

開発者向けオンラインセミナー

ベクトルであそぼう！

インターシステムズジャパン株式会社
ソリューションアーキテクト
鵜浦 永光
2025年7月29日





Agenda



01 写真から魚の名前をあててみる

02 ベクトルを“見える化”する

03 データの集まりを見る

04 変なデータ(=アノマリ)を見つける



写真から魚の名前をあてる

この魚は何だろう？

釣り人ビギナー Webinar サポートAIを作ろう！

ベクトル検索やRAGの活用で答えの精度を向上

第1回 ベクトル検索入門

第2回 ベクトルであそぼう！

← 本日開催！

第3回 RAGと生成AIであそぼう！

この魚は何だろう？

これは
アジです

参考情報：
美味しい季節は夏！



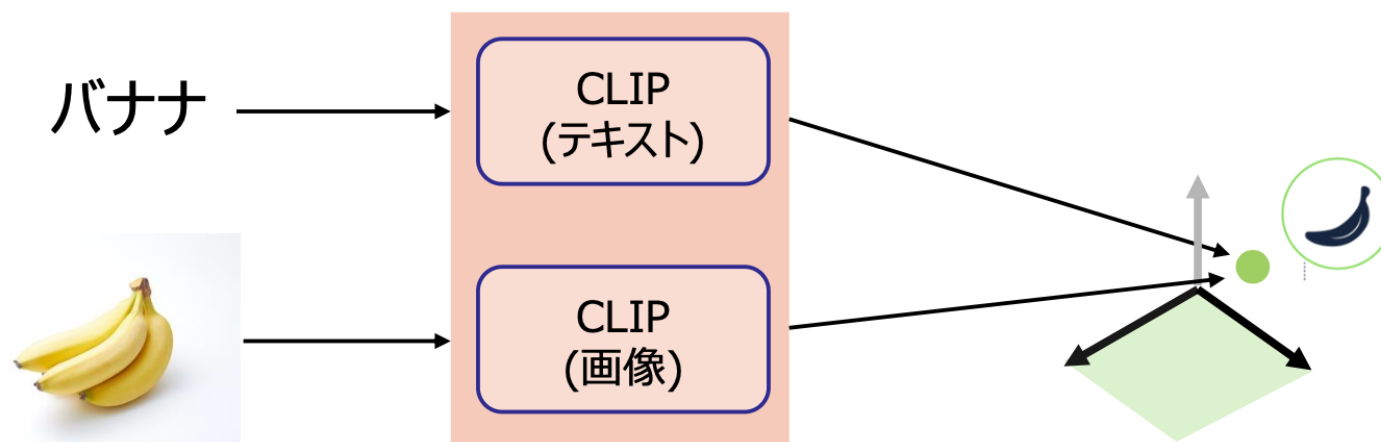
アジの塩焼きの
作り方です。



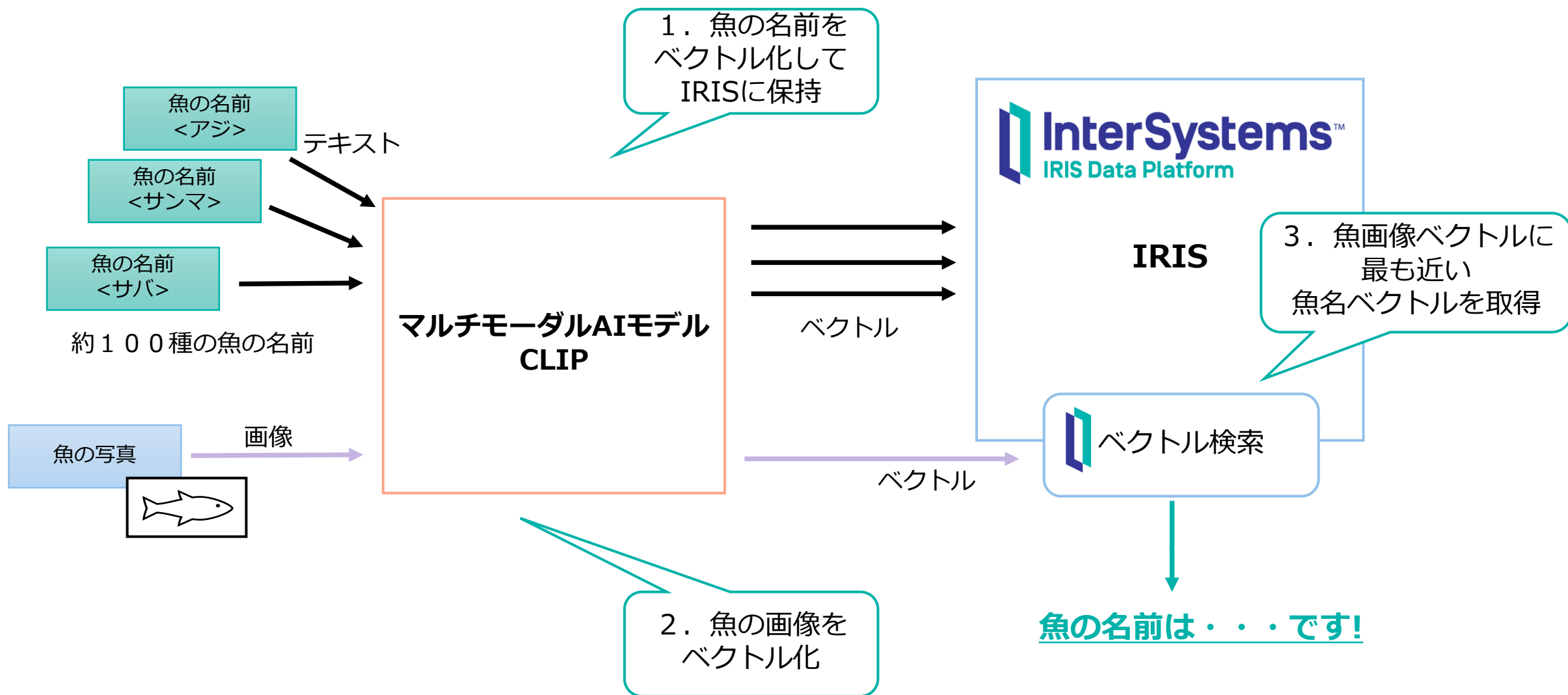
CLIP



- CLIP – Contrastive Language-Image Pre-training
- Open AIが開発したマルチモーダル基盤モデル。
- 画像とその説明文（テキスト）のペアを予測できるように学習したもの。その結果、与えられたテキストに最も合致する画像を見つけたり、逆に、与えられた画像を最もよく説明するテキストを見つけたりすることが可能になった。



写真から魚の名前をあてる実験構成





写真から魚の名前をあてる

DEMO

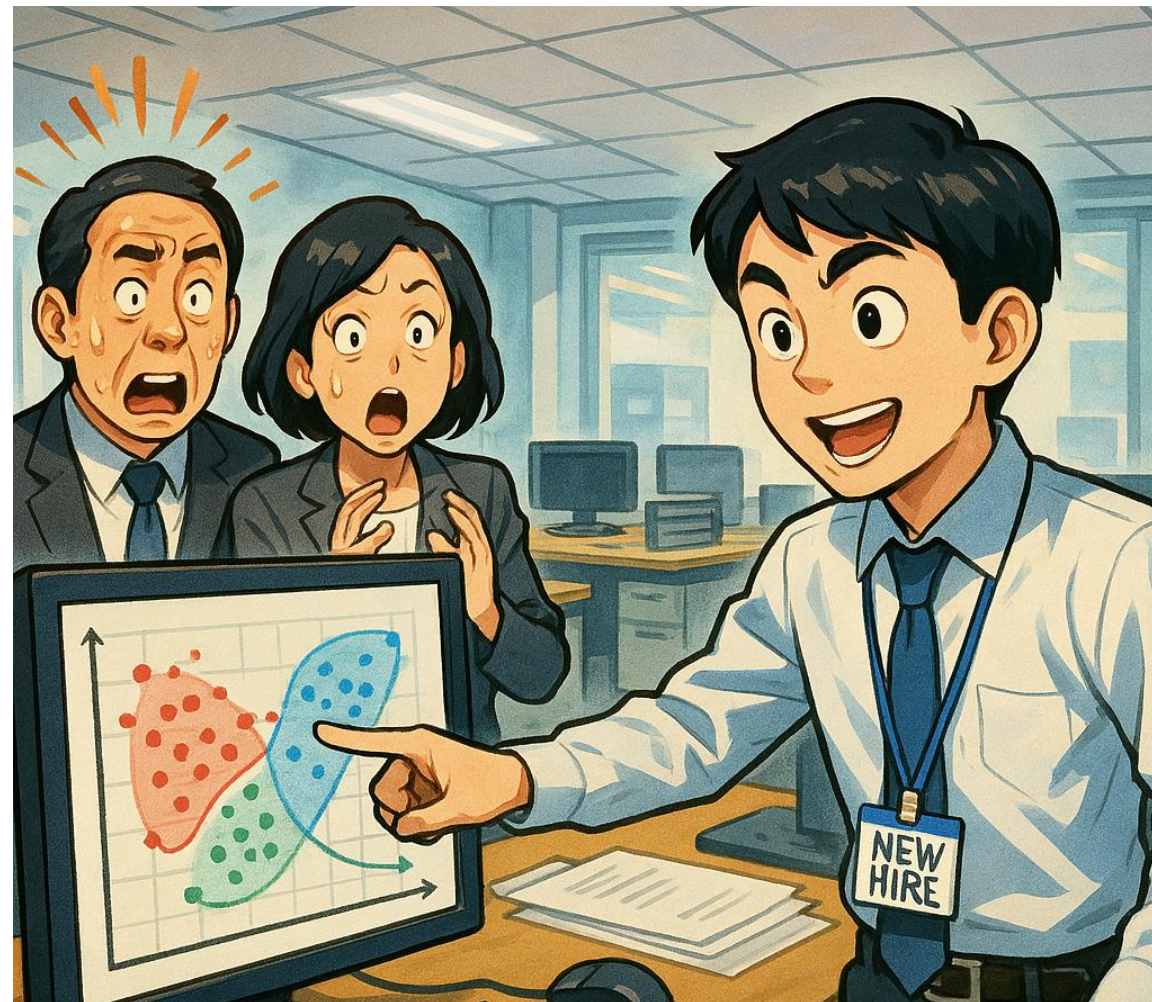


ゼロショット分類



特定のラベル（例：魚の名前）に対して事前に学習していなくても、モデルが正しく推論・分類できる能力のこと

※従来の画像分類モデルでは、
「サバ」を当てるには、サバの画像を多数用意し
「これはサバです」と正しくラベル付けして
学習させる必要があった



ファインチューニング



すでに事前学習されたAIモデルに対して
特定の目的やタスクに合わせた
追加学習を行うプロセス

※今回のデモでは行っていません





ベクトルを“見える化”する

ベクトル - AIモデルにとっては特徴、人にとっては謎の数字



Name	Vectors	
マアジ	.0030519193969666957856,.019338253885507583618, -.026509894058108329772,-.0056376024149358272552,-.00...	
カツオ	.013707441277801990509,.020661199465394020081, -.044357202911,-.01809...	
サクラマス	-0.032789113331, -0.04844971461,-0.0410...	
アユ	.014848685823380947113,.0040654172189533710479, -.033739592880010604858,-.030808333307504653931,-.015...	
ギンダラ	.052557989954948425292,.012833195738494396209, .011079670861363410949,-.012871518731117248536,-.00877...	

数百 ~ 数万次元 の数値

ベクトルを“見える化”する利点



データ / AIモデル に対するより深い理解



人に伝える (レポート)

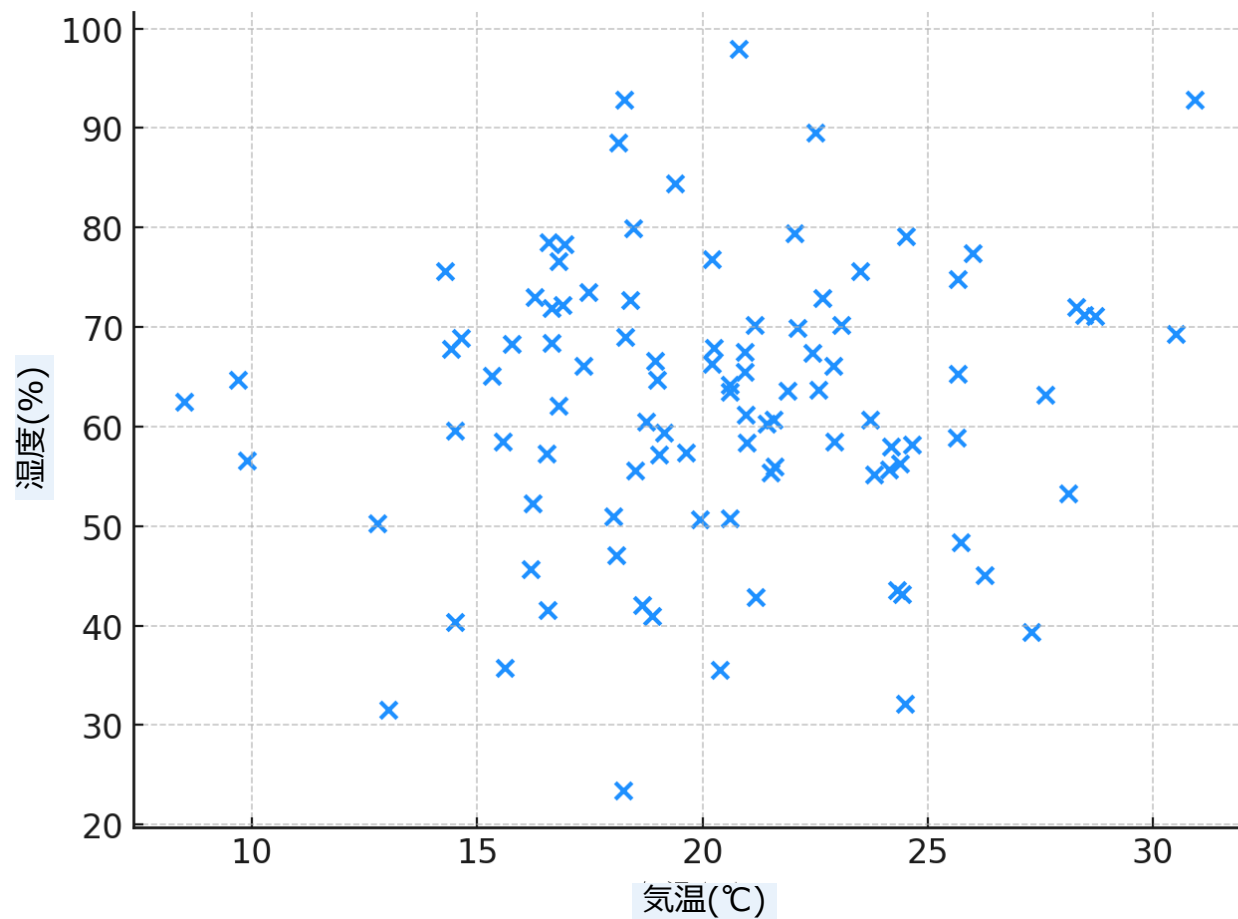


トラブルシューティング

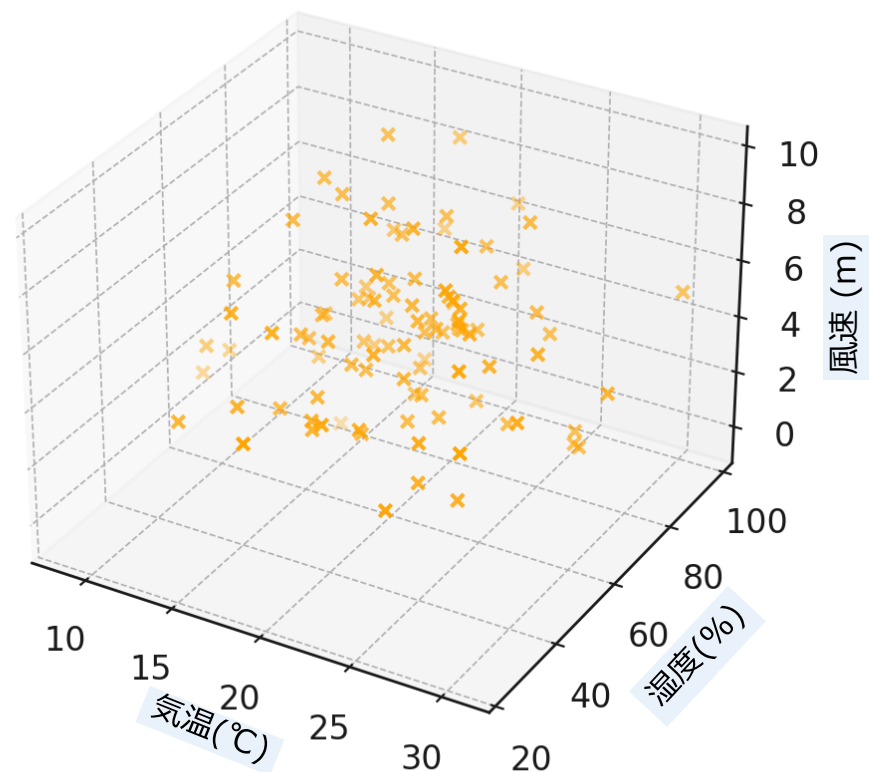
人間が目で見えて把握できる次元は2次元か3次元



2次元



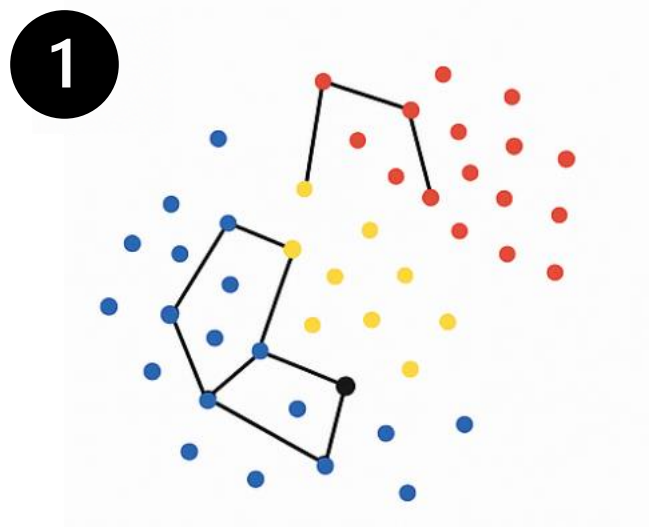
3次元



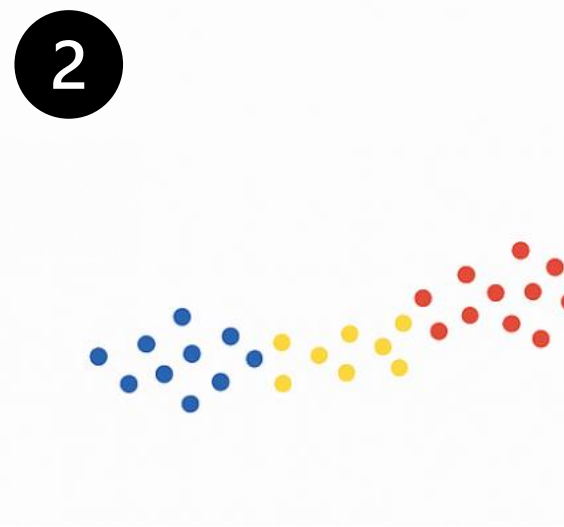
次元を削減



UMAP - 次元削減のアルゴリズム



高次元空間での“近さ”を
グラフにする



低次元空間(2次元や3次元)で
同じ構造になるように最適化する

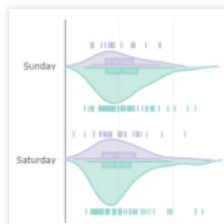
インタラクティブなグラフ描画 - Plotly

<https://plotly.com/python/>

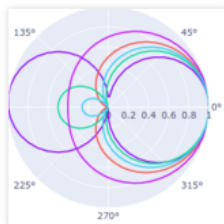
plotly | Graphing Libraries

Fundamentals

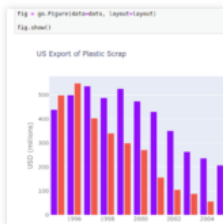
More Fundamentals »



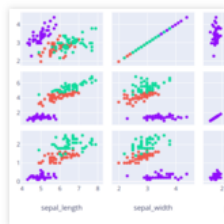
The Figure Data Structure



Creating and Updating Figures



Displaying Figures



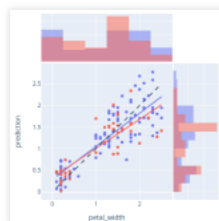
Plotly Express



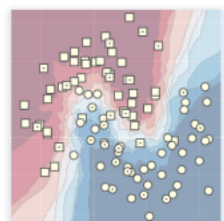
Analytical Apps with Dash

Artificial Intelligence and Machine Learning

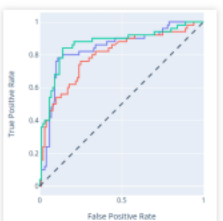
More AI and ML »



ML Regression



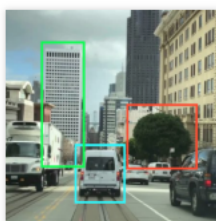
kNN Classification



ROC and PR Curves



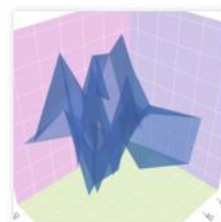
PCA Visualization



AI/ML Apps with Dash

3D Charts

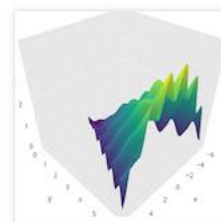
More 3D Charts »



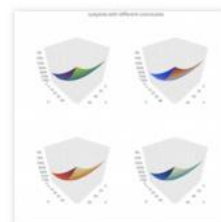
3D Axes



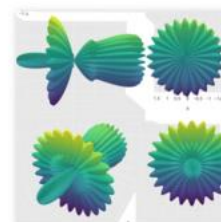
3D Scatter Plots



3D Surface Plots



3D Subplots



3D Camera Controls



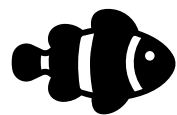
ベクトルを“見える化”する

DEMO



データの集まりを見る

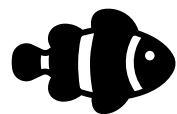
クラスタリングのユースケース



顧客セグメンテーション




商品のグルーピング





アンケート結果の傾向分析

クラスタリングの代表的なアルゴリズム – “K-Means”



 データをK個のクラスタ(似た者同士のグループ)に分類する

 教師なし学習
「こういうグループ分けをなさい」といった分類ルールを与えず、
モデル自身がデータの特徴に基づいてグループ分けを行う

 クラスタごとの平均(中心点) がクラスタを代表する値となる

最初にKの数(クラスタの数)を決める

K-Means クラスタリングのステップ

K=4



K個の中心点をランダムに決める

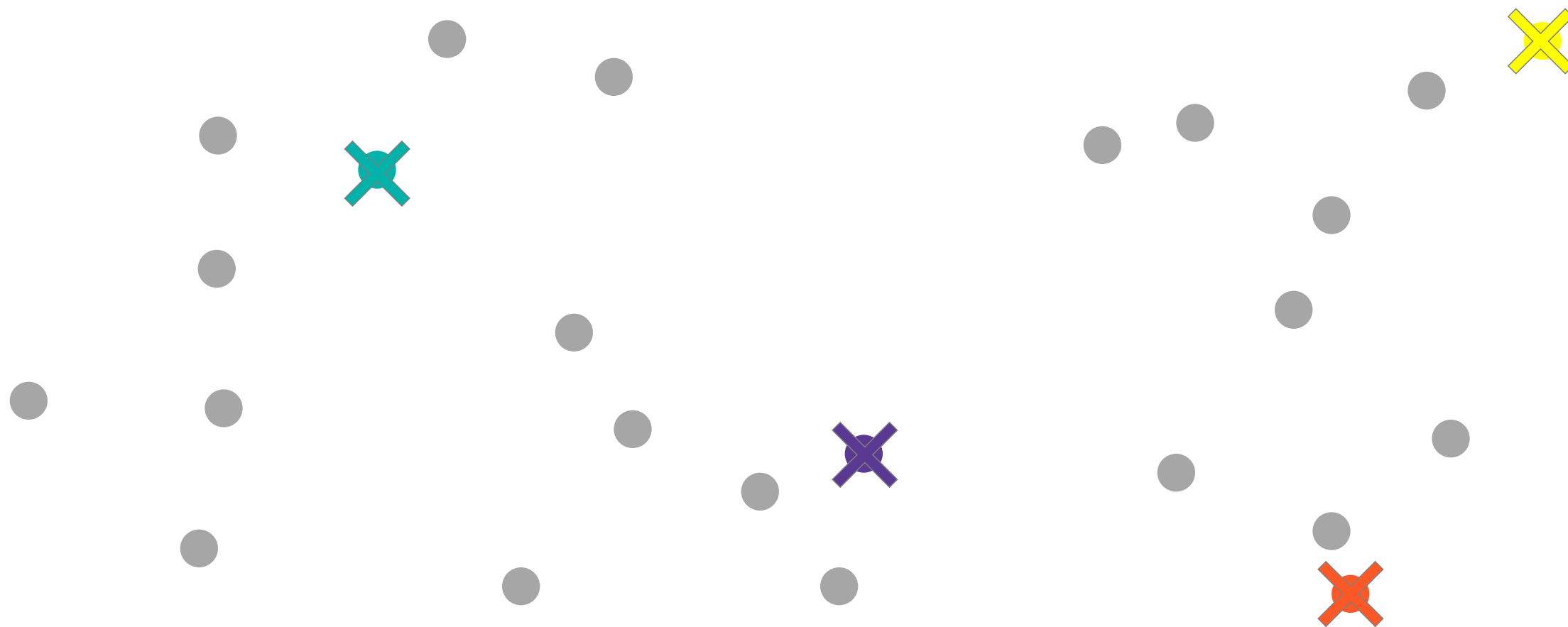
K-Means クラスタリングのステップ



すべての点を、いちばん近い中心点のクラスターに割り当てる

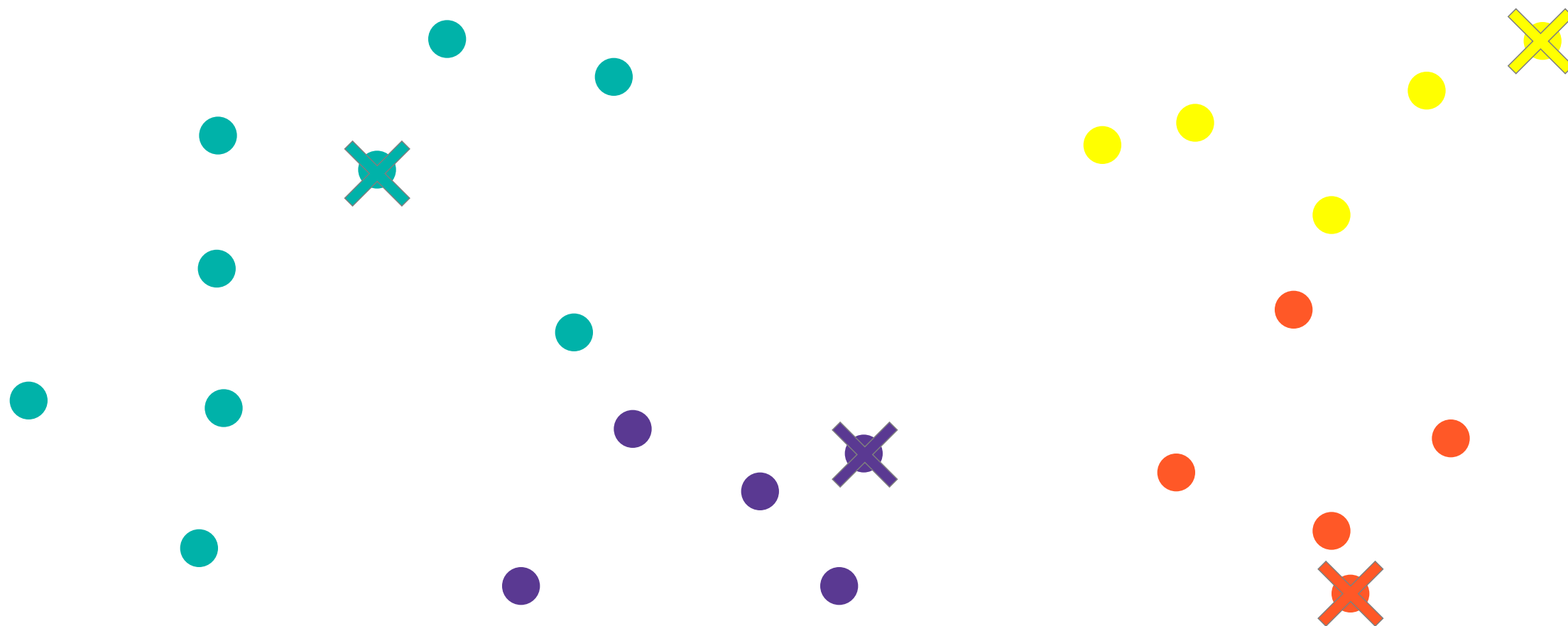


K-Means クラスタリングのステップ



各クラスタの真ん中を計算し、中心点を更新する

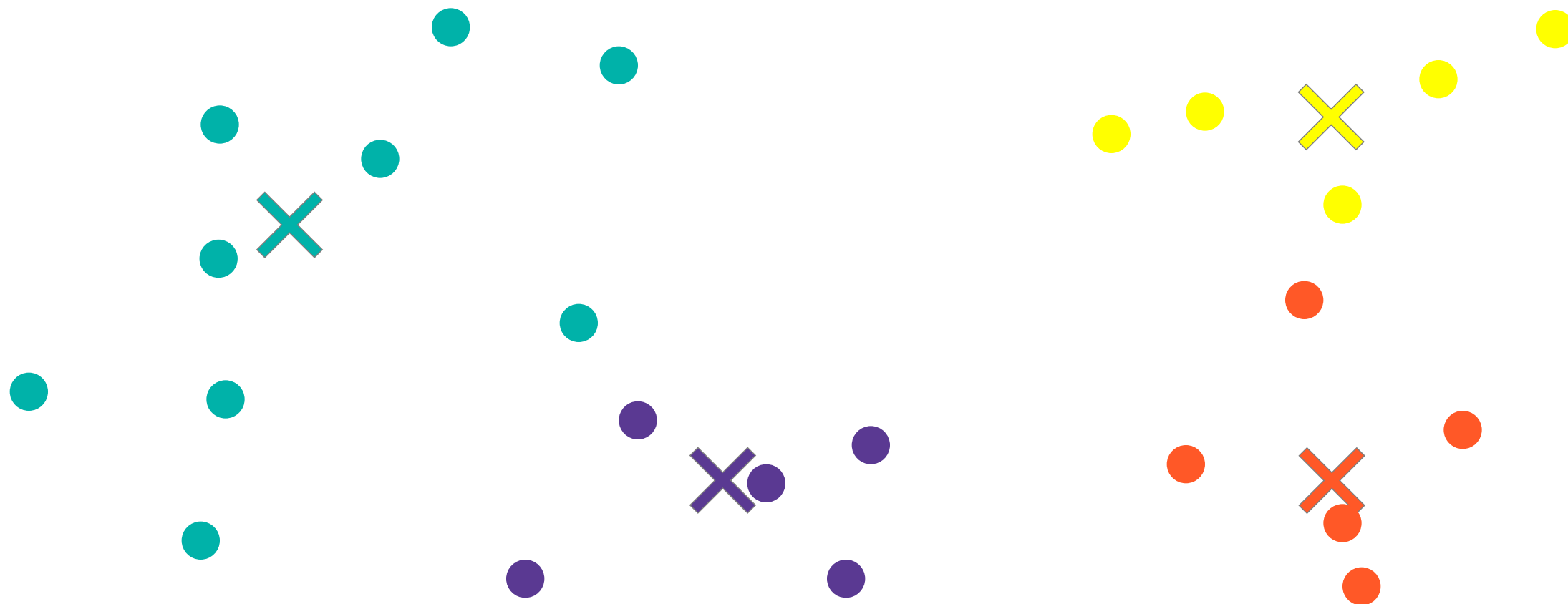
K-Means クラスタリングのステップ



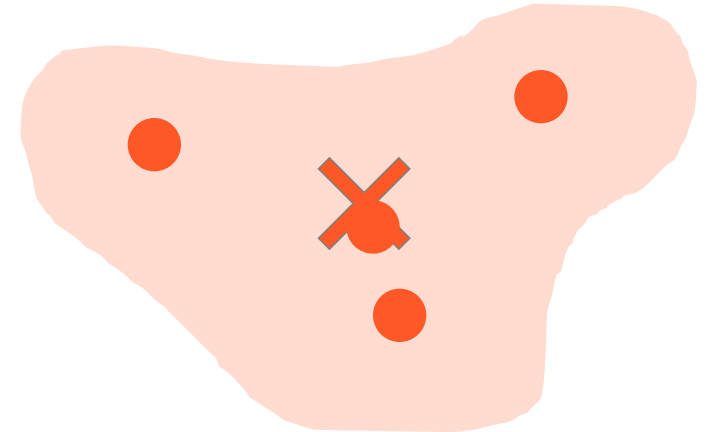
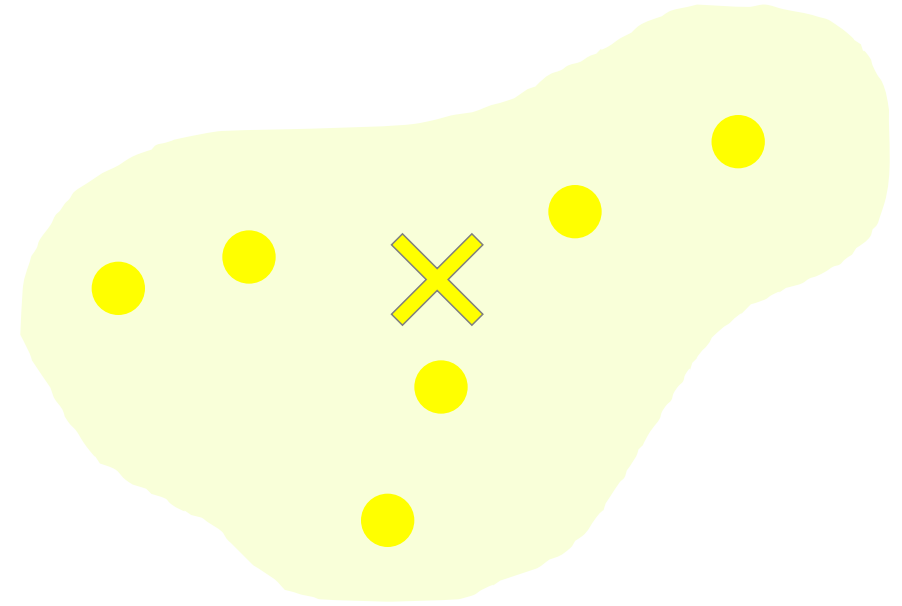
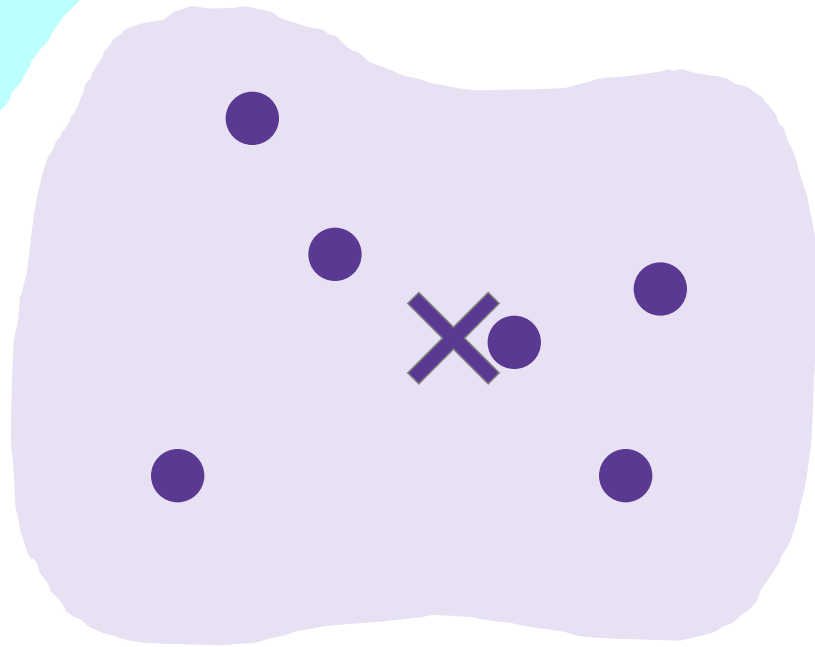
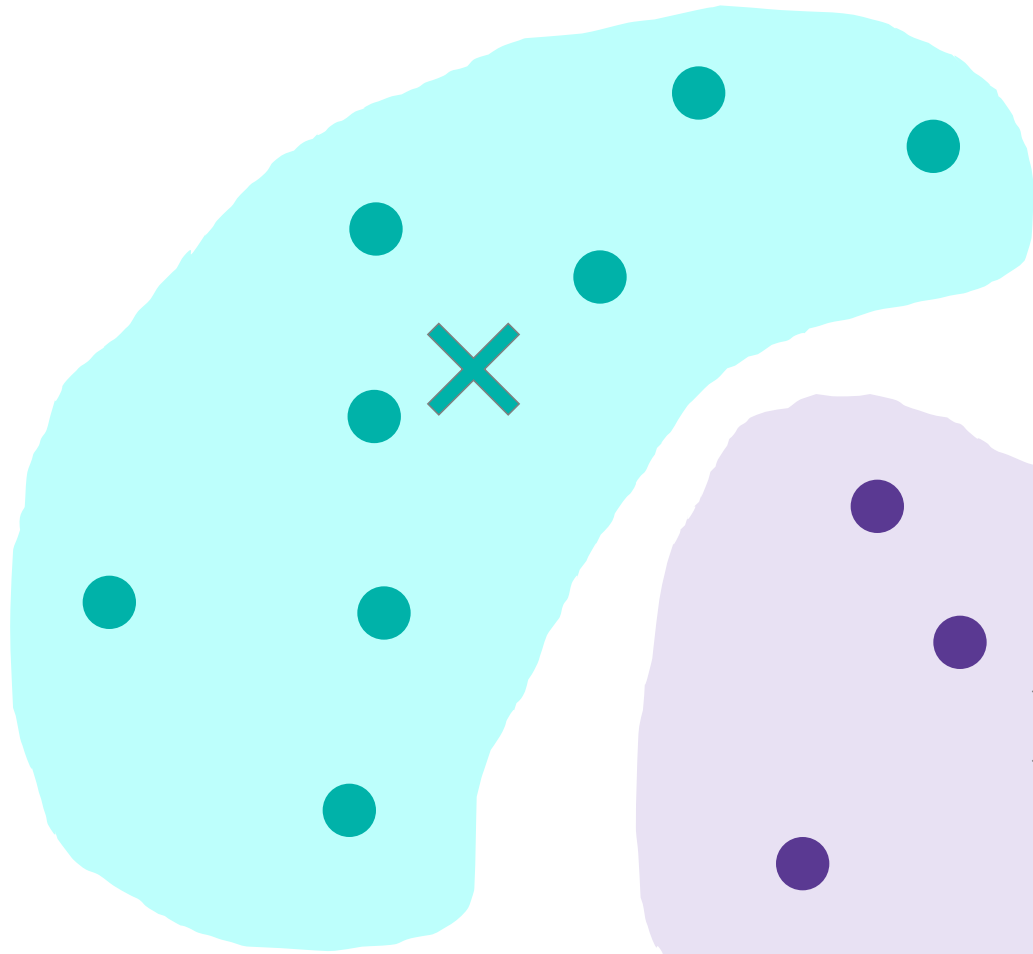
再び全ての点を、いちばん近い中心点のクラスターに割り当てる



K-Means クラスタリングのステップ



K(4)個のクラスタが完成

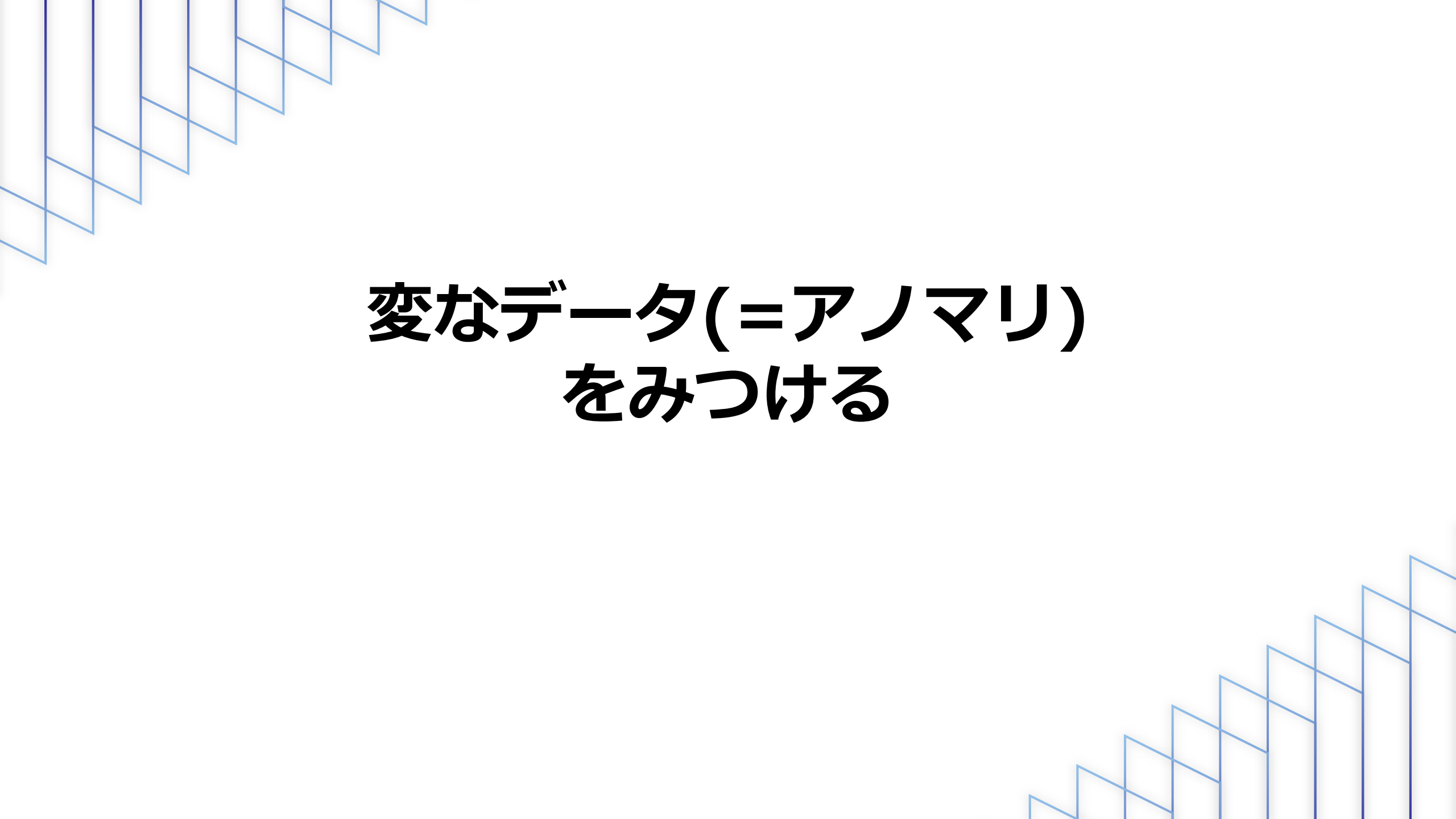




データの集まりを見る

DEMO





**変なデータ(=アノマリ)
を見つける**

アノマリ検知のユースケース



製造ラインの画像検査



セキュリティ攻撃検知



SNS等での不適切コンテンツ検出

アノマリ検知の方法の例

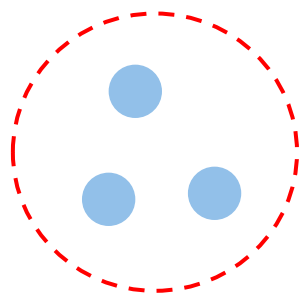
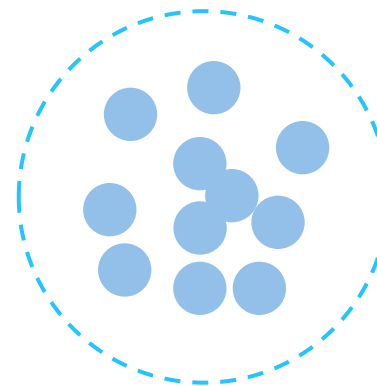
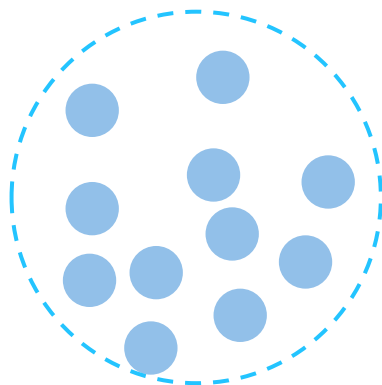
K-Meansを活用したアノマリ検知

教師なし アノマリ検知

- ・ ラベルなしデータを使用
- ・ データの分布特性に基づいて異常を検出

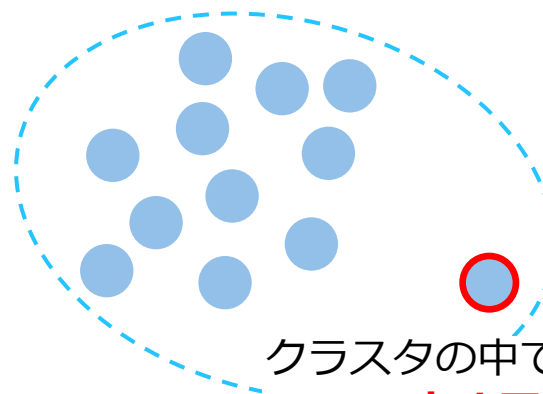
どのクラスタの中心点からも遠い

グローバルアノマリ



データの少ないクラスタ(=少数派)

グループアノマリ

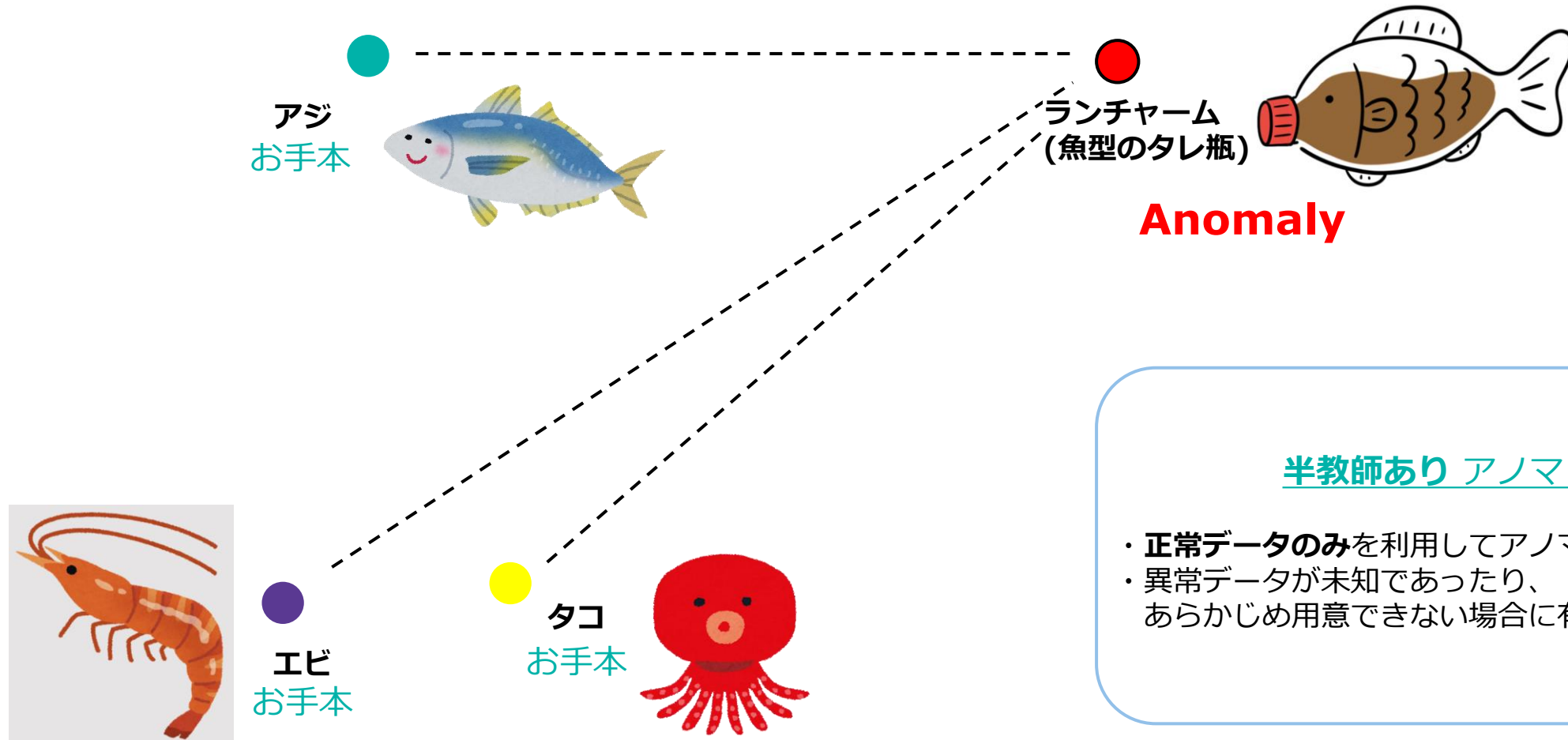


クラスタの中での外れ値

ローカルアノマリ

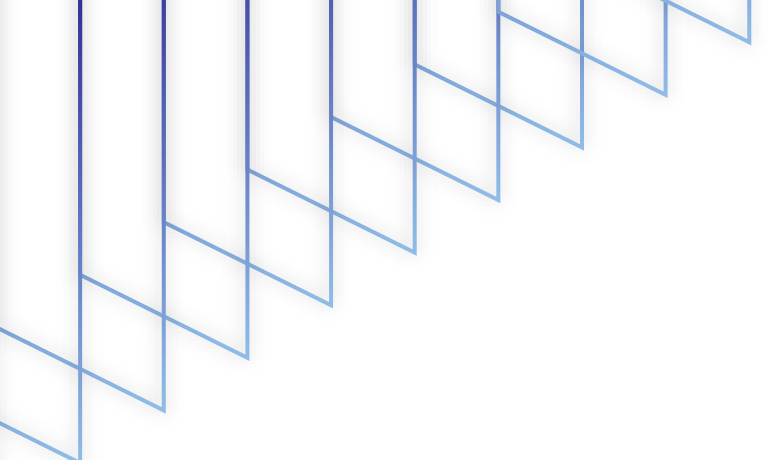
アノマリ検知方法の例

お手本データのいずれにも類似していないものをアノマリとする



半教師あり アノマリ検知

- ・ 正常データのみを利用してアノマリ検知
- ・ 異常データが未知であったり、あらかじめ用意できない場合に有効



変なデータ(=アノマリー) を見つける

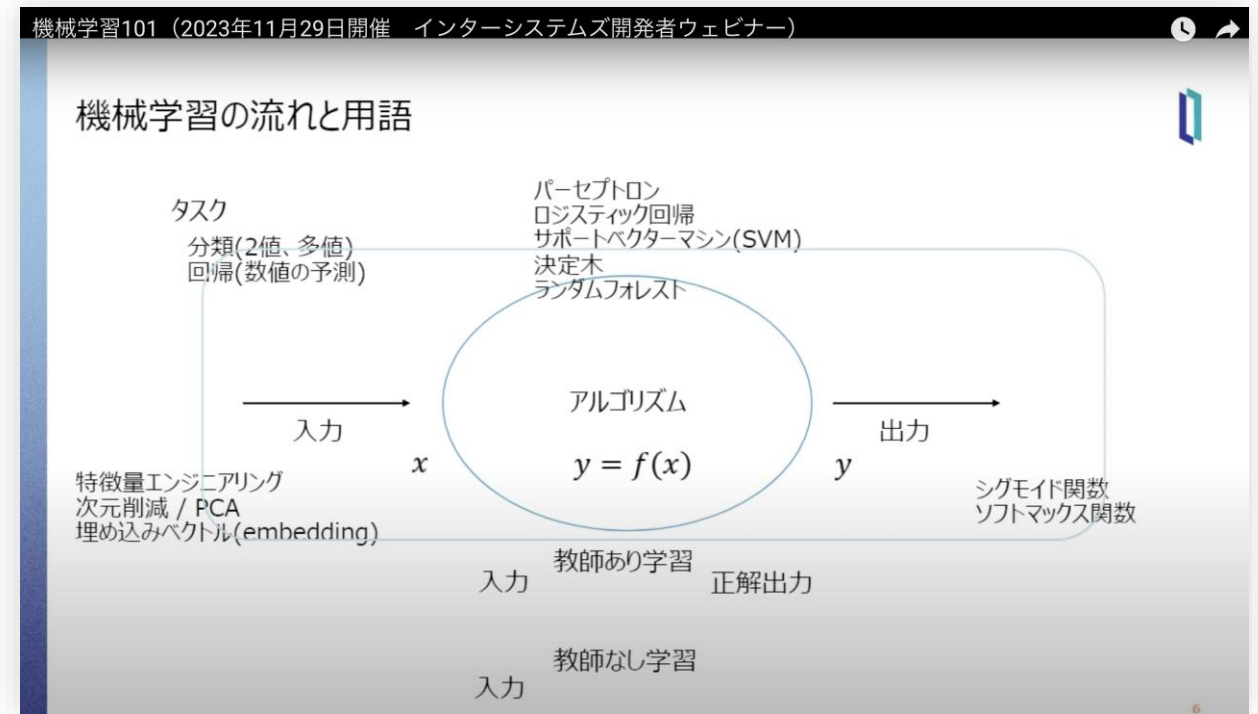
DEMO



「機械学習101」もおすすめ

インターシステムズ 開発者ウェビナー
『機械学習101』

[内容]
機械学習 / 深層学習 /
教師あり学習 / 教師なし学習 /
性能評価の基本（分類・回帰）など



<https://www.youtube.com/watch?v=47bP5-AtBVU>

インターシステムズ 機械学習101



コンテスト／ミート
アップに向けた
ウェビナーシリーズ
この後も続きます！

釣り人ビギナー Webinar サポートAIを作ろう！

ベクトル検索やRAGの活用で答えの精度を向上

第1回 ベクトル検索入門

第2回 ベクトルであそぼう！ 7/29開催！

第3回 RAGと生成AIであそぼう！

この魚は何だろう？

これは
アジです

参考情報：
美味しい季節は夏！



アジの塩焼きの
作り方です。

