



チーム：ウマ天気（総会版）

2018年10月01日

<メンバー（2017年当時）>

二見 加代子

根岸 颯

赤池 奈奈

大石 直人

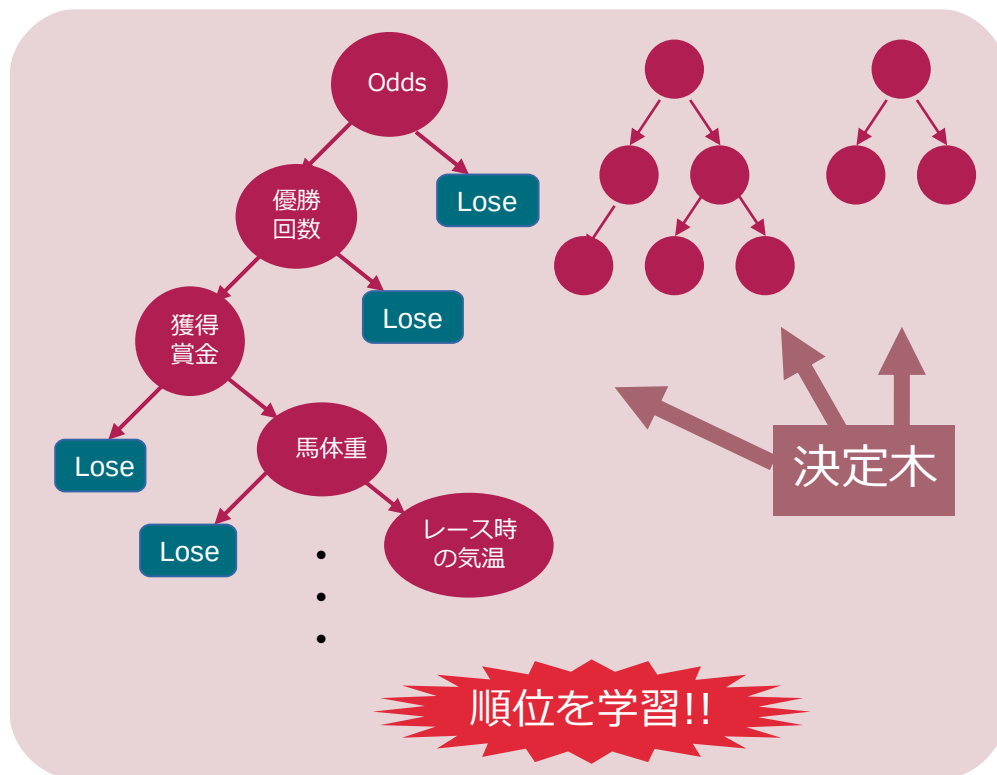
吉田 光輝

発表者：舟城 亮一

そもそもどうやって予想しているのか？



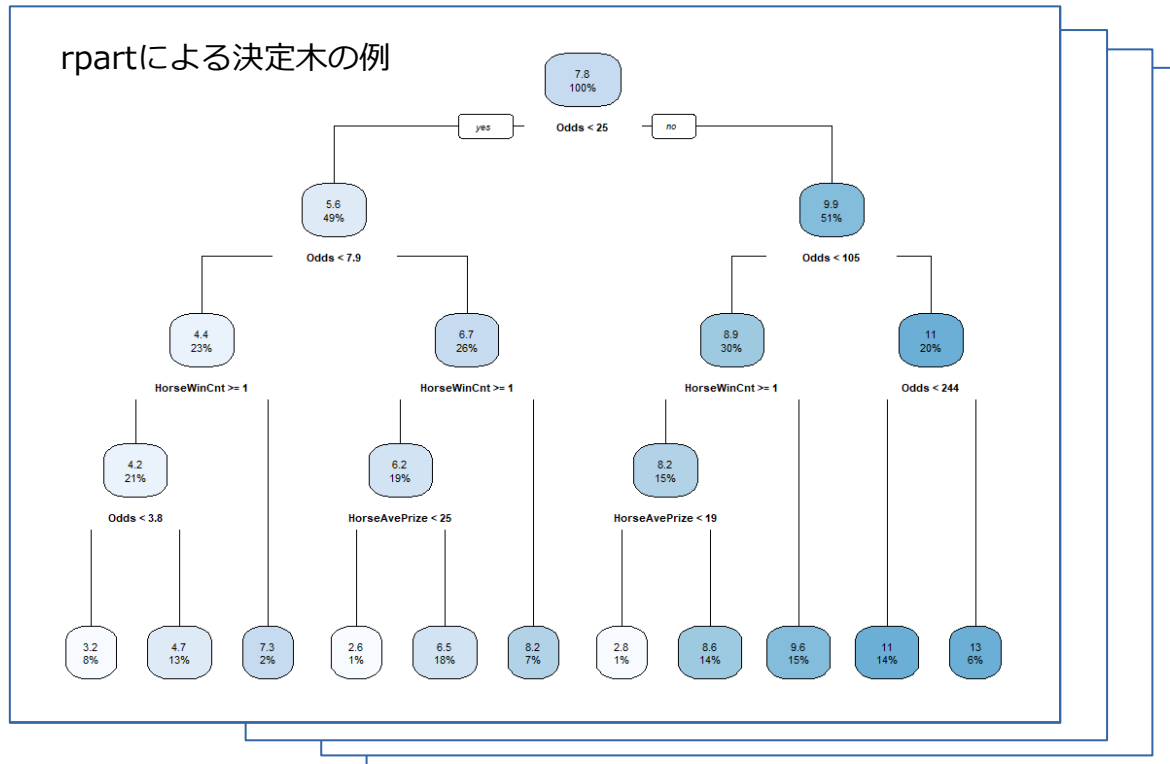
レース前日までに作っておく！



学習データ (RandomForestによる決定木)

決定木による分析とは？

- 学習データから、予想したい目的変数(順位)に強く影響する要因を見つけ、目的変数を導くルールを木構造で示したもの
- Random Forestは、学習データからランダムにデータを選び複数の決定木を作成して、総合的に目的変数を予想



今回のrandomForestでは1000の決定木を使用

予想実行！（デモ）

インターネット

- ・今日のレース情報
- ・レース時の天気

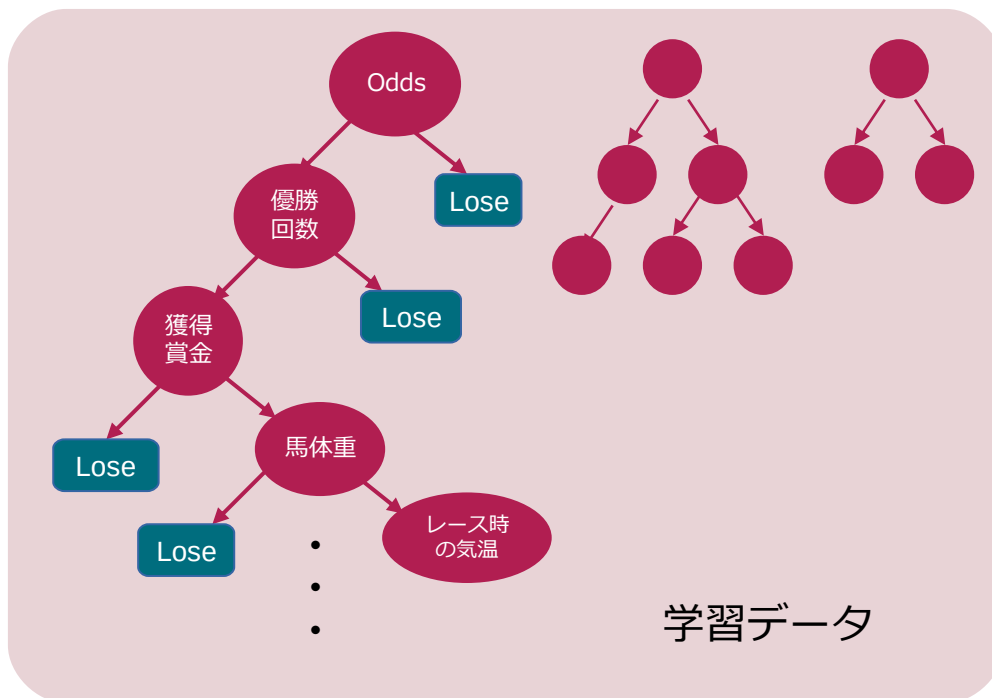


レース開始時点での
成績の計算！



予測

predict()



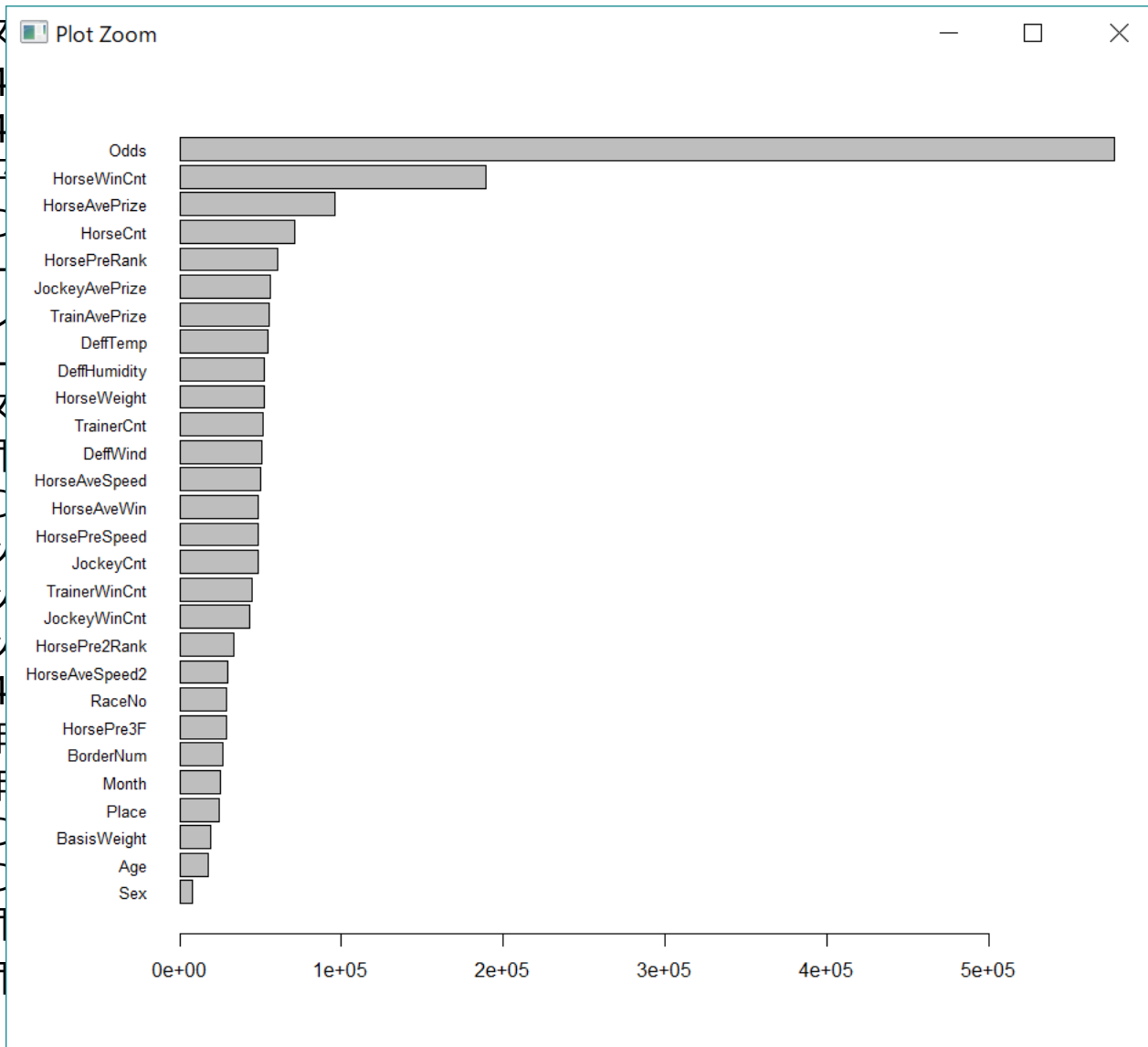
順位を予想！

（最も小さな値を1位と想定）

デモ

学習で使っているデータ

- レース
- 過去4
- 過去4
- 馬の平
- 騎手の
- 前レー
- 前々レ
- 前レー
- レース
- 調教師
- 前回の
- 3位以
- 3位以
- 3位以
- 過去4
- 馬の年
- 馬の年
- 騎手の
- 騎手の
- 調教師
- 調教師



(ちょっと前の) 成績

2017/12/24 有馬記念 (GI) 「キタサンブラック」的中！

2018/5/27 日本ダービー (GI) 「ワグネリアン」的中！


舟城亮一
 2017年12月24日 · 2人

有馬記念！
機械学習が出した予想はキタサンブラック。
最近当ててないから自信ないなあ・・・

Console ~/Documents/AITCデータ分析/SQL/


```

> source("Nakayama.R")
[1] "Race No"
[1] 11
[1] "RaceName"
[1] "有馬記念 (GI)"
      
```

ヤマカツエース	キタサンブラック	クイーンズリング	プレスジャーニー	トーセンビクトリー	
6.713100	4.083300	9.445400	10.419067	10.013833	
サトノクロニクル	シャクトラ	レインボーライン	サクラアンブルール	シュヴァルグラン	
8.811933	7.728733	9.170600	9.027933	5.290967	
ルージュバック	サトノクラウン	ミッキークイーン	スワーヴリチャード	カレンミロティック	
9.356600	4.753433	6.348167	5.729767	10.519000	
サウンドオブアース					
10.206667					

```

[1] "Yoso"
      
```

キタサンブラック 2

他12人
 コメント3件

いいね！
 コメントする
 シェアする


舟城亮一 やった！！買っときやよかった！

いいね！ · 返信 · 37週間前


舟城亮一
 5月27日 · 2人

日本ダービー発走一時間前。
また、過去5年間のレース結果と気象情報から機械学習が予想した結果を書いておきます。
Macbookがはじき出した勝者は「ワグネリアン」。
当たるかな？

Console Terminal


```

~/Documents/AITCデータ分析/SQL/
> source("Tokyo.R")
[1] "Race No"
[1] 10
[1] "RaceName"
[1] "東京優駿 (GI)"
      
```

クラウンディバイダ	エアウインザー	コーカス	アップクォーク	セセリ	
6.055600	3.952133	6.353033	6.873033	9.245767	
ブライトエンブレム	プロフェット	ミッキークローリー	コスモナインボール	エンジン	
9.165767	7.861267	7.449733	8.468167	4.372533	

```

[1] "Yoso"
      
```

エアウインザー 7

```

> source("Tokyo.R")
[1] "Race No"
[1] 10
[1] "RaceName"
[1] "東京優駿 (GI)"
      
```

ダノンプレミアム	タイムフライヤー	デーオーエナジー	アドマイヤアルバ	キタノコマンドール	
5.907000	10.188433	12.168367	12.339000	8.073533	
ゴーフォザサミット	コズミックフォース	ブラストワンピース	オウケンムーン	ステイファースリッシュ	
6.824833	11.205333	5.757733	9.495733	8.515433	
ジャンダルム	エポカドーロ	グレイル	エタリオウ	ステルヴィオ	
10.117000	7.043800	8.661300	9.802067	7.016600	
ジェネラルレーヴ	ワグネリアン	サンリヴァル			
7.506833	5.710900	10.057533			

```

[1] "Yoso"
      
```

ワグネリアン 17

他17人
 コメント5件

いいね！
 コメントする
 シェアする

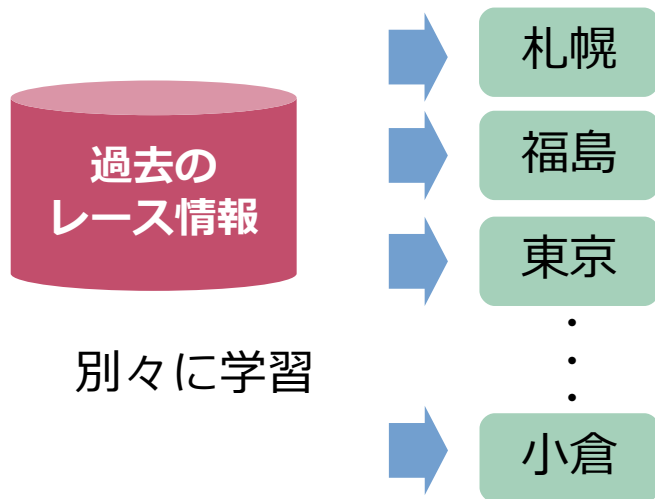

舟城亮一 やった！！大的中！
買っときや良かった！！

いいね！ · 返信 · 15週間前

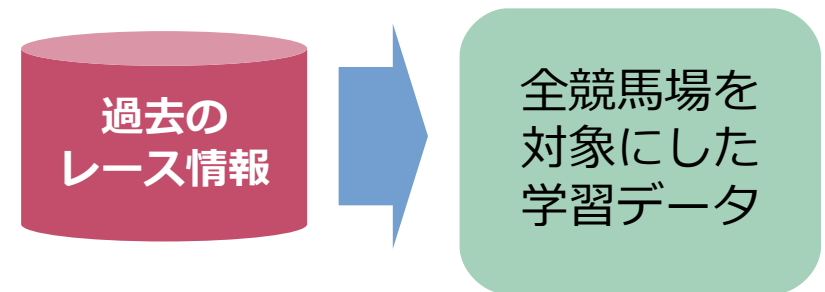
学習プロセスの「ranger」化

- 「randomForest」より速いと評判の「ranger」に変更
- JRA中央競馬（10競馬場）のデータを一度に学習

《今まで（randomForest）》



《改良後（ranger）》



Rによる競馬予想をやってみて

- マーフィーの法則が一番強い
→ 賭けると当たらない・・・
- 学習方法を改良しても、当たるとは限らない
→ 期待すると当たらない・・・
- それでも、たまに「すごい」予想が出る
→ 未来を予測するのは楽しい！

■ Rと競馬データで学ぶ統計学

<http://dokoka.org/wiki.cgi?page=R%A4%C8%B6%A5%C7%CF%A5%C7%A1%BC%A5%BF%A4%C7%B3%D8%A4%D6%C5%FD%B7%D7%B3%D8%A1%A1%CC%DC%BC%A1>

■ stockedge.jpの技術メモ 競馬の予測をガチでやってみた

<http://stockedge.hatenablog.com/entry/2016/01/03/103428>

■ JRA-VAN データマイニング予測の仕組み

<http://jra-van.jp/fun/dm/mining.html>

最新情報は
こちらをご参照ください



<http://aitc.jp>



<https://www.facebook.com/aitc.jp>



ハルミン
AITC非公式イメージキャラクター