

# みんなのKaggle講座

## Section4





# Section4の概要




# 講座の内容

Section1. Kaggleの概要

Section2. 機械学習とKaggle

Section3. 精度向上のためのテクニック

 **Section4. Titanicの先へ**

# 今回の内容

1. Section4の概要
2. Tabular Playground Seriesにトライしよう！
3. AutoMLの概要
4. AutoMLによる機械学習の自動化
5. 最後に

# 教材の紹介

- **Pythonの基礎:**

python\_basic

- **Section4の教材:**

01\_tabular\_playground\_series.ipynb

02\_automl\_houseprices.ipynb

# 演習の解答 Section3

section\_3/04\_exercise.ipynb



**Tabular Playground Seriesにトライしよう！**



# Tabular Playground Seriesとは？

- **Tabular Playground Series**

- 毎月開催。上位3チームにKaggleグッズを提供。
- Titanicよりやや難易度が高く、次の課題として最適。



# Tabular Playground Series - Jan 2022とは？

- **Tabular Playground Series - Jan 2022**

- Kaggleグッズを販売する2つの（架空の）店舗チェーン
- 北欧の3カ国、3つの商品
- 数年分の販売データから、未来の販売数を予測
- <https://www.kaggle.com/c/tabular-playground-series-jan-2022/>

# Tabular Playground Seriesにトライしよう！

- 01\_tabular\_playground\_series.ipynb

# LOMILの概要

# AutoMLとは？

- **AutoML (Automated Machine Learning、自動機械学習)**
  - 機械学習モデルの設計や構築を自動化すること、もしくはそのための概念全般
  - 機械学習の専門家でなくても、高機能な機械学習の機能を利用可能に
  - データの前処理、モデリングなどを自動化し、開発にかかるコスト削減や期間短縮を可能に



# AutoMLで自動化できること

- データの準備

- データを分析に適した形式に整えます

- モデリング

- モデルを構築、訓練し予測モデルを作成します

- 評価

- 作成したモデルが、目標の達成に適しているかどうかを評価します

- デプロイ

- 作成したモデルをサービスで利用可能な状態にします

- データ分析の結果を、目標の達成に向けたアクションに落とし込みます

# AutoMLを提供するWebサービス

- **Cloud AutoML (Google)**
  - クラウド上で画像認識、動画認識、自然言語処理、翻訳、etc...
  - 有料 (無料枠あり)
- **Automated ML (Microsoft)**
- **AutoAI (IBM)**
- **Prediction One (Sony)**
- **DataRobot**
- **H2O Driverless AI**
- **etc...**

# AutoMLのOSSライブラリ

- **PyCaret**  
→ <https://pycaret.org/>
- **TPOT**  
→ <https://github.com/EpistasisLab/tpot>
- **H2O**  
→ <http://docs.h2o.ai/>
- **Auto-Keras**  
→ <https://autokeras.com/>
- **Auto-sklearn**  
→ <https://www.automl.org/automl/auto-sklearn/>
- **AutoGluon**  
→ <https://auto.gluon.ai/>
- **etc...**

# PyCaretとは

- **PyCaret**

- 簡単に使えて、ビジネスや研究などの現場に導入しやすい
- 10行にも満たないコードで、前処理、モデルの作成とチューニング、予測、分析などが可能
- アソシエーション分析、異常検知に対応
- 自然言語処理にも対応（2021年6月現在は英語のみ対応）
- <https://pycaret.org/>



# AutoMLによる機械学習の自動化



# AutoMLの実践 -回帰-

- **House Prices - Advanced Regression Techniques**

住宅価格の予測にトライしよう！

<https://www.kaggle.com/c/house-prices-advanced-regression-techniques>

02\_automl\_houseprices.ipynb

# AutoMLによる機械学習の自動化



- 02\_automl\_houseprices.ipynb



# 最後に





# さらに学びたい方へ...

- **PythonではじめるKaggleスタートブック**（講談社）

→ <https://www.kspub.co.jp/book/detail/5190067.html>

- **Kaggleで勝つデータ分析の技術**（技術評論社）

→ <https://gihyo.jp/book/2019/978-4-297-10843-4>

# さらに学びたい方へ...



## Udemy講座

みんなのデータサイエンス講座 -  
Python、Colab、Kaggleで基礎から学び親しむ「データ」の世界-



## Udemy講座

【AutoML】自動化された機械学習を学ぼう！  
【PyCaret / Google Colab / Kaggle】



受講いただき

ありがとうございました！

