

輪読会(ゼロつく)成果物
競馬予想AI_阪神競馬場ver.

...

yazu yuichiro

Goal ゼロつくの内容を自分のものにする(扱える)

❑ 苦労

- ❑ 意外とコピペではいけないことが多い
- ❑ いざ扱ってみても意外と細かいソースコードの理解ができておらず、エラーが出るとclassとかややこしくて修復に時間がかかる
- ❑ ゼロつくのTwoLayerNetは多クラス分類用なので二値分類用にカスタムしないといけない

❑ 成長

- ❑ シンプルにpythonのclass記法についての理解が深まった
- ❑ ネットワークの構造の理解が深まった

Target 阪神競馬場の過去データで競馬予想AI(DL)を実装

- ❑ 仮説: 競馬場によって特徴があるので、1つのコースに絞って学習したら精度が高いAIができるのではないかな。
- ❑ 方法: 3着以内を正例、4着以降を負例として予測
 - ❑ 複勝での的中を狙う想定
- ❑ 参考: 競馬予想で始めるデータ分析・機械学習
 - ❑ リンク→https://youtube.com/playlist?list=PL_fbLtvVldtnv6q99s3CPLBTQXI53XPNL



Why なぜ競馬？→スポーツとして魅了されている

□ 幼少期

- 父親の影響で〇〇ライダーや〇〇マンよりジョッキーごっこで遊ぶ
- 誕生日プレゼントに競馬用ムチ(レプリカ)をもらう
- 馬名をどんどん覚えひらがなよりカタカナが得意になる

□ 学生時代

- スポーツとしての競馬に魅了されジョッキーを目指す
- 想定よりも身長が伸びて断念
- 数々の競馬ゲームを嗜む→血統やレースの詳細などより詳しくなる
- 持ち前の知識を活かして馬券に手を出す
- 友達に馬券を教えて言われ競馬場に行くも未成年だったため補導

Task 分析の流れ・工数

□ スクレイピング(40時間※放置含む)

- 阪神競馬場の過去10年間のレース結果、そのレースに出走していた競走馬の過去成績
- 参考から阪神競馬場のデータだけ抽出できるようカスタムし実行

□ データクレンジング(3時間)

- スクレイピングしたデータを分析できるようクレンジング(ほぼ参考からコピー)

□ データ分析(20時間)

- TwoLayerNetを二値分類用にカスタム

Task TwoLayerNetモデルの構築

- 復習
- 二値分類対応へのカスタム
 - SigmoidWithLoss
 - cross_entropy_error
 - TwoLayerNet . accuracy

Foward 自動化、運用・WEBアプリ化

□ 自動化、運用

- 回収率の計算
- 様々な賭け方の追加
- レース開始前の出馬データをスクレイピングし予測
- 成績データのスクレイピング、情報更新

□ WEBアプリ化

- Seleniumなどを勉強しWEBアプリ化
- ポートフォリオとして公開