



SIVA(シヴァ)AI競馬開発エンジニアに
よる

「競馬で始める機械学習」 ハンズオン

株式会社GAUSS R&Dセンター

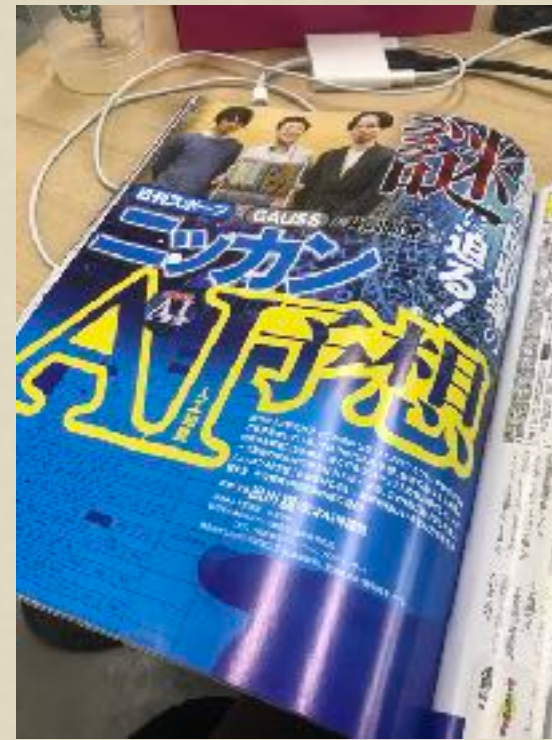
Yuta Miyawaki

Index

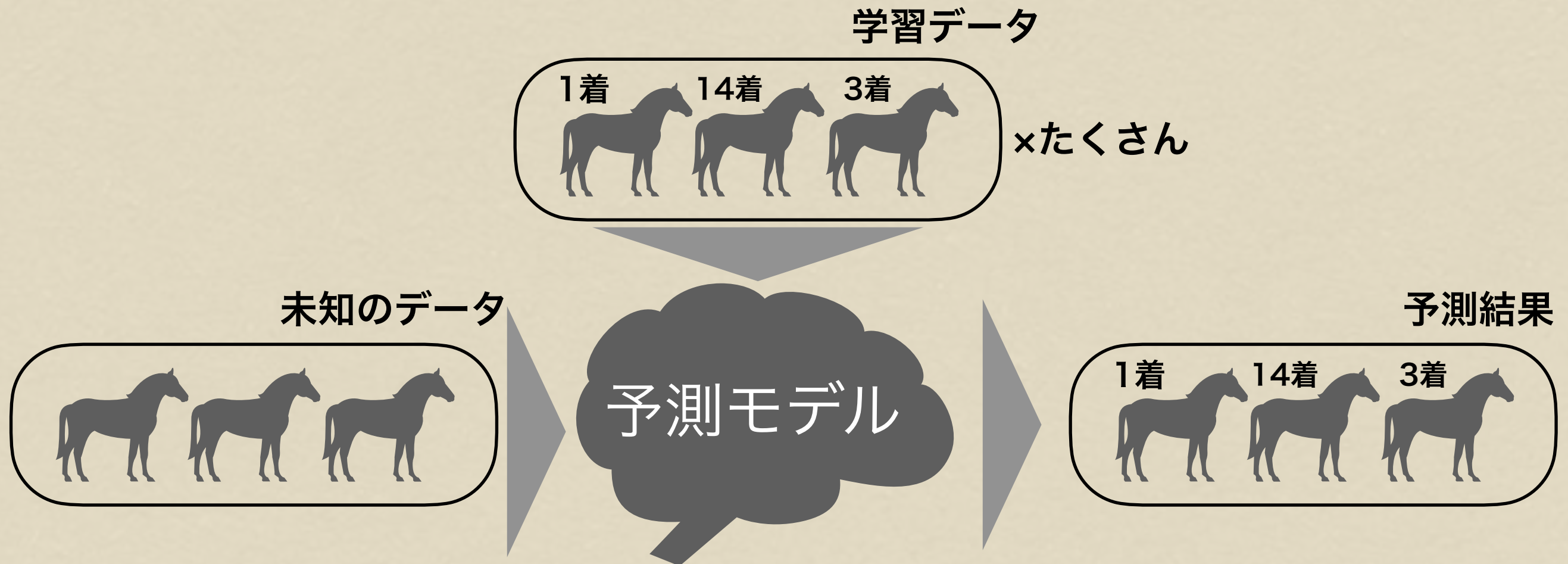
- 自己紹介
- 機械学習について
- 競馬予測の手順
- 環境構築
- 機械学習アルゴリズムの紹介
- 予測AIのカスタマイズ

自己紹介

- 宮脇 祐太 (@nami73nbj)
 - 株式会社GAUSS
 - SIVAの的中率を100%にするためGAUSSに入社
 - 競馬歴は4年半
 - 趣味は、競馬・自転車

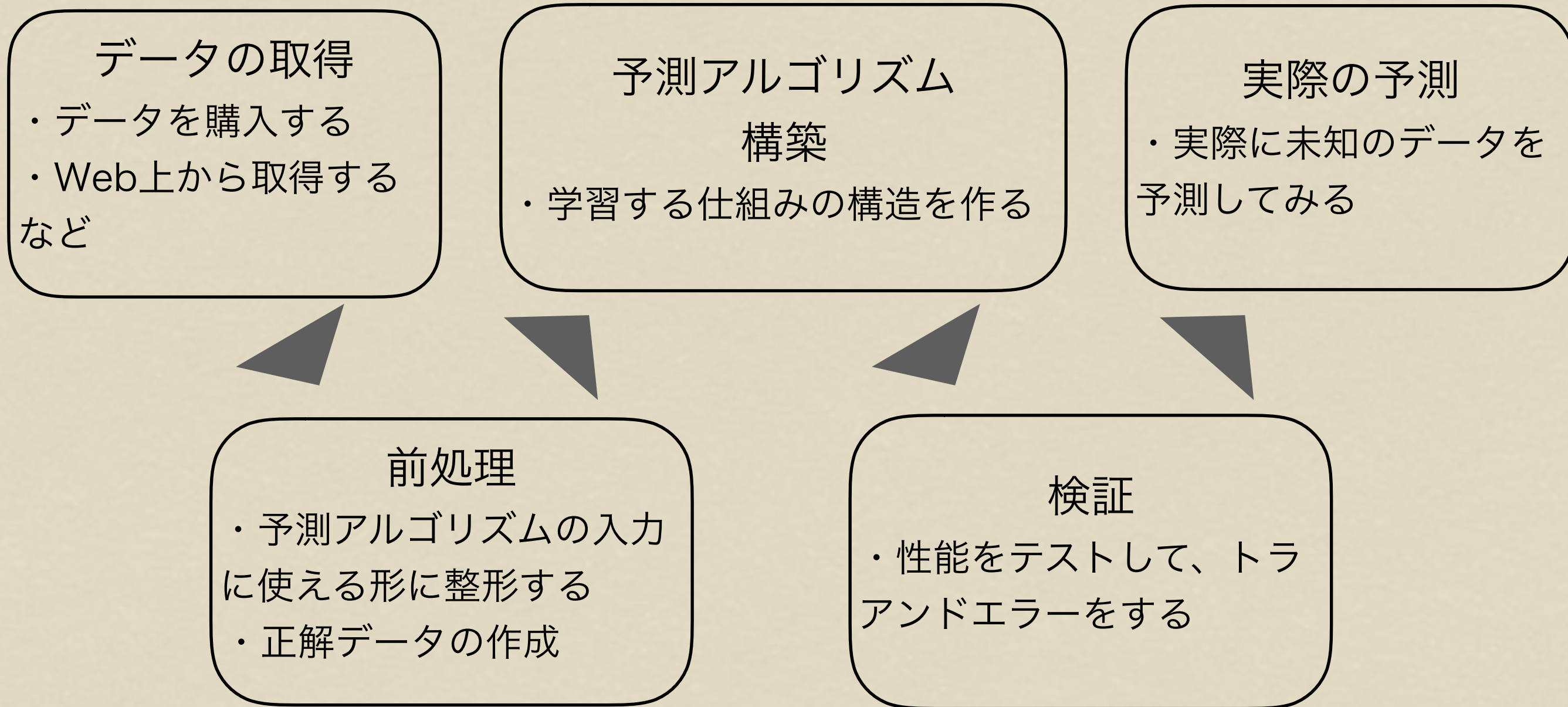


機械学習について

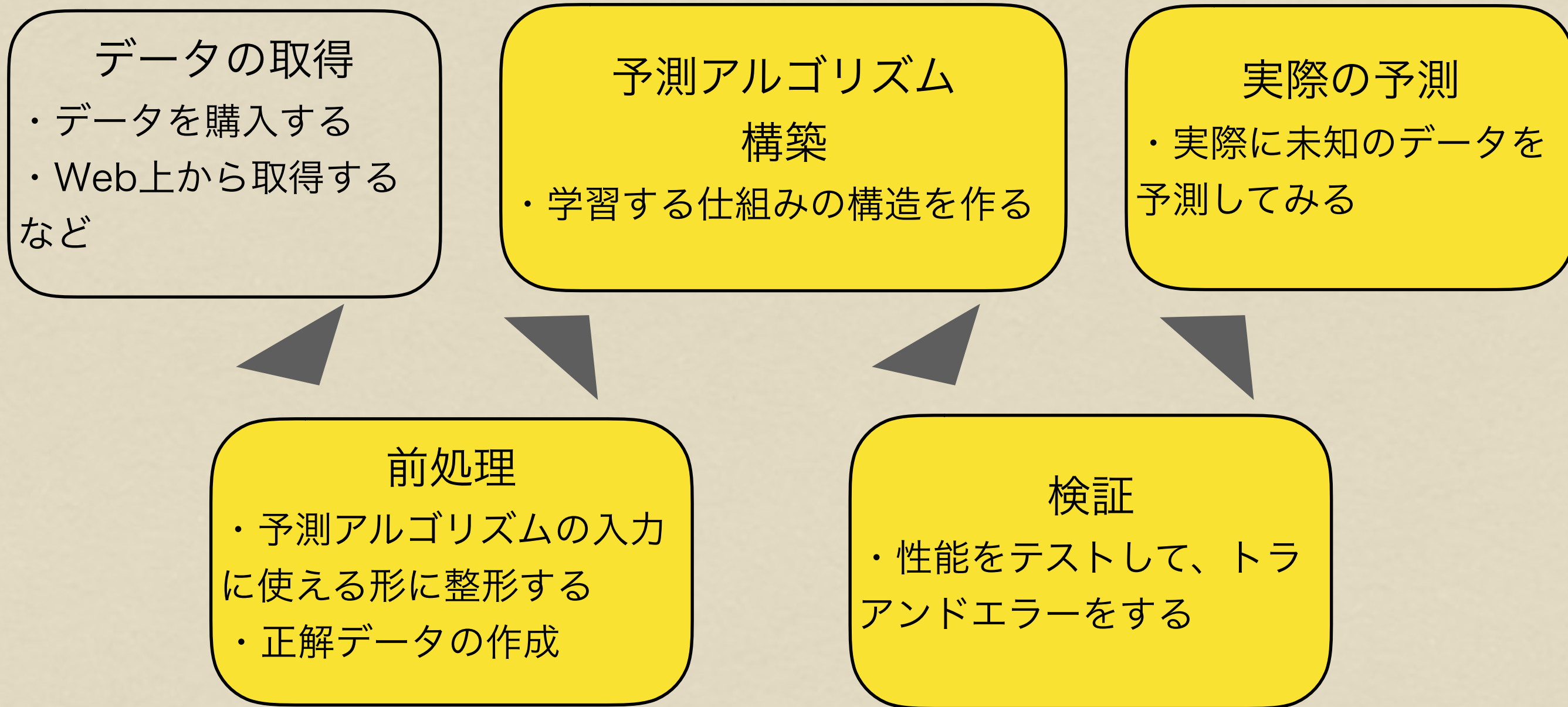


学習データから特徴・パターンを見つけ出し、
未知のデータに対する答えを見つけ出す。

競馬予測の手順

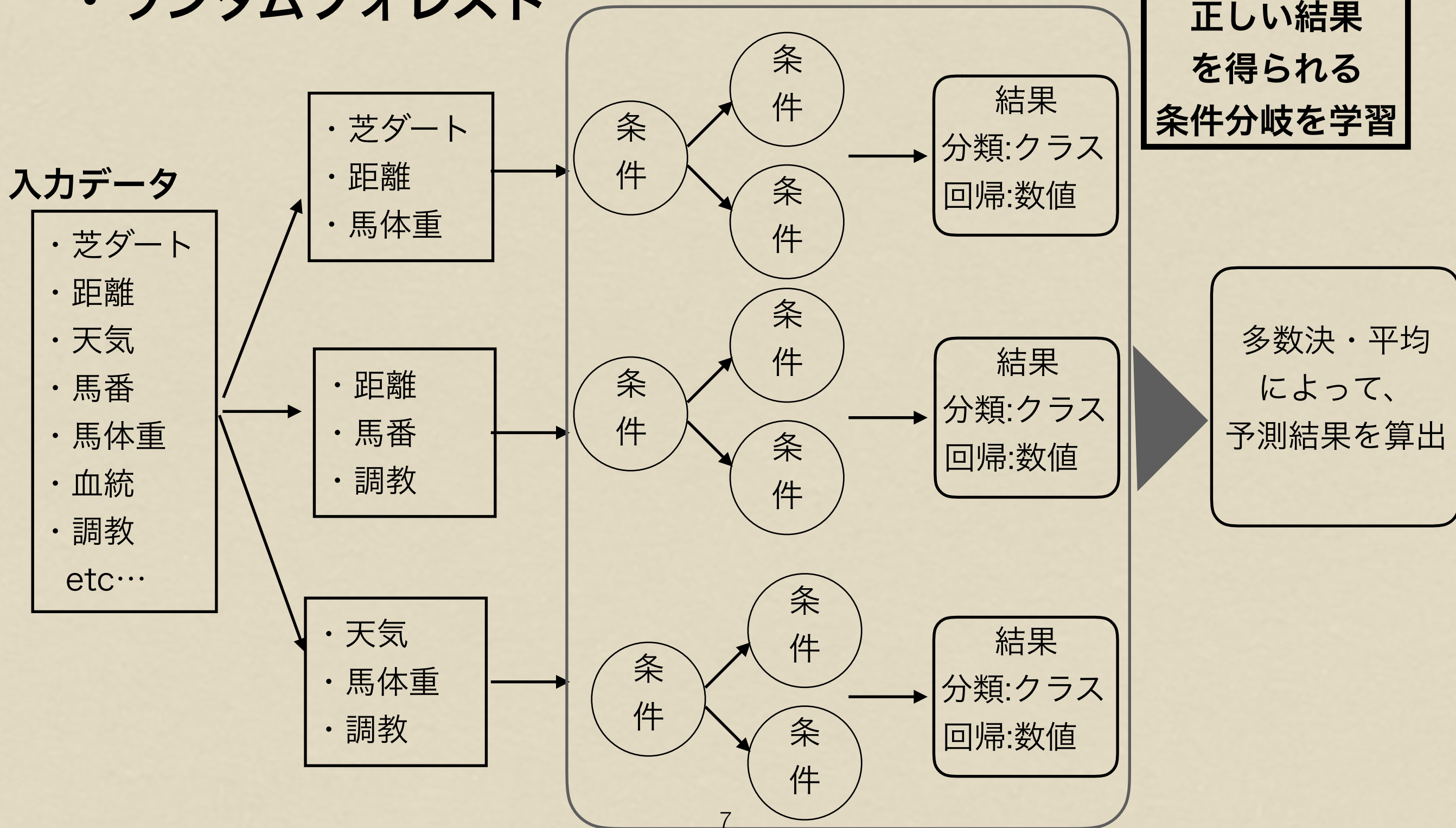


競馬予測の手順



機械学習アルゴリズムの紹介

・ランダムフォレスト

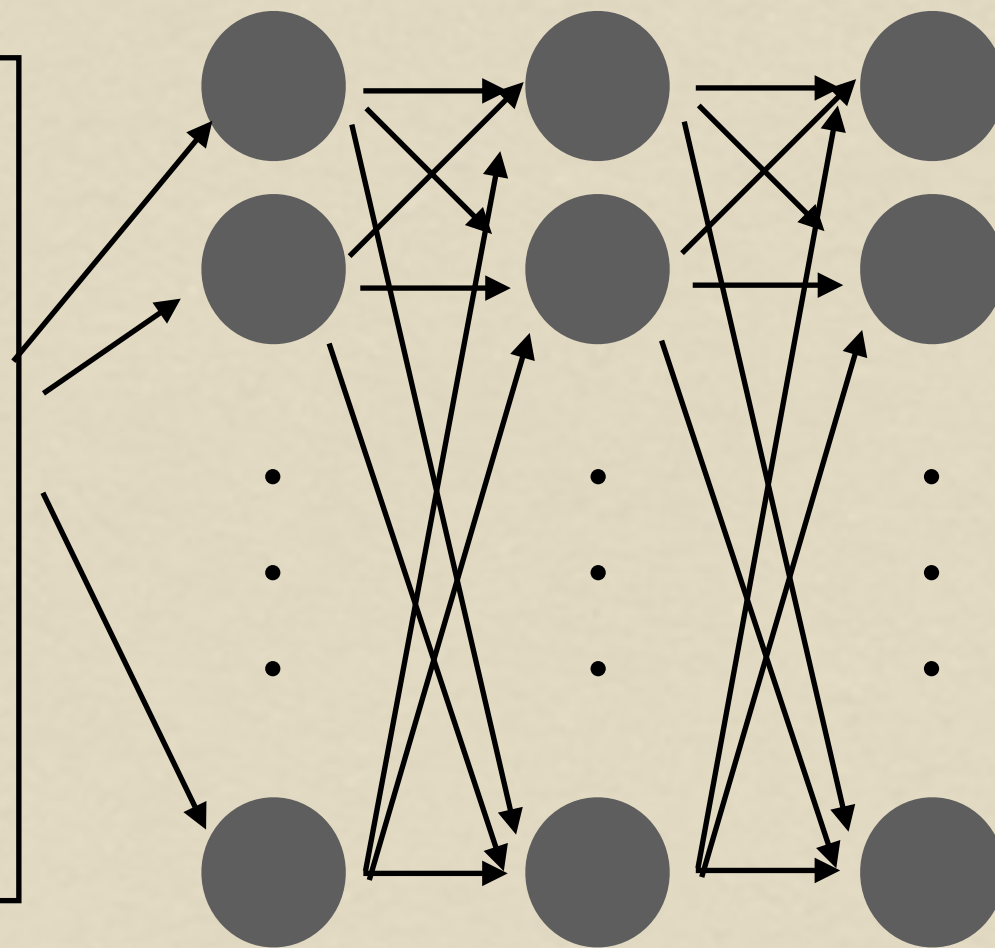


機械学習アルゴリズムの紹介

・ニューラルネット

競馬データ

- ・ 芝ダート
- ・ 距離
- ・ 天気
- ・ 馬番
- ・ 馬体重
- ・ 血統
- ・ 調教
- etc...



入力データから
正しい結果を得られる
パラメータ(矢印)を学習

予測結果

環境構築

では早速、環境構築から始めましょう

環境構築

- リポジトリからGit clone

`'git clone https://github.com/nami73b/siva_hands_on'`

- Docker Imageのビルド

`'docker build ./ -t siva_handson'`

- Dockerコンテナの起動

`'docker run -it -p 8888:8888 siva_handson'`

- jupyter notebookの起動

`http://127.0.0.1:8888/?`

`token=c64a22b43aa93db59626bb27353aa8a22580dc4b2ea47a55`

予測モデルのカスタマイズ

前処理編

[例1]

馬体重と斤量(レースの際に馬が背負う重量)の比率をファクタに追加する

[例2]

開催日付ファクタを月と日付に分割する

開催日付:1123 -> 開催月:11,開催日:12

[例3]

前走からの経過日数を計算する

予測モデルのカスタマイズ

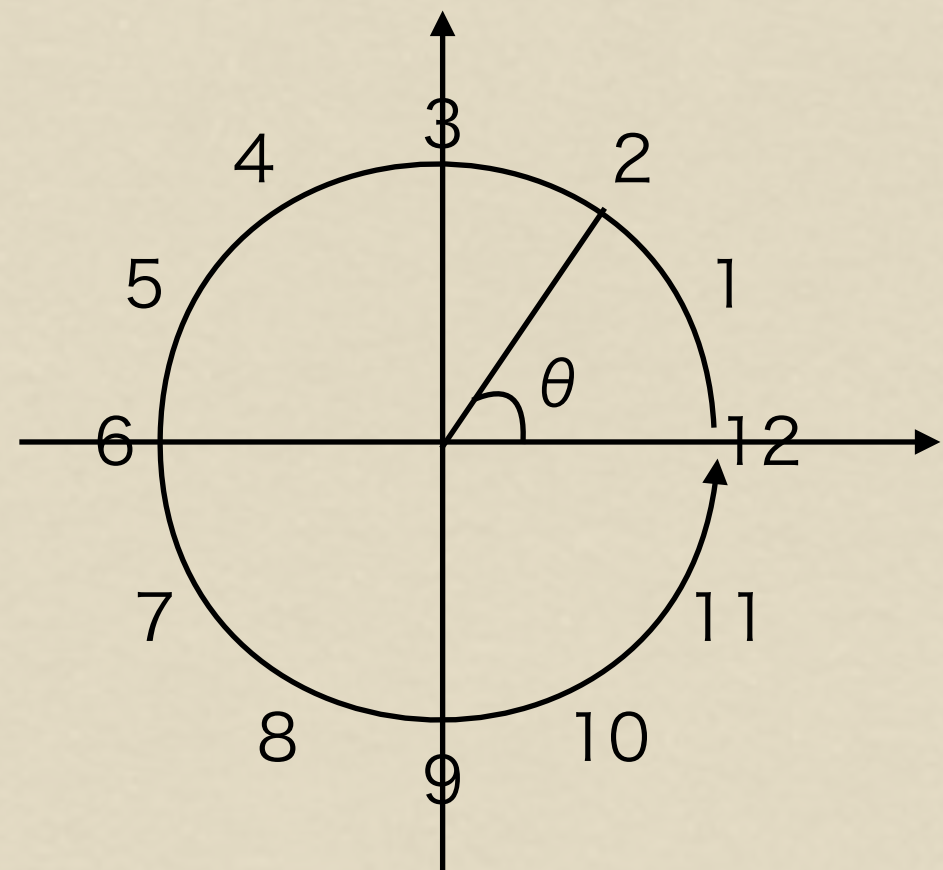
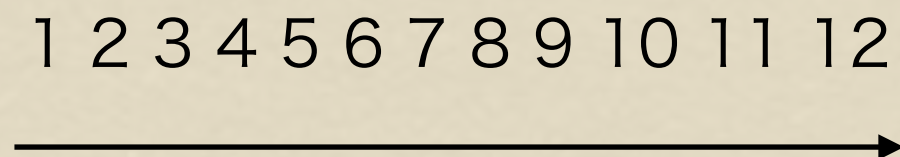
前処理編

[例4]

開催月、開催日を三角関数で表現する

$$\text{df}[\text{'開催月_sin'}] = \sin(\theta)$$

$$\text{df}[\text{'開催月_cos'}] = \cos(\theta)$$



予測モデルのカスタマイズ

アルゴリズム編

- ・ ランダムフォレスト

fit関数のハイパーパラメータを調整する

-n_estimators : 木の本数

-max_depth : 木の深さの最大

-max_features : 1つの木がランダムで選択するファクタ数
など

予測モデルのカスタマイズ

自由時間です。

本日のメインレース「中山牝馬ステークス」を各自カスタマイズした予測モデルで予測して推奨馬を上限3頭選んでください！

- ・ スケジュール

14:00~15:15 自由時間

15:15~15:30 予測結果の発表

15:45 中山牝馬ステークスの観戦

16:00 解散