

実践 Kaggle Titanic 生存者予測

目次

1. はじめに
2. python の基本
3. Kaggle の使い方
4. 機械学習の流れ

1. はじめに

1-1. 本日の目標

- 機械学習の流れを体験する

1-2. 本日の予定

- Python の基本
- Kaggle の使い方、用語説明
- 分類問題（Titanic）
- 機械学習の流れを確認

python の基本

Kaggle 上で利用できる Jupyter notebook を利用して説明します。

<https://www.kaggle.com/tatsuyateshima/kaggle-titanic>

もっと知りたい場合はKaggle がサポートしてくれます。

<https://www.kaggle.com/learn>

機械学習で主に使用するライブラリ

- NumPy - 数値計算を効率的に行うための拡張モジュール。
- Pandas - データ解析を支援する機能を提供するライブラリ。
- matplotlib - グラフ描画のライブラリ。Pandas オブジェクトのグラフ描画メソッドで利用している。
- Seaborn - matplotlib のラッパー。matplotlib をそのまま使うと面倒なグラフが簡単に描ける。
- scikit-learn - オープンソース機械学習ライブラリ。

Kaggle とは(1/2)

- 機械学習を実際にしようにも題材が無い。
- 環境構築が大変そう。
- そもそも何から手をつけたら良いのか分からない。

そこで **Kaggle** です。

<https://www.kaggle.com/>

- 世界中のデータサイエンス・機械学習に携わる人が参加するコミュニティーサイト。
- サイト上にコーディング環境があり、Python 言語でのデータ分析および機械学習が可能。

Kaggle とは(2/2)

- Google アカウントだけでログインできる。
- 環境構築も簡単 ← ブラウザ上で実行可能
- コンペ形式で様々な題材がある。
- 先人の知恵がたくさんある。
 - ex. Qiita での情報、Kaggle の他の提出者の解答も閲覧可能
- ただし、英語

過去のコンペに参加してみよう

タイタニック号の生存者予測

[Titanic: Machine Learning from Disaster](#)

機械学習の流れ

<https://www.kaggle.com/tatsuyateshima/kaggle-titanic>

終わり