活用段階

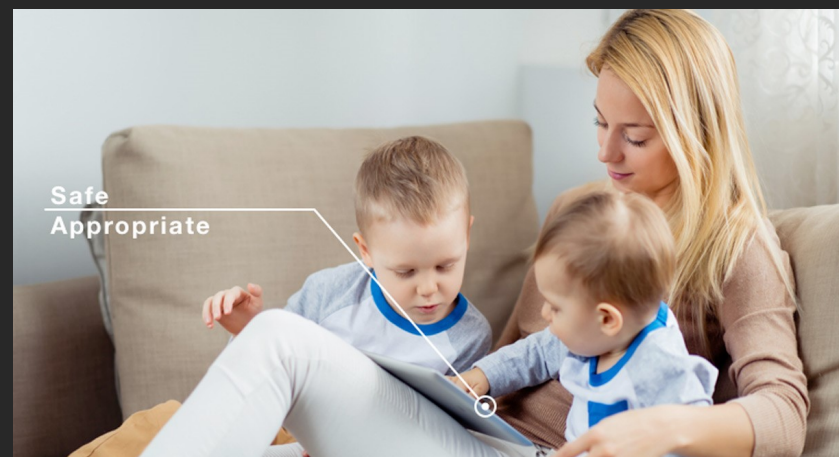


機械学習モデル

# Amazon Rekognition とは



対象物体、シーン、アクティビティ検出



安全でないコンテンツの検出



顔認識

- 深層学習に基づく画像・動画分析サービス
- Rekognition API に画像・動画を送るだけで簡単に対象物・人やテキストの検出、シーン・アクティビティ認識や不適切コンテンツの検出ができる
- 顔認識の API を利用すると、大規模な画像のコレクションの中から似た顔を見つけることができる



# Amazon Rekognition で猫の検出が可能

## オブジェクトとシーンの検出

Rekognition は、イメージ内の物体、概念、およびシーンに自動的にラベルを付け、信頼スコアを提供します。

Need to detect specific objects and scenes unique to your business?  
Use **Rekognition Custom Labels** to quickly build a custom model, no machine learning experience required.



サンプルイメージを選択

自前の画像を使用

[Read feature documentation to learn more](#)

Issues or questions? Use feedback button on bottom-left.

### ▼ 結果

|        |        |
|--------|--------|
| Mammal | 98.7 % |
| Cat    | 98.7 % |
| Animal | 98.7 % |
| Pet    | 98.7 % |
| Manx   | 98 %   |
| Wood   | 95.4 % |

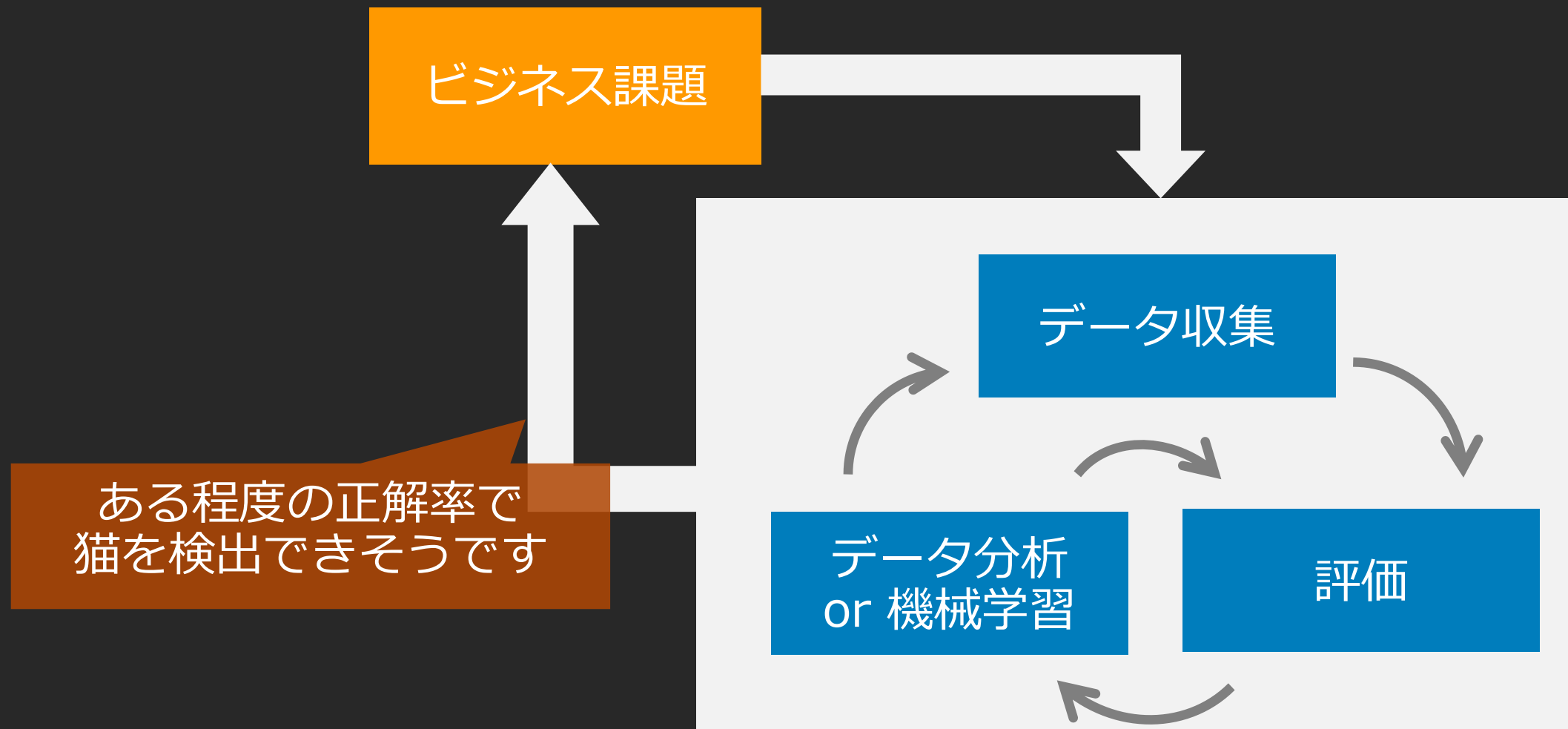
[さらに表示](#)

▶ リクエスト



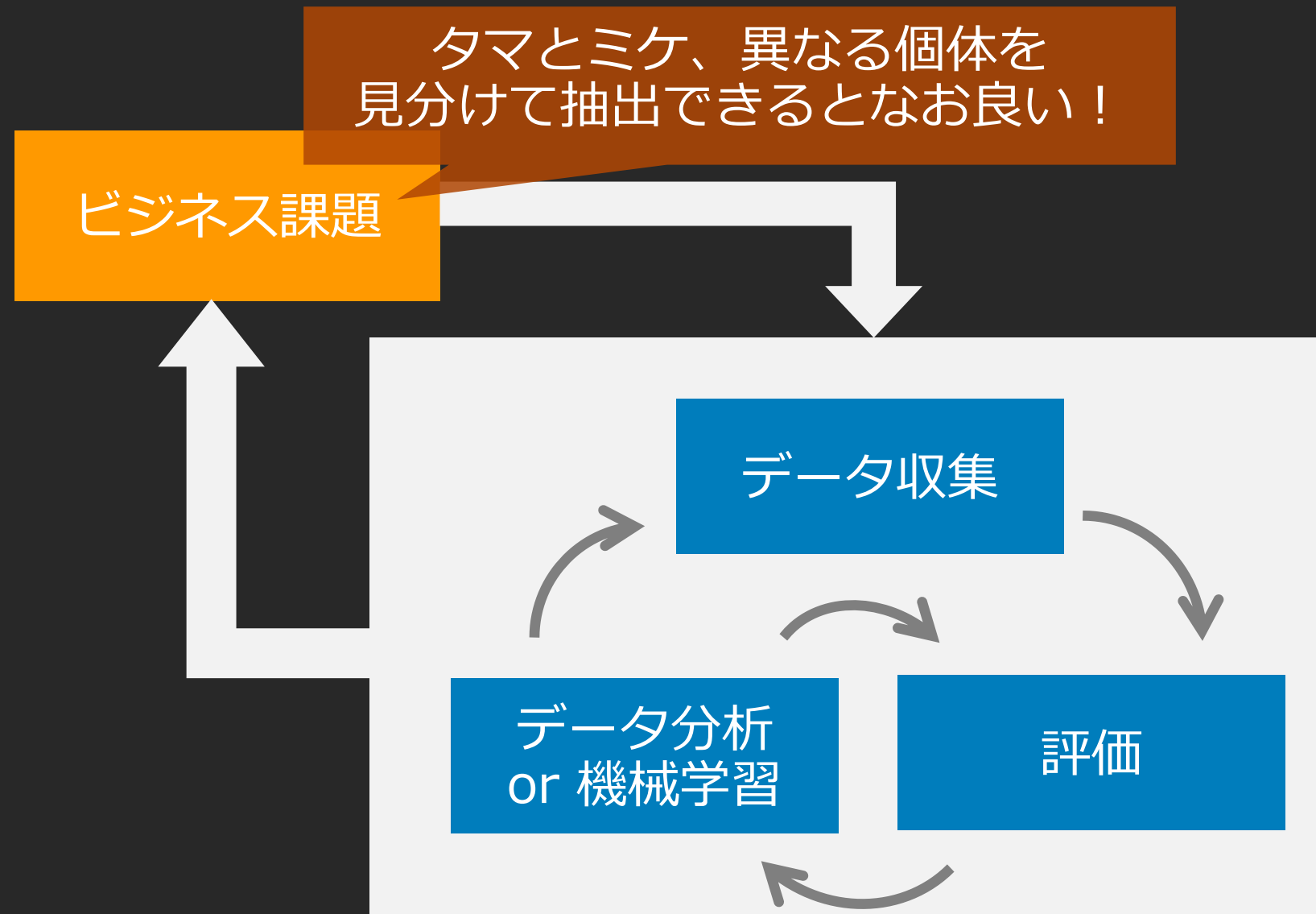
# 猫識別ソリューションを開発する手順

実現したい内容に従って機械学習モデルを開発し、実際に適用してみて課題解決できているかを評価します



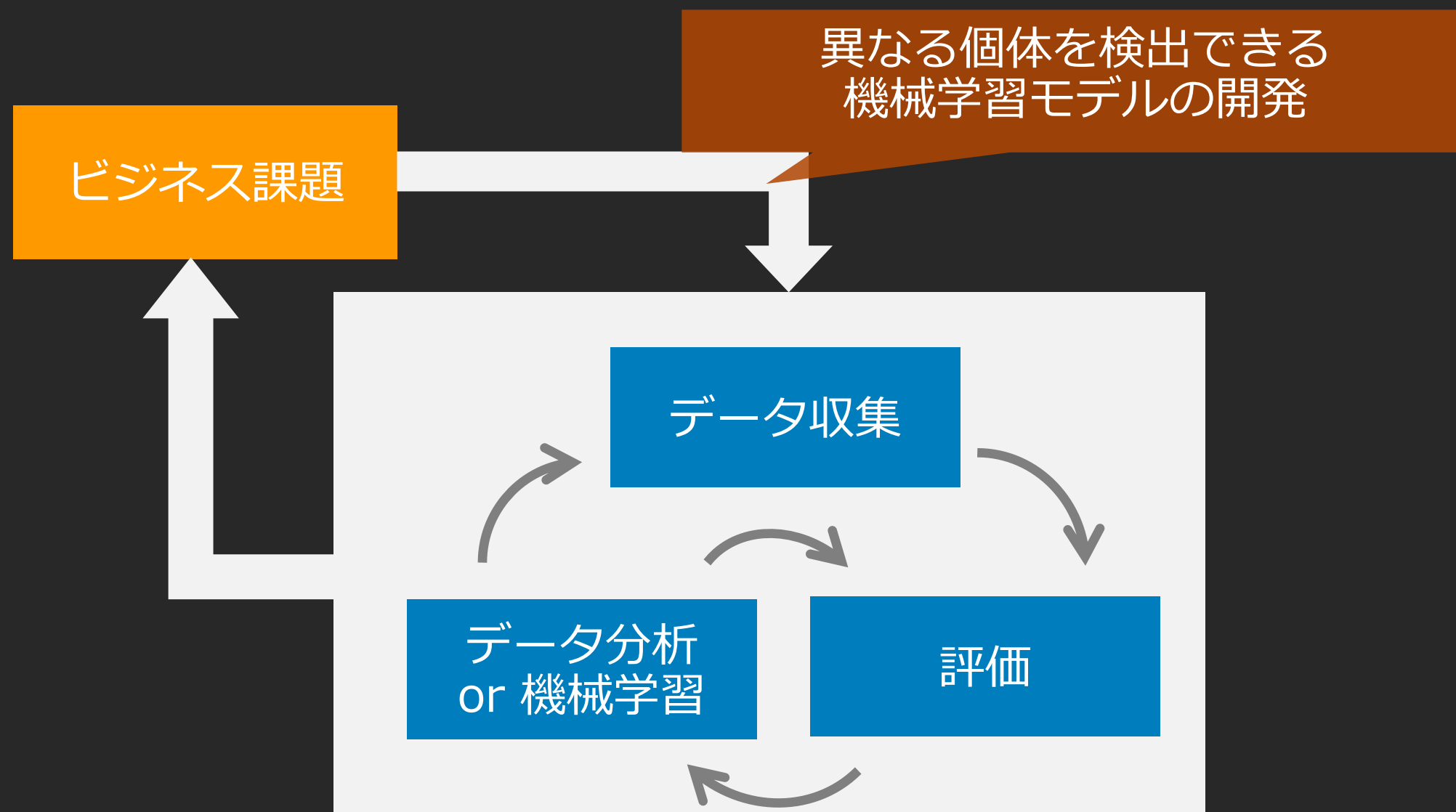
# 猫識別ソリューションを開発する手順

実現したい内容に従って機械学習モデルを開発し、実際に適用してみて課題解決できているかを評価します



# 猫識別ソリューションを開発する手順

実現したい内容に従って機械学習モデルを開発し、実際に適用してみて課題解決できているかを評価します



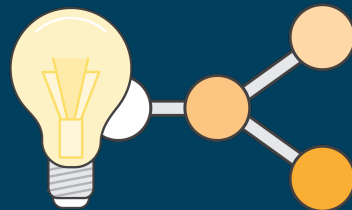
# 猫識別のための機械学習モデルを開発する手順

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する

## 開発段階



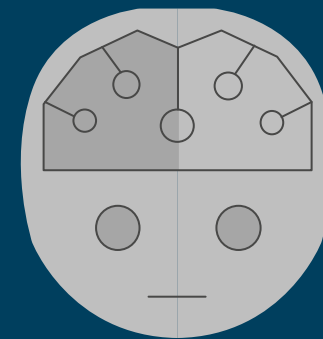
学習用データ



機械学習アルゴリズム

学習

## 活用段階



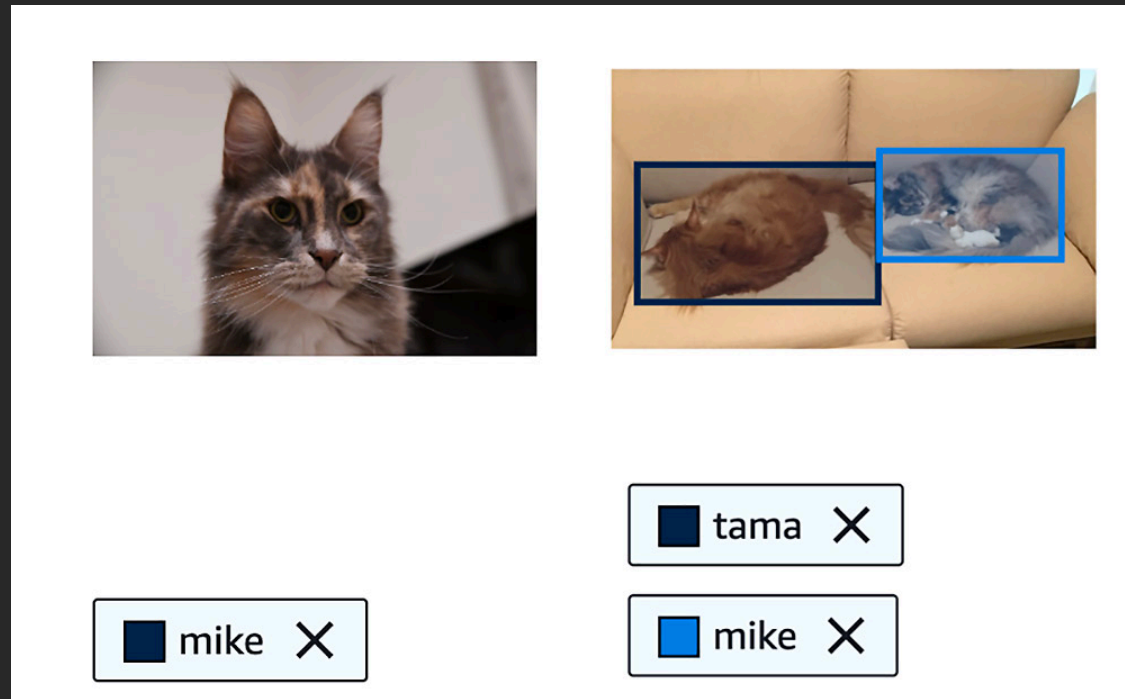
機械学習モデル

# 猫識別のための機械学習モデルを開発する手順

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する



# Amazon Rekognition Custom Labels とは



学習用データ



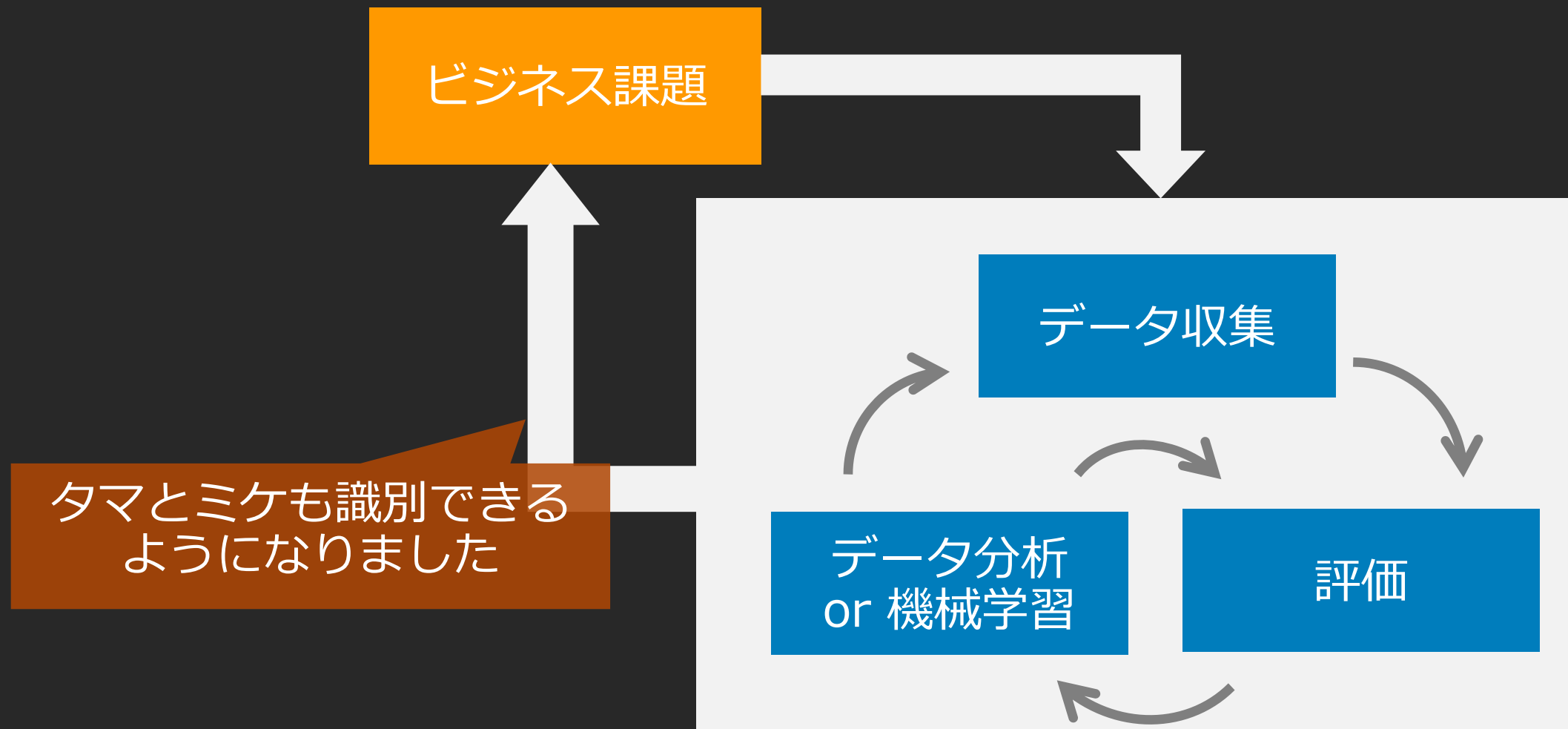
推論結果

- Amazon Rekognition で用意されているアルゴリズムに独自の画像データを学習させて機械学習モデルを作成することができる
- ラベル付けを行うために SageMaker Ground Truth をご活用頂くこともできます



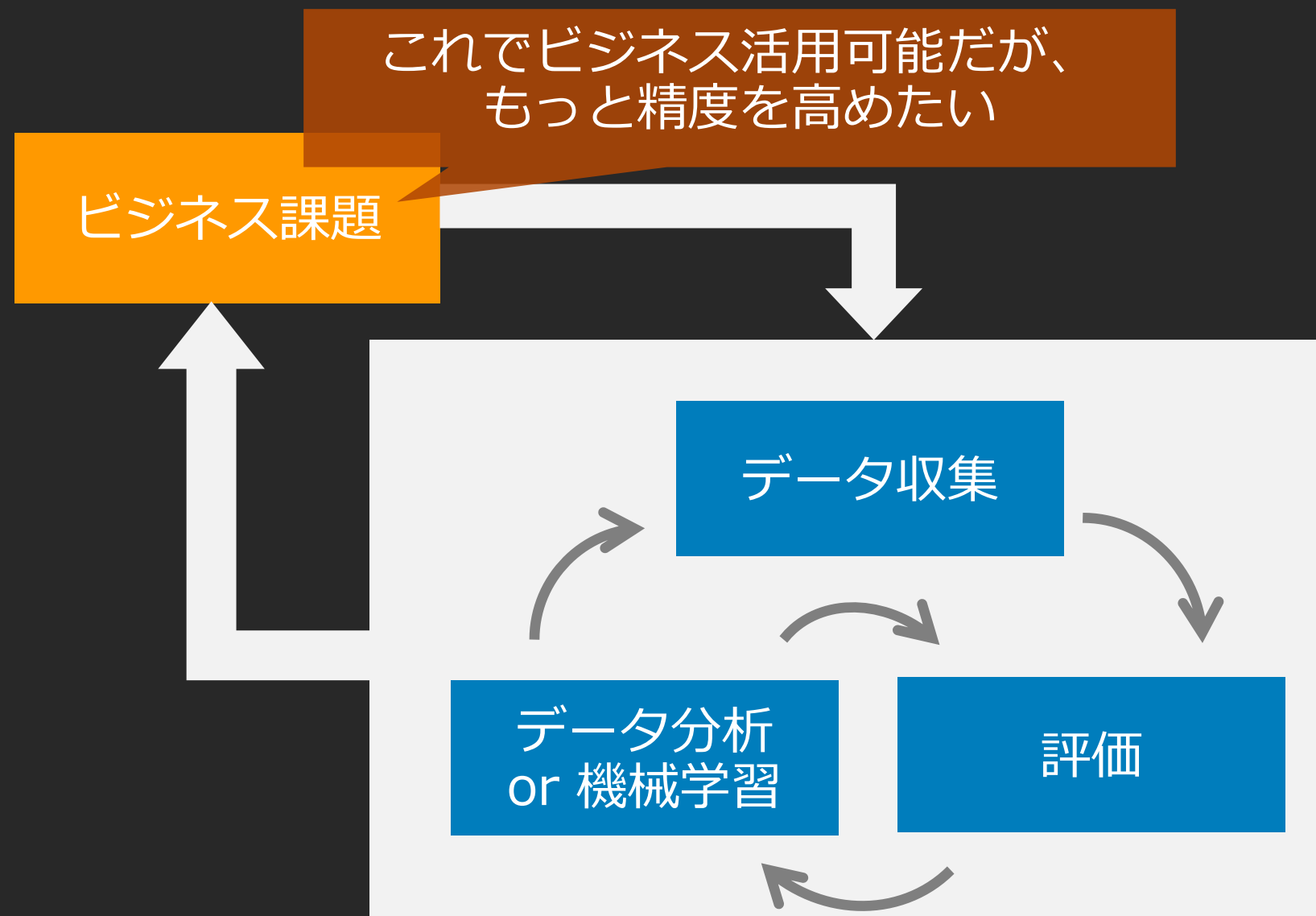
# 猫識別ソリューションを開発する手順

実現したい内容に従って機械学習モデルを開発し、実際に適用してみて課題解決できているかを評価します



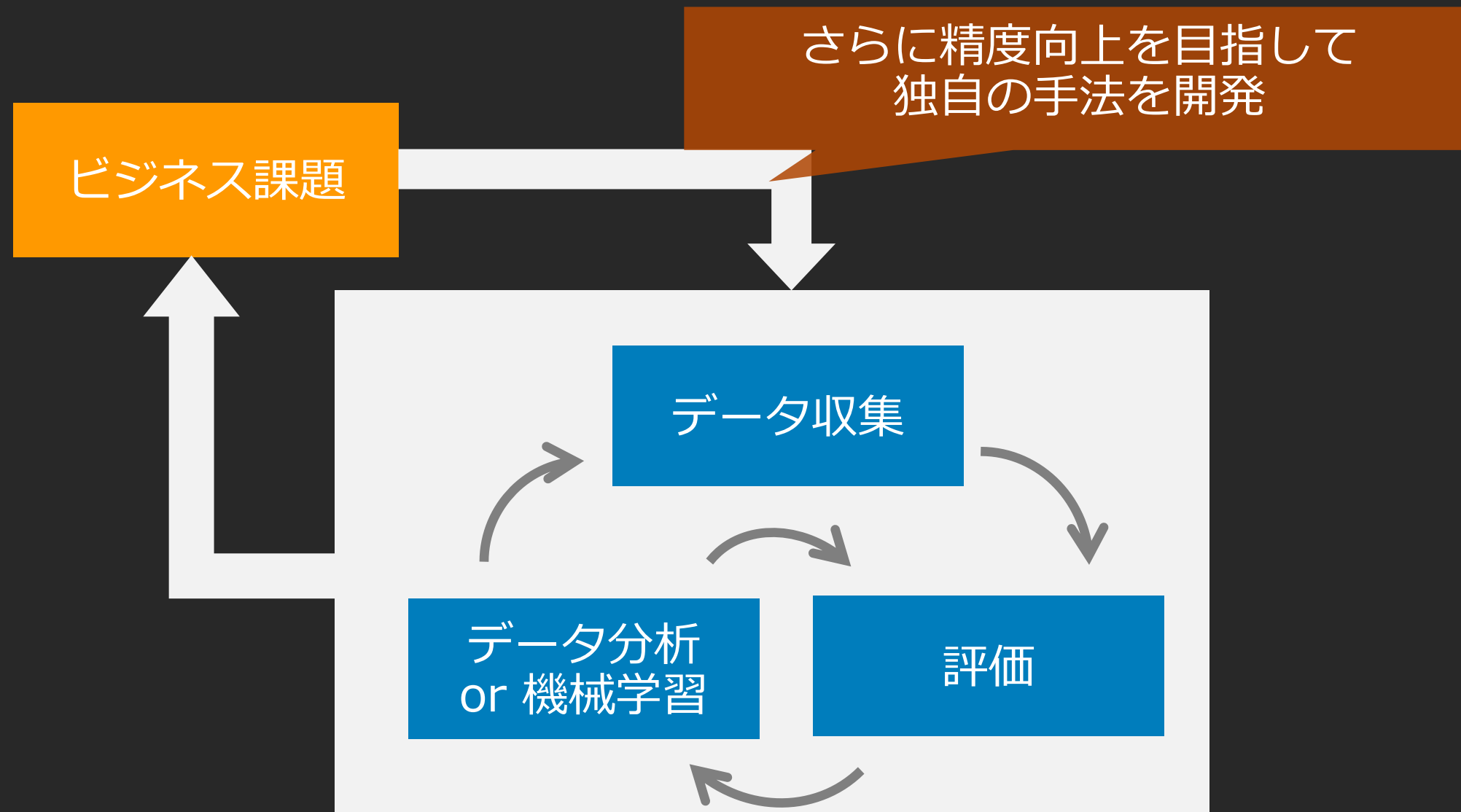
# 猫識別ソリューションを開発する手順

実現したい内容に従って機械学習モデルを開発し、実際に適用してみて課題解決できているかを評価します



# 猫識別ソリューションを開発する手順

実現したい内容に従って機械学習モデルを開発し、実際に適用してみて課題解決できているかを評価します



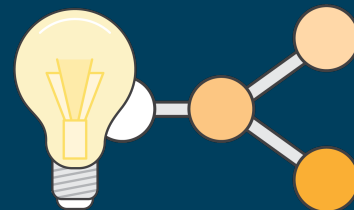
# 猫識別のための機械学習モデルを開発する手順

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する

## 開発段階



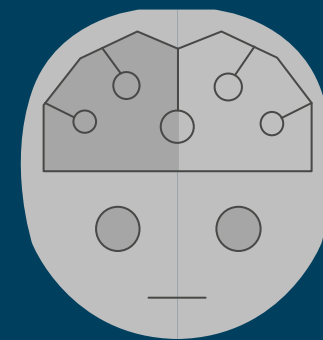
学習用データ



機械学習アルゴリズム

学習

## 活用段階



機械学習モデル

# 猫識別のための機械学習モデルを開発する手順

1. モデル開発を行わず学習済のモデルを活用する
2. データを準備して提供されているモデルを学習させて活用する
3. 独自のモデルとデータを準備して学習して活用する



# Amazon SageMaker でできること

準備

構築

トレーニング&チューニング

デプロイ&管理

機械学習のための Web ベース統合開発環境 (IDE)

モデルの自動構築・トレーニング

完全マネージドな  
データ処理ジョブと  
ラベリングの  
ワークフロー

101011010  
010101010  
000011110

データ収集・準備

1クリックの  
コラボレーションと  
ビルトインの高性能な  
アルゴリズム・モデル



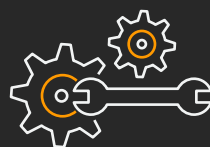
機械学習  
アルゴリズム構築

1クリックで  
トレーニング



トレーニング環境の  
セットアップ

デバッグ・最適化



モデルの  
トレーニング・  
デバッグ・  
チューニング

可視化して  
トラッキング・比較



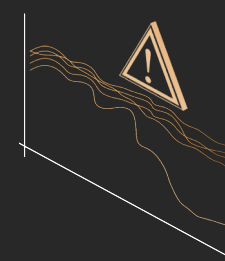
実験管理

1クリックで  
デプロイと  
オートスケーリング



本番環境への  
デプロイ

コンセプト  
ドリフトの  
自動検知



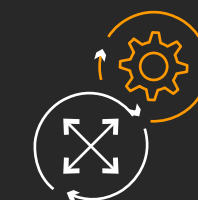
モデルの  
監視

予測結果に  
人手のレビュー



推論結果の検証

オートスケーリング  
によるコスト削減



本番環境での  
スケーリングと管理



まとめ

# 本セッションのまとめ

## 1. 常にビジネスKPIへの貢献へ集中する

- なるべく解決しなければいけない本質的な課題、ビジネス価値の高いものに取り組みめるようにするべき

## 2. 2種類の試行錯誤を高速に行えるように

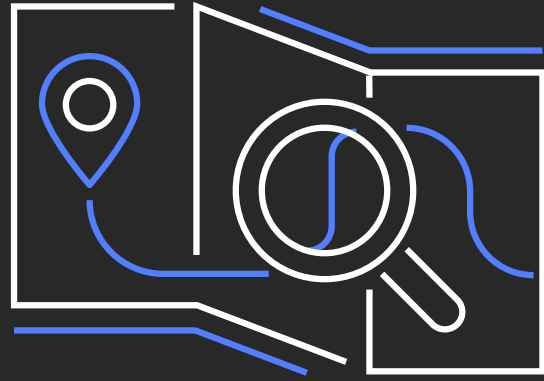
- 機械学習モデルの開発は手間のかからないものから徐々に高度化
- 差別化につながらない重労働は最小限に

## 3. 試行錯誤を柔軟に行える環境づくり

- 沢山ある選択肢から、適切なサービスを選んで活用
- データレイクを構築することで、多様なサービスでのデータ活用を容易に

# Thank you!

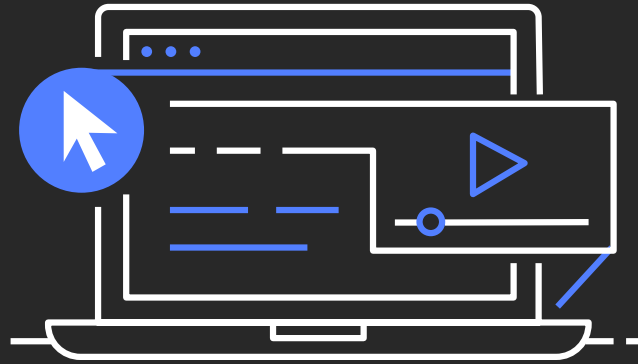
# AWS トレーニングと認定



## 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](https://aws.amazon.com/training)