機械学習手法を用いたサッカーの試合結果予測モデルの構築とその活用

日本大学文理学部　情報科学科

谷聖一研究室　田中聡太朗　細井祐希

画像：https://unsplash.com/photos/8ZvTRSRdsOA

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

1.はじめに

目次

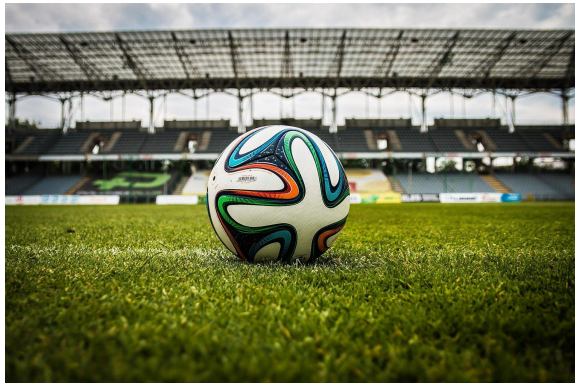
1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

1. はじめに

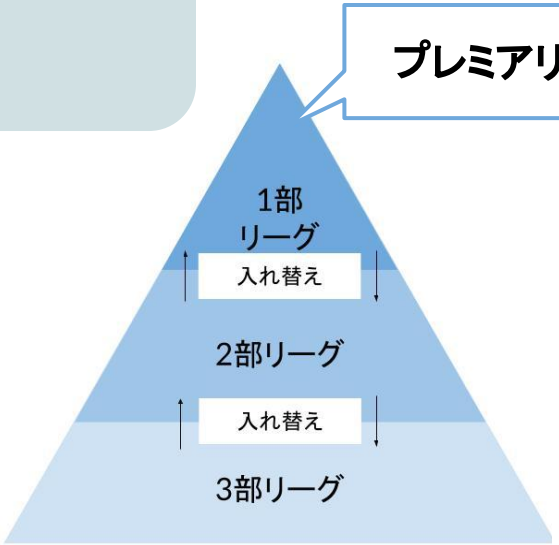
1.1. 動機

サッカー観戦が好き

特に**プレミアリーグ**を観ている



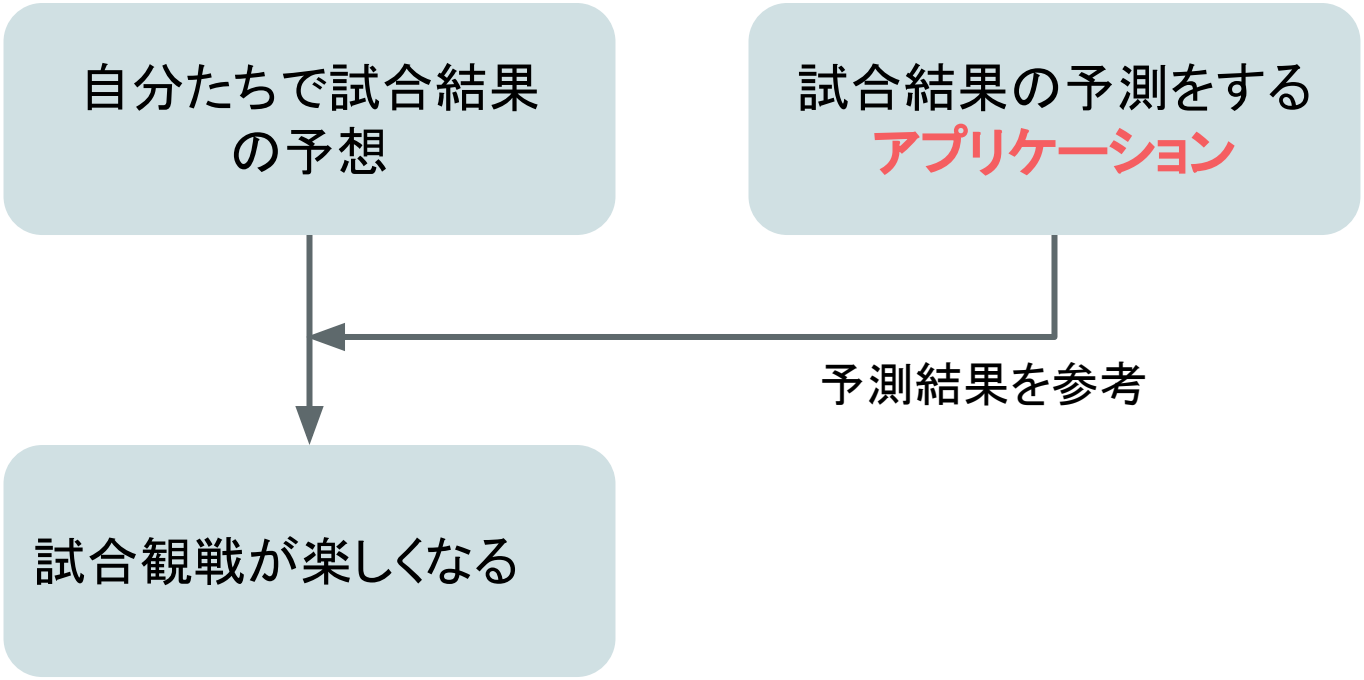
引用:https://pixabay.com/ja/photos/%E3%82%B5%E3%83%83%E3%82%AB%E3%83%BC-%E3%83%9C%E3%83%BC%E3%83%AB-%E3%82%B9%E3%82%BF%E3%82%B8%E3%82%A2%E3%83%A0-488700/



引用: A cross-sectional survey of asymptomatic meningococcal carriage amongst UK adolescents at a period of low invasive meningococcal disease incidence

1. はじめに

1.1. 動機



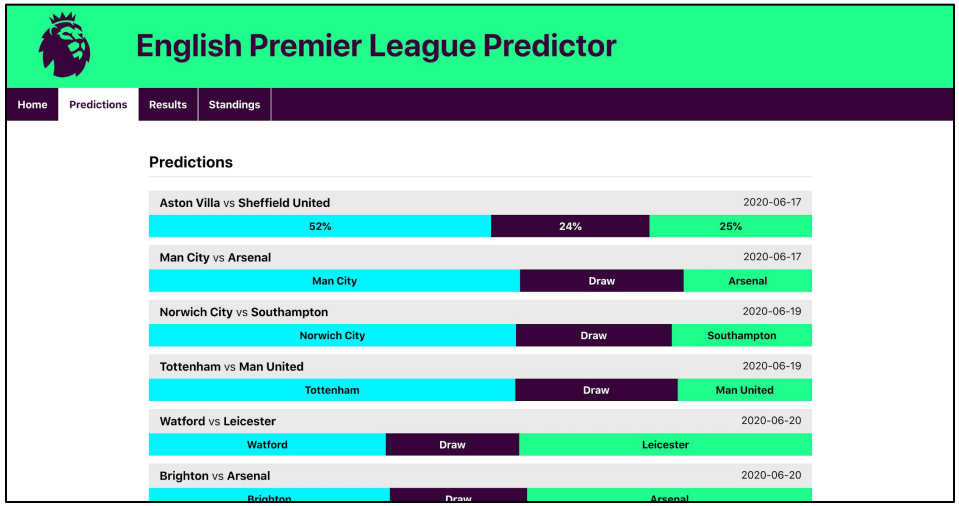
目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

1. はじめに

1.2. 先行研究の紹介

プレミアリーグの試合結果を予測するアプリケーション



2020年6月18日以降予測結果が未更新

MIT License
Copyright (c) 2018 Richard Woongbin Choi

github : https://github.com/woongbinchoi/English-Premier-League-Prediction

demo app : https://epl-client.herokuapp.com/

1. はじめに

1.2. 先行研究の紹介

機械学習を用いたサッカーの試合結果の予測(2019年度)

明治大学 理工学部 情報科学科

野村空生氏

概要

サッカー**J1リーグの試合結果の予測**をする学習モデルを作成

使用していた機械学習手法

- 勾配ブースティング決定木
- サポートベクターマシン
- ニューラルネットワーク etc...

参照:
http://www.th.cs.meiji.ac.jp/assets/researches/2019/nomurasorai/research_nomurasorai.pdf

1. はじめに

1.2. 先行研究の紹介

Predictive analysis and modelling football results using machine learning approach for English Premier League(April–June 2019)

Department of Computer Science and Engineering, School of Engineering Sciences and Technology , Guru Gobind Singh Indraprastha University

Rahul Baboota , Harleen Kaur

概要

サッカー**プレミアリーグの試合結果の予測**をする学習モデルを作成

使用していた機械学習手法

- 勾配ブースティング決定木
- サポートベクターマシン
- ランダムフォレスト etc...

参照: https://www.sciencedirect.com/science/article/abs/pii/S0169207018300116

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

1. はじめに

1.3.演習目的・内容

既存のアプリケーション
機能していない

先行研究
誰でも手軽に予測できない

→ 誰でも手軽に予測できる
アプリケーションを作成

プレミアリーグの試合結果予測モデルを機械学習手法を用いて構築し、アプリケーションの作成・活用を行う

2.準備

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

2.準備

2.1.機械学習

機械学習とは

明示的にプログラムしなくても学習する能力をコンピュータに与える研究分野

(Arthur Lee Samuel 1959)

Arthur Lee Samuel

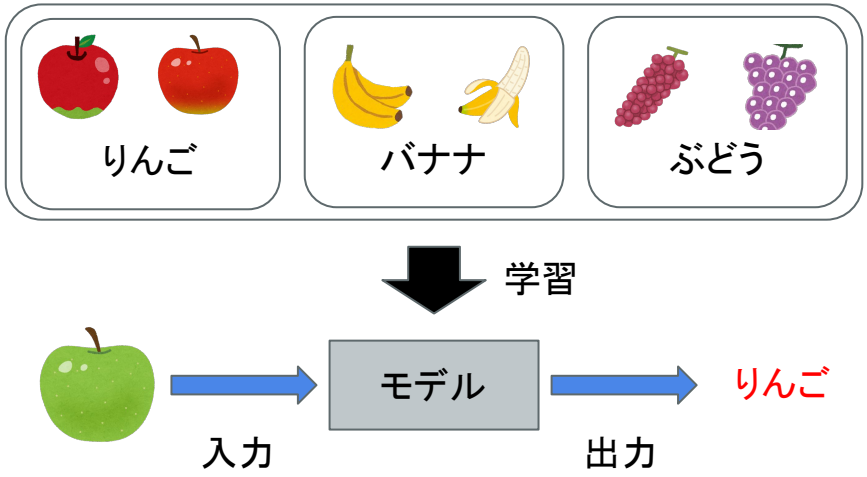
アメリカの計算機科学者で、世界初の学習型プログラム であり、人工知能(AI)の基本概念をいち早く示した人物。

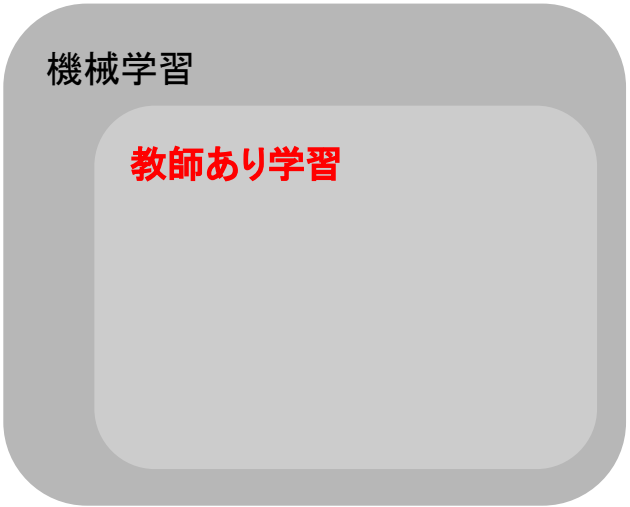
2.準備

2.1機械学習

教師あり学習とは

学習データに正解を与えた状態で学習させる機械学習手法



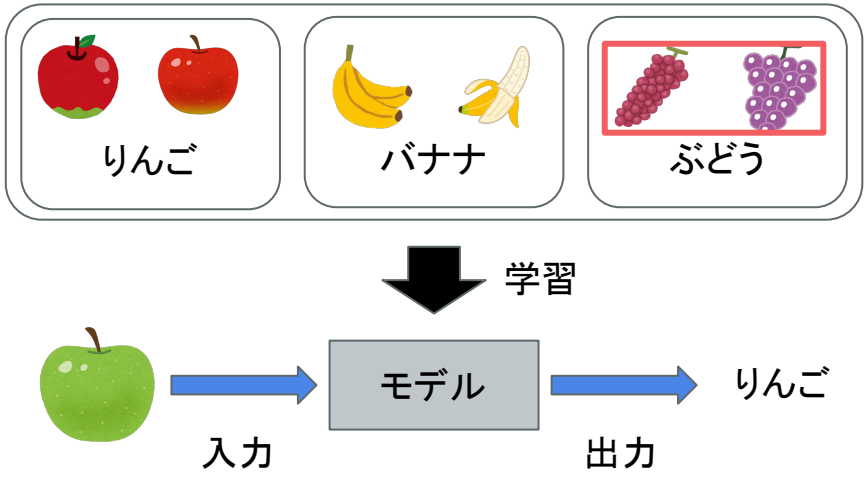


2.準備

2.1機械学習

教師あり学習とは

学習データに正解を与えた状態で学習させる機械学習手法



- 特徴量

対象の特徴が数値化されたもの

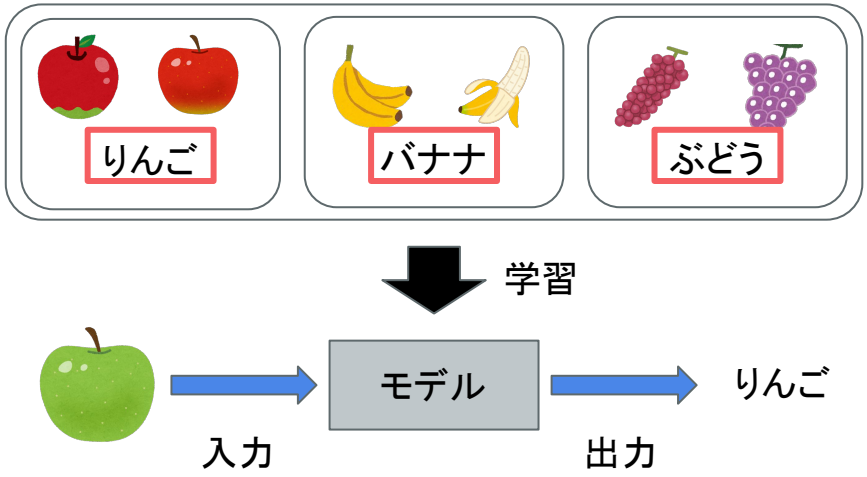
ex)RGB値,高さ,幅

2.準備

2.1機械学習

教師あり学習とは

学習データに正解を与えた状態で学習させる機械学習手法



- 正解ラベル

あるデータがどのカテゴリに属するかを表した数値

ex)りんご→0，バナナ→1，ぶどう→2

2.準備

2.1.機械学習

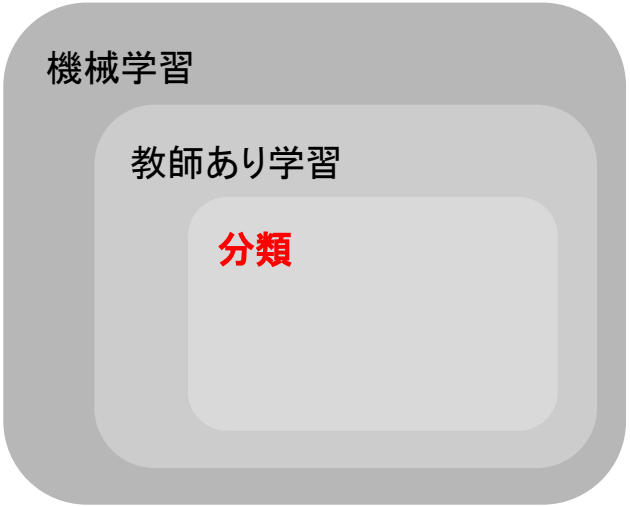
分類とは

分析したいデータが属するカテゴリーやクラス,

種類が何なのかを判定する教師あり学習手法

代表的手法

- ランダムフォレスト
- サポートベクターマシン
- ニューラルネットワーク
- 勾配ブースティング決定木　etc...



目次

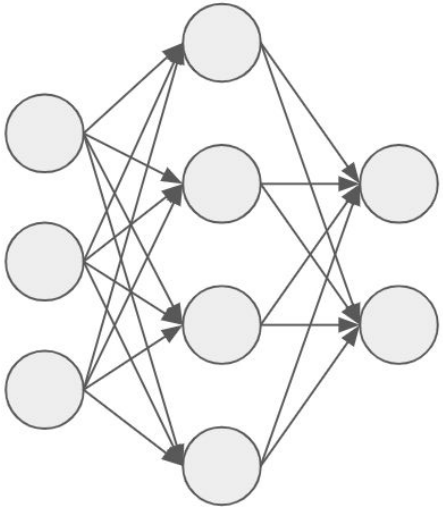
1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

2.準備

2.1.ニューラルネットワーク

ニューラルネットワークとは

脳の神経回路・情報処理機構を模した数理モデル

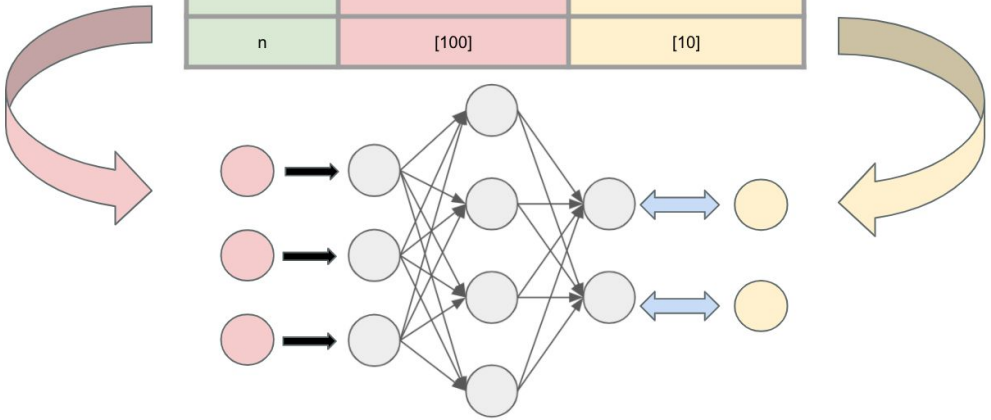


2.準備

2.1.ニューラルネットワーク

ニューラルネットワークの学習

データ番号	データ（特徴量）	データ（正解ラベル）
1	[101]	[10]
2	[110]	[01]
⋮	⋮	⋮
n	[100]	[10]



出力と正解ラベルの誤差を計算し，出力が正解ラベルに近づくように重みw_iを調整

→ には重みw_iがつく

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

2.準備

2.1.勾配ブースティング決定木

勾配ブースティング決定木とは

- 勾配降下法
- アンサンブル学習
- 決定木

の3つの手法が組み合わされた機械学習の手法

2.準備

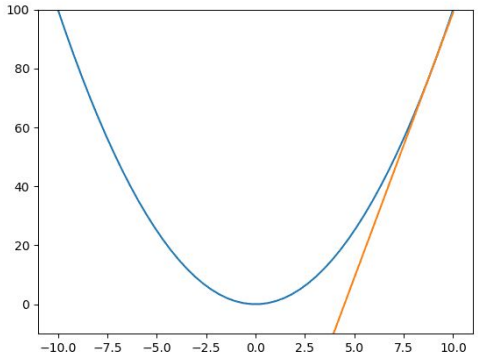
2.1.勾配ブースティング決定木

勾配降下法とは

機械学習を用いた予測モデルの学習の目的は

実際値と予測値の誤差を小さくすること

誤差関数：実際値と予測値のズレの大きさを表す関数



Qiita 勾配降下法をアニメーション化した

https://qiita.com/ophhdn/items/1640f16d161b3dde67c5

最小化したい誤差関数がある場合にその関数の勾配を求めその勾配を元に関数の最小値を探索する手法

2.準備

2.1.勾配ブースティング決定木

アンサンブル学習とは

複数の弱学習器を用いてそれらを相互に作用させることで

大きな一つの学習モデルとしてモデルの精度を向上させる手法

2.準備

2.1.勾配ブースティング決定木

決定木とは

木構造を用いて分類や回帰を行う機械学習の手法の一つ

2.準備

2.1.勾配ブースティング決定木

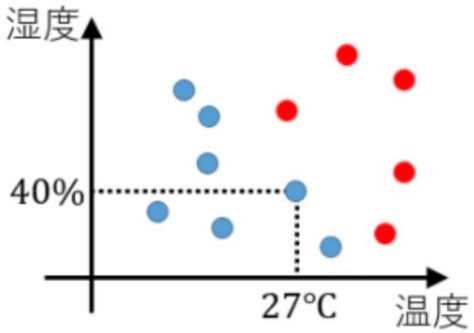
決定木(分類木)とは

1. 日々の温度と湿度のデータ（特徴量）
2. その日Aさんが暑いと感じたか暑くないと感じたかかのデータ（正解ラベル）

が与えられた状況を考えてみる

赤い点が暑いと感じた日

青い点が暑くないと感じた日



2.準備

2.1.勾配ブースティング決定木

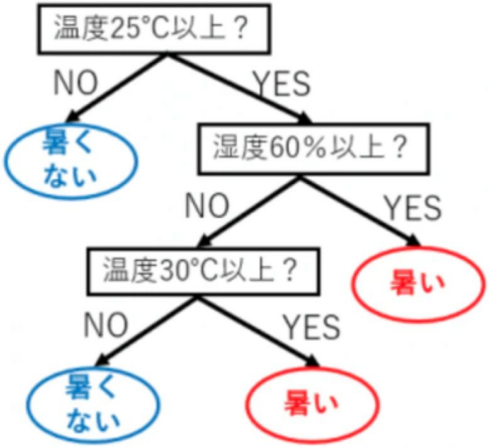
決定木(分類木)とは

このデータから

「温度と湿度がどのようなときにどう感じるのか？」

といったことを木で表現できる

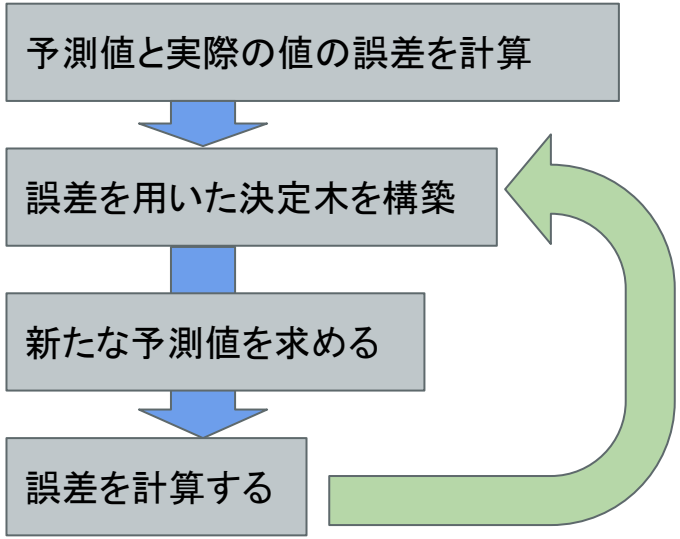
この木のことを分類木と言う



2.準備

2.1.勾配ブースティング決定木

勾配ブースティング決定木とは



左の図のサイクルを繰り返すことで
誤差を小さくしていく

つまり精度を改善していく機械学習
アルゴリズム

3.予備実験

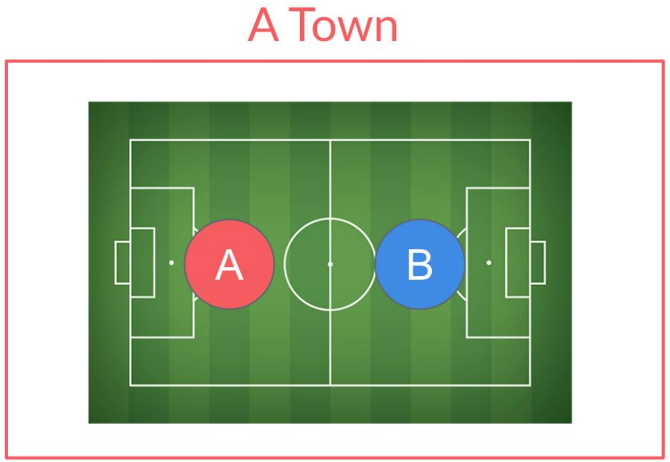
目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

3.予備実験

3.1.予備実験の準備

学習モデル構築の目的



チームA: ホームチーム

チームB: アウェイチーム

3.予備実験

3.1.予備実験の準備

学習モデル構築の目的

3.予備実験
3.1.予備実験の準備

学習データ

特徴量

- チームの能力値の差
- ホームorアウェイ
- イロレーティングの差
- 主審

チームの強さを示す指標

先行研究で使用

プレミアリーグの試合結果を予測する先行研究
Predictive analysis and modelling football results using machine learning approach for English Premier League(April–June 2019)

3.予備実験

3.1.予備実験の準備

学習データ

特徴量

- チームの能力値の差
- ホームorアウェイ
- イロレーティングの差
- 主審

- 慣れた芝生
- 観客の数
- 移動の負担の少なさ

ホームが有利

3.予備実験

3.1.予備実験の準備

学習データ

特徴量

- チームの能力値の差
- ホームorアウェイ
- イロレーティングの差
- 主審

チームの調子の良さを示す指標

先行研究で使用

J1の試合結果を予測する先行研究
機械学習を用いたサッカーの試合結果の予測(2019年度)

3.予備実験
3.1.予備実験の準備

学習データ

特徴量

- チームの能力値の差
- ホームorアウェイ
- イロレーティングの差
- 主審

チームによって主審との
相性の良さがある

主審(29人)をリスト化しindexを使用

参照:https://plaza.rakuten.co.jp/reds2004v/diary/202008120000/

3.予備実験

3.1.予備実験の準備

学習データ

特徴量

- チームの能力値の差
- ホームorアウェイ
- イロレーティングの差
- 主審

- チーム能力値データ
  FIFAIndexからスクレイピング

トッテナム	85	83	82	83
マンチェスター・U	85	83	81	83
チェルシー	84	86	82	83
アーセナル	84	80	81	82

- 過去の試合結果データ
  SkySportsからスクレイピング

スウォンジー	0	4	マンチェスター・U	Jonathan Moss
ボーンマス	0	2	ワトフォード	Roger East
バーンリー	0	1	WBA	Martin Atkinson
レスター・シティ	2	0	ブライトン	Lee Probert

FIFAIndex:https://www.fifaindex.com/ja/players/top/fifa21/
SkySports:https://www.skysports.com/

3.予備実験

3.1.予備実験の準備

学習データ

特徴量

- チームの能力値の差
- ホームorアウェイ
- イロレーティングの差
- 主審

2017～2020のデータを使用

- チーム能力値データ
  FIFAIndexからスクレイピング

トッテナム	85	83	82	83
マンチェスター・U	85	83	81	83
チェルシー	84	86	82	83
アーセナル	84	80	81	82

過去の試合結果データ
rtsからスクレイピング

スワ	0	4	マンチェスター・U	Jonathan Moss
ボーンマス	0	2	ワトフォード	Roger East
バーンリー	0	1	WBA	Martin Atkinson
レスター・シティ	2	0	ブライトン	Lee Probert

FIFAIndex:https://www.fifaindex.com/ja/players/top/fifa21/
SkySports:https://www.skysports.com/

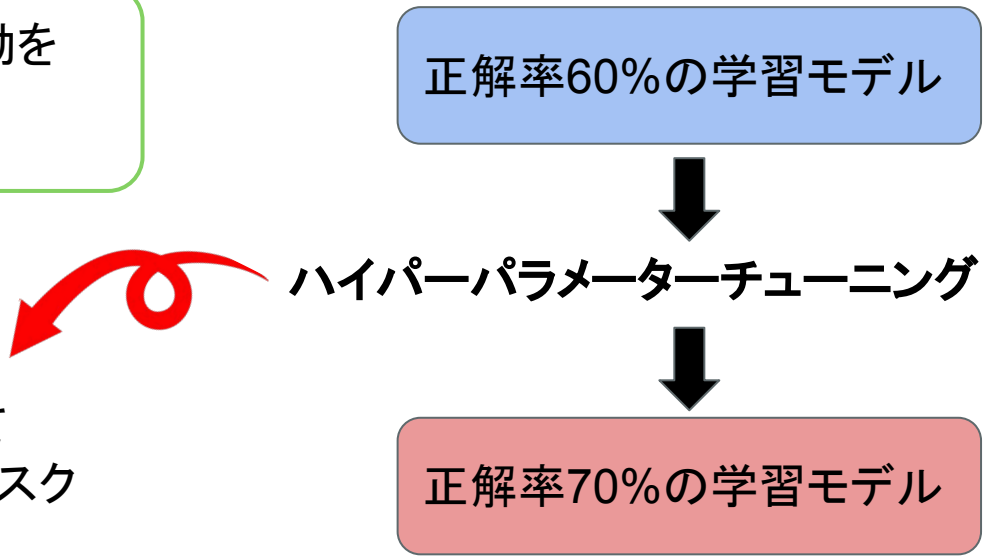
3.予備実験

3.1.予備実験の準備

ハイパーパラメーター

機械学習アルゴリズムの挙動を

設定するパラメータ



機械学習の実装において
避けて通れない必須のタスク

3.予備実験

3.1.予備実験の準備

チューニング手法

グリッドサーチ
- **手動**で行う手法
- 本演習では10000通り

optuna
- **自動**で行う手法

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

3.予備実験

3.2.予備実験内容①

3値分類

予備実験として3値分類を行った

↓

引き分けの予測精度が良くない

プレミアリーグの試合結果を予測する先行研究
Predictive analysis and modelling football results using machine learning approach for English Premier League(April–June 2019)

先行研究でも3値分類を行っており
引き分けの予測精度が良くない

3.予備実験

3.2.予備実験内容①

2値分類から3値分類

引き分けをうまく予測したい

ホームチーム
予測：勝ちか引き分けor負け

アウェイチーム
予測：勝ちか引き分けor負け

	ホームの予測	アウェイの予測	試合全体の予測
パターン1	勝ちか引き分け	負け	ホームの勝ち
パターン2	負け	勝ちか引き分け	アウェイの勝ち
パターン3	勝ちか引き分け	勝ちか引き分け	引き分け
パターン4	負け	負け	引き分け

本演習ではこの手法を用いることに決定

3.予備実験

3.2.予備実験内容②

機械学習手法とチューニング手法の選択

NN　:ニューラルネットワーク
GBDT:勾配ブースティング決定木

	精度
NN ＋ グリッドサーチ	0.69
GBDT ＋ グリッドサーチ	0.71
GBDT ＋ optuna	0.73

- 2つの機械学習手法
- 各チューニング手法

組み合わせて2値分類のモデルの精度を検証

4.演習

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

4.演習

4.1.演習の準備

特徴量と正解ラベル

特徴量

- ATTの差
- MFの差
- DEFの差
- OVRの差
- ホームorアウェイ：1or0
- イロレーティングの差
- 主審のindex

正解ラベル

- 勝ち引き分けor負け：1or0

4.演習

4.1.演習の準備

特徴量と正解ラベル作成

ホーム

特徴量　[-11,-7,-9,-9,1,-16,0]

正解ラベル　0

アウェイ

特徴量　[11,7,9,9,0,16,0]

正解ラベル　1

サウサンプトン	76	76	76	76

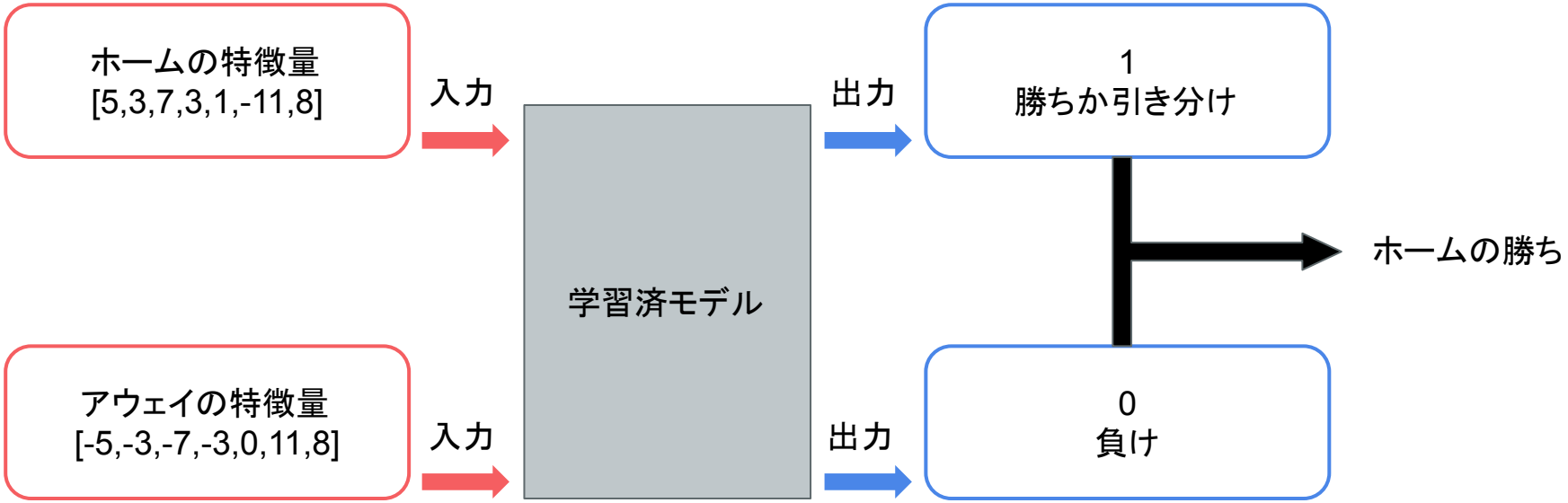
リバプール	87	83	85	85

サウサンプトン	1	2	リバプール	1492	1508	Andre Marriner

4.演習

4.1.演習の準備

入力と出力の例



4.演習

4.1.演習の準備

学習モデルの評価手法

クロスバリデーションを用いたmacro平均

macro平均

- F値

目的のクラスにどれだけ良く分類されているかを示す指標

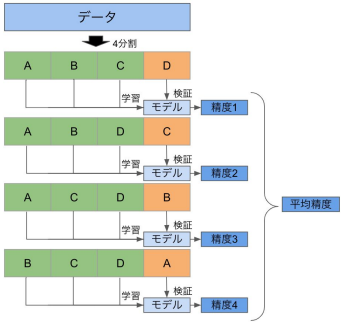
- macro平均

各クラスごとのF値の平均
モデルがどれだけ優れているかを示す指標

クロスバリデーション

データを複数の
グループに分割

学習用データと
検証用データを
入れ替えながら精度を検証



4.演習

4.1.演習の準備

予備実験から採用したチューニング手法

optunaを使用

ハイパーパラメータの最適化を自動化するための
ソフトウェアフレームワーク

ハイパーパラメータの値に関する試行錯誤を自動的に行いながら
優れた性能を発揮する値を自動的に発見する

4.演習

4.1.演習の準備

予備実験から採用した機械学習手法

勾配ブースティング決定木を使用

LightGBMとは

Microsoftによって開発された機械学習用の無料のオープンソース分散型勾配ブースティングフレームワーク

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

4.演習

4.2.演習内容

精度の向上

予備実験

- プレミアリーグ2017～2020のデータ
- 勾配ブースティング決定木
- optuna

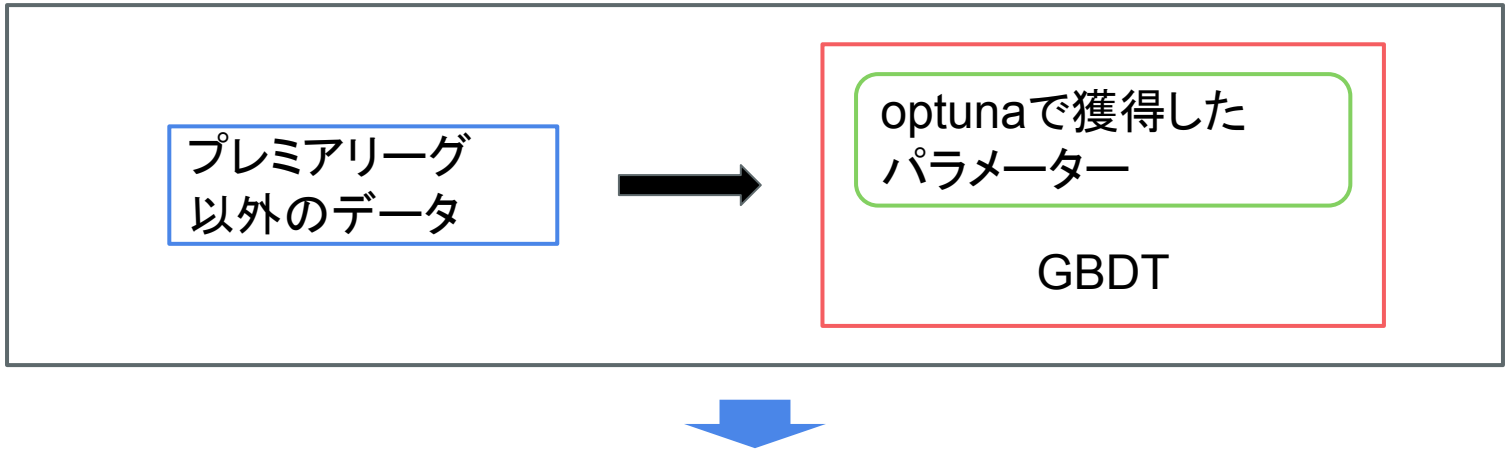
プレミアリーグ以外の
データを使用

更なる精度の向上を狙う

4.演習

4.2.演習内容

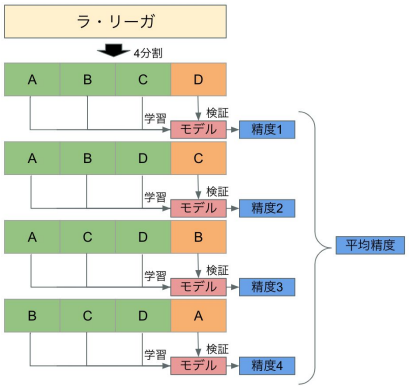
プレミアリーグと似た性質を持つデータを探す



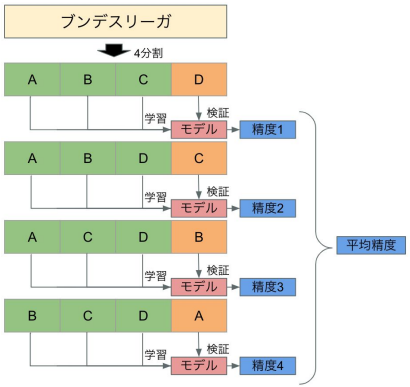
4.演習

4.2.演習内容

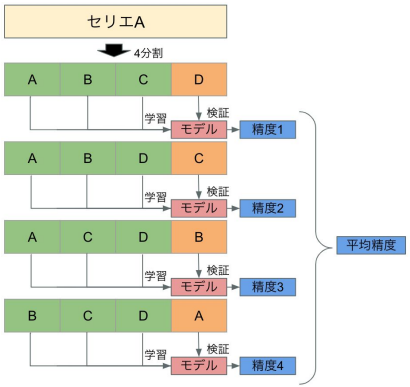
プレミアリーグと似た性質を持つデータを探す



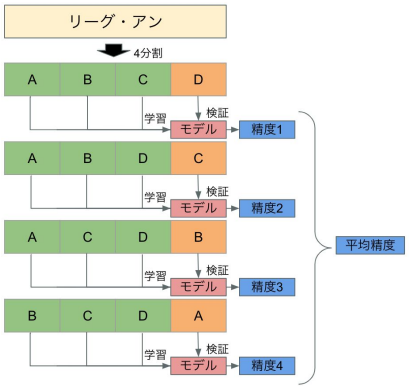
0.59



0.61



0.68



0.61

4.演習

4.2.演習内容

セリエAのデータを足したデータで学習

	macro平均
ラ・リーガ	0.59
ブンデスリーガ	0.61
セリエA	0.68
リーグ・アン	0.61

プレミアリーグ
＋
セリエA

のデータで再学習

	F値
勝ちか引き分け	0.81
負け	0.69
macro平均	0.75

4.演習

4.2.演習内容

精度の比較とアプリケーション化モデルの決定

	2値分類macro平均	3値分類macro平均
NN ＋ グリッドサーチ	0.69	0.48
GBDT（プレミアリーグのみ）＋ optuna	0.73	0.50
GBDT（プレミアリーグ＋セリエA）＋ optuna	0.75	0.55

5.予測モデルを用いたアプリケーション

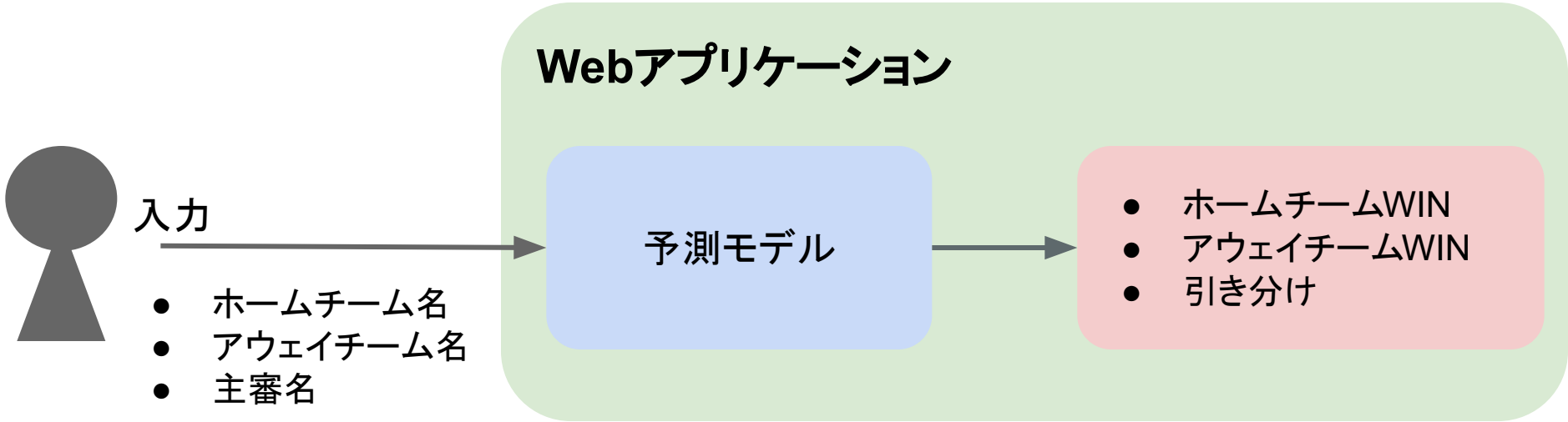
目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

5. 予測モデルを用いたアプリケーション

5.1. Webアプリケーションの作成

アプリケーション内容



5. 予測モデルを用いたアプリケーション

5.1. Webアプリケーションの作成

予測モデルの保存方法

pickle形式で予測モデルを**保存**

pickle

Pythonの標準ライブラリにあるpickleモジュール
Pythonのオブジェクトを直列化・非直列化するための機能
直列化とは、プログラミング言語においてオブジェクトをビット列に変換する事を指す

参照：https://blog.amedama.jp/entry/2015/12/05/132520

5. 予測モデルを用いたアプリケーション

5.1. Webアプリケーションの作成

作成方法

Flaskを利用しWebアプリケーションを**作成**

Flask

PythonのWebアプリケーションフレームワークで、軽量で実装できる機能が必要最低限に収まっている

5. 予測モデルを用いたアプリケーション

5.1.Webアプリケーションの作成

公開方法

Herokuを利用しWebアプリケーションの**公開**

Heroku

Herokuとは、いわゆるPaaS、Webアプリケーションの構築から保守、スケールの拡張などの運用作業までができるクラウドサービスで、Salesforce社の製品の1つ

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 学習モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

5. 学習モデルを用いたアプリケーション

5.2.Webアプリケーションデモ

Premier League Predictor 2021

URL : https://testml2021.herokuapp.com

6.アプリケーションの活用

目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

6.アプリケーションの活用

6.1. 20/21シーズン1月の試合結果予測

試合結果予測の結果

正解率　　:58%
macro平均　:0.57
引き分けのF値:0.46

6.アプリケーションの活用

6.1. 20/21シーズン1月の試合結果予測

既存のアプリケーションとの比較

本演習

訓練データ：2017～2020/12

テストデータ：2021/1

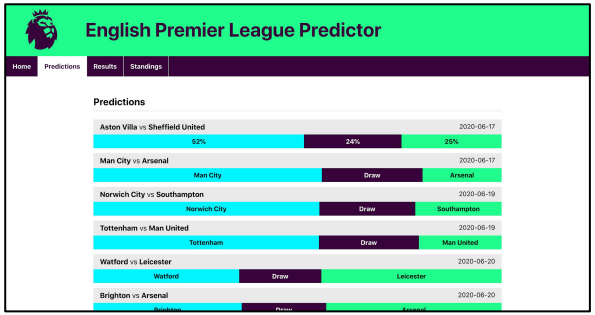
⟷

既存のアプリケーション

訓練データ：？

テストデータ：？

	本演習	既存のアプリケーション
正解率	58%	54.37%



6.アプリケーションの活用

6.1. 20/21シーズン1月の試合結果予測

先行研究との比較

本演習

訓練データ：2017～2020/12

テストデータ：2021/1

既存のアプリケーション

訓練データ：2005～2014

テストデータ：2015～2016

	本演習	先行研究
引き分けのF値	0.46	0.29
macro平均	0.57	0.476

プレミアリーグの試合結果を予測する先行研究
Predictive analysis and modelling football results using machine learning approach for English Premier League(April–June 2019)

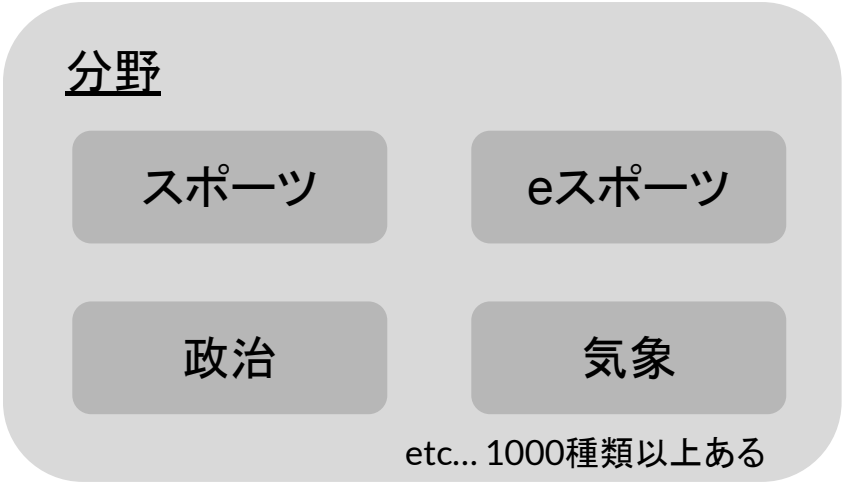
目次

1. はじめに
   - 1.1. 動機
   - 1.2. 先行研究の紹介
   - 1.3. 演習目的・内容
2. 準備
   - 2.1. 機械学習
   - 2.2. ニューラルネットワーク
   - 2.3. 勾配ブースティング決定木
3. 予備実験
   - 3.1. 予備実験の準備
   - 3.2. 予備実験内容
4. 演習
   - 4.1. 演習の準備
   - 4.2. 演習内容
5. 予測モデルを用いたアプリケーション
   - 5.1. Webアプリケーションの作成
   - 5.2. Webアプリケーションデモ
6. アプリケーションの活用
   - 6.1. 20/21シーズン1月の試合結果の予測
   - 6.2. ブックメーカーと勝負
7. まとめ

6.アプリケーションの活用
6.2.ブックメーカーと勝負

ブックメーカーとは

欧米における「**賭け屋**」



6.アプリケーションの活用
6.2.ブックメーカーと勝負

BET CHANNELとは

プレミアリーグの各試合結果にベットすることで

オッズに応じた払戻金を得ることができる

BET CHANNEL

オッズ

オッズの初期値は
BET CHANNELが決定

ユーザーのベットに応じて変動

6.アプリケーションの活用
6.2.ブックメーカーと勝負

<u>損益をどのように出すか</u>

本演習で作成したアプリケーションから得られた1月の全試合結果に
1000円ずつベットすることでどれほどの損益が出るか試してみた

例）2/4トッテナムvsチェルシー

予測	損益(円)
トッテナムWIN	-1000
引き分け	-1000
チェルシーWIN	+850

6.アプリケーションの活用
6.2.ブックメーカーと勝負

毎試合1000円Betした時どれほどの損益をあげられるか

結果は「-3190円」で3190円の損

49試合
総投資：49000円

28試合的中
回収：45810円

6.アプリケーションの活用
6.2.ブックメーカーと勝負

毎試合1000円Betした時どれほどの損益をあげられるか

予測が難しい試合

オッズが高い

予測が難しい試合結果を的中できなかったため払戻金が少なくなってしまった

7.まとめ

7.まとめ

演習内容

- プレミアリーグ試合結果予測モデルを機械学習手法を用いて構築
- プレミアリーグ試合結果予測モデルを用いてアプリケーションを作成
- アプリケーションを使って1月の試合結果の予測

今後の課題

- イロレーティングの出力結果への影響力の向上
- 試合結果だけでなく結果と得点の予測
- リアルタイムで学習モデルを更新する仕組み