

みんなのKaggle講座

Section3



Section3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. Section3の概要
2. 特徴量エンジニアリング
3. ハイパーパラメータの調整
4. 交差検証
5. 演習

教材の紹介

- **Pythonの基礎:**

python_basic

- **Section3の教材:**

01_feature_engineering.ipynb

02_hyperparameter.ipynb

03_cross_validation.ipynb

04_exercise.ipynb

演習の解答 Section2

section_2/03_exercise.ipynb

特徴量エンジニアリングとは？

- 特徴量エンジニアリング (Feature Engineering)
 - データを機械学習で扱える形に変換
 - 値の変換、欠損値の穴埋め、不要なデータの削除、etc...
 - 精度向上に有用な、新たな特徴量を作成

特徴量エンジニアリング

- 01_feature_engineering.ipynb

ハイパーパラメータの調整



ハイパーパラメータの調整

- 02_hyperparameter.ipynb

交差検証

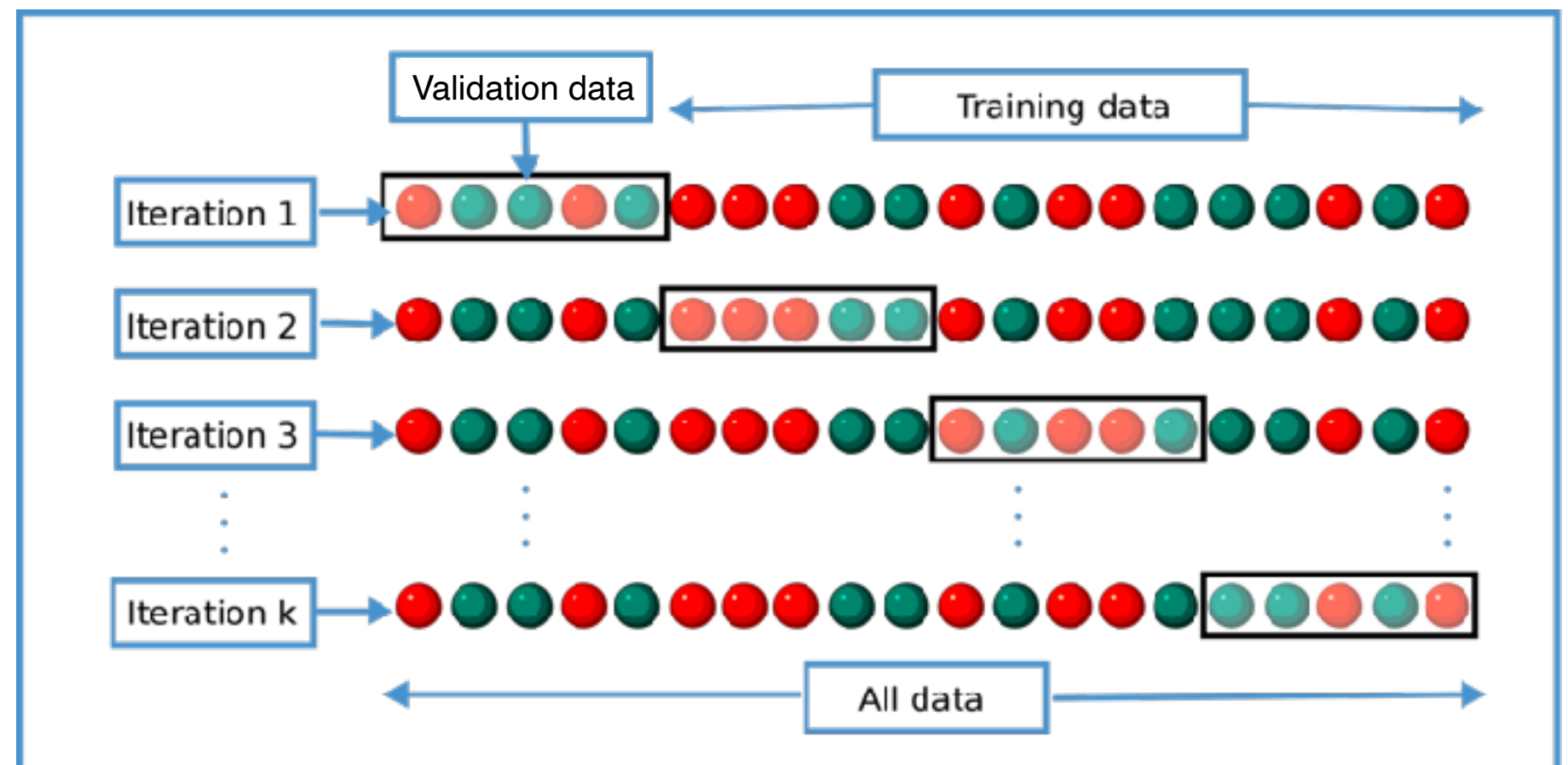


交差検証とは？

- 交差検証 (cross-validation)

→ データを分割して、そのうち1つをテストデータに、残りを学習データに使用してモデルを評価

→ これを全ての組み合わせで行い平均を取る



画像の引用: Wikipedia Cross-validation (statistics) [https://en.wikipedia.org/wiki/Cross-validation_\(statistics\)](https://en.wikipedia.org/wiki/Cross-validation_(statistics))

Gufosowa - Own work CC BY-SA 4.0

交差検証

- 03_cross_validation.ipynb

澳白

演習


- 04_exercise.ipynb

次回の内容

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**