

ブートキャンプ° for Data Science

4回目講義資料

Tryeting Inc. ながえゆうき

アジェンダ

- イン트로ダクション
 - ラスト2回！エンディングに向けて…！
 - 回帰の精度向上のためのデータハンドリング
- 早速やってみよう
 - データスクレイピング…株価の取得
 - 株価の取得～1ヶ月分の予測
 - 生データの回帰における精度・データ前処理後の精度比較
- 最後の宿題

- 前回までで、できるようになったこと
- 回帰・分類の機械学習は計算できる
- データ取得して、自分でモデル構築・予測までできる
- お客様に向けて、機械学習を用いたご提案ができる

- 前回までで、**まだ足りないところ**
- 回帰・分類の機械学習は計算できる
- データ取得して、自分でモデル構築・予測までできる
- お客様に向けて、機械学習を用いたご提案ができる
- **データの前処理による精度の違いを説明できる**

今日のお題

株価予測しませんか？

- 株価取得から予測の流れ
- 株価のスクレイピング ←NEW!
- 株価モデリング
 - データ前処理
- 株価予測
- 時系列データの評価
 - 決定係数(R^2)
 - 売買シミュレーション

キーワード：データ前処理、関数、評価

- 株価取得…WEBで落ちてるもの使おうぜ

半額タイムセール 18時スタート！お見逃しなく
年末プレミアム会員セール
Yahoo! JAPAN ヘルプ
ウェブ検索

ファイナンス
IDでもっと便利に新規取得
ログイン 5年目のツール・ド・東北で体感したこと

現在の日時：12月 10日 23:26
日本の証券市場はお休みです。
[利用明細](#) - [アラート](#) - [表示設定](#) ▶ [サイトマップ](#)

[トップ](#)
[株式](#)
[FX・為替](#)
[投資信託](#)
[株価予想](#)
[ニュース](#)
[金利](#)
[企業情報](#)
[証券口座](#)
[不動産投資](#)
[保険](#)
[有料情報](#)

[株式ランキング](#) | [ETF](#) | [掲示板](#) | [株主優待](#) | [決算スケジュール](#) | [レポート](#) | [IPO](#) | [みんなの株式](#) | [REIT](#) | [PTS](#) | [初心者](#) | [中国株](#) | [米国株](#)

[やさしい投資のはじめかた](#)

[ポートフォリオ](#) [機能紹介]
[ログインしてポートフォリオを表示](#)

[注目の好利回りトルコリラ建社債](#)
外貨ベースで利回り／年10.71%（税引前・複利）。10年満期償還。
www.hs-sec.co.jp

[名古屋駅の洋食のオススメ](#)
名古屋駅より徒歩3分。ナナちゃん人形からスグ。錦通りセンチュリー豊田ビル横…
little-carioca.owst.jp

東証1部・
PTS (6,191円)
OTC (27.34ドル)
アラート設定
自動 | 手動

+ 追加
6902 輸送用機器
(株)デンソー

12/08 リアルタイム株価
6,197 前日比 ↑ +87 (+1.42%)

[詳細情報](#)
[チャート](#)
[時系列](#)
[ニュース](#)
[企業情報](#)
[掲示板](#)
[株主優待](#)
[レポート](#)
[業績予想](#)
[みんなの](#)

[株価時系列](#)
[信用残時系列](#)

年初来高値：6,697（2017年11月9日） 年初来安値：4,551（2017年4月17日）

日付	始値	高値	安値	終値	出来高	調整後終値*
2017年12月8日	6,116	6,197	6,103	6,197	2,370,200	6,197
2017年12月7日	6,060	6,117	6,055	6,110	1,607,500	6,110

Q. あなたなら「A B C D」どこでこの株を買いますか？

正解はコチラ>>

特集！ 今年最大のIPO
佐川急便、上場迫る！

秋は投資シーズン？ 詳しくはこちら▶
ハ田ウィンと投資にいったいどんな関係が？

株価取得…WEBで落ちてるもの使おうぜ

2017年11月27日	6,348	6,362	6,278	6,295	2,151,000	6,295
2017年11月24日	6,318	6,374	6,302	6,343	2,328,100	6,343
2017年11月22日	6,548	6,563	6,447	6,453	1,777,700	6,453
2017年11月21日	6,502	6,536	6,435	6,498	1,498,100	6,498
2017年11月20日	6,439	6,484	6,421	6,455	1,252,200	6,455
2017年11月17日	6,500	6,537	6,425	6,440	1,935,700	6,440
2017年11月16日	6,344	6,472	6,313	6,446	1,706,500	6,446
2017年11月15日	6,426	6,437	6,336	6,345	1,944,700	6,345
2017年11月14日	6,415	6,483	6,379	6,442	1,714,000	6,442
2017年11月13日	6,551	6,591	6,470	6,477	1,411,100	6,477
2017年11月10日	6,541	6,592	6,492	6,547	2,408,400	6,547

⇒ [株価時系列データをもっと見る](#)
⇒ [信用残時系列データを見る](#)

2017 年 11 月 10 日 から 2017 年 12 月 10 日 表示

☒ デイリー ☐ 週間 ☐ 月間

この銘柄を見た人はこんな銘柄も見ています

1位 [アイシン精](#) 2位 [豊田織](#) 3位 [小糸製](#) 4位 [東海理化](#) 5位 [ケーヒン](#)

⇒ [もっと見る](#)



アルプス電気の株価が大化け？

www.masters-trade.jp

今が仕込み時の大化け株候補を無料公開銘柄診断、濱村比呂史のコラムも絶賛配信



三菱重工業株の株価が買い時

www.kabu-theme.jp

3. [理研グリーン](#) [セフと見る](#)
+21.11%
⇒ [もっと見る](#)

プロが選ぶチャンス銘柄 予想中一覧

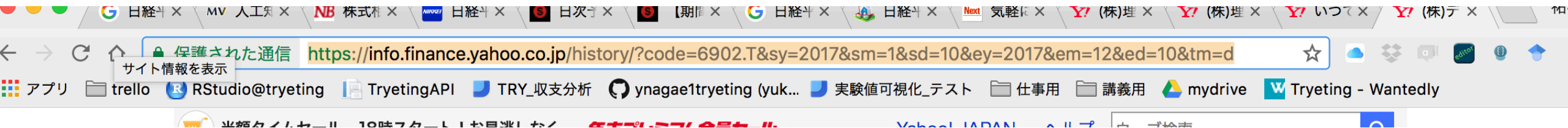
▶ [\[買い\] オムロン](#) [自信!](#) [加藤あきら](#)
▶ [\[買い\] エディア](#) [本田隆一郎](#)
▶ [\[買い\] コムシスH](#) [NEW!](#) [自信!](#) [4連勝中](#) [西村剛](#)

⇒ [日本株の予想をもっと見る \(予想中: 74件\)](#)

ヘッドライン

- ▶ [来週の相場で注目すべき3つのポイント：米FOMC、IPOラッシュ、ビットコイン先物（フィスコ）](#)
- ▶ [国内株式市場見通し：個人主体で年末高を意識した物色へ（フィスコ）](#)
- ▶ [株式週間展望＝日経平均持ち直し、レンジ昇段視野＝米利上げへの懸念は不要、ビットコイン先物には注意（モーニングスター）](#)
- ▶ [【北浜流一郎のズバリ株先見！】－銘柄に恋するなかれ！試し買いで“成功の種”を撒け（株探ニュース）](#)
- ▶ [今週の決算発表予定 シーズHD、パーク24、東京ドームなど\(12月11日～15日\)（株探ニュース）](#)
- ▶ [59歳教育費で貯蓄ゼロ。住宅ローン残債が2000万円（あるじゃん（All About マネー））](#)
- ▶ [「相続税対策の不動産投資」で見過ごせないリスク（不動産投資の楽待）](#)

- 株価取得…WEBで落ちてるもの使おうぜ



- 20170110~20171210の期間のデータ表示 = webAPI
- URL末尾のdはページ数

- 昔はこれでやったが、参考まで。禁止されてる。

- 株価取得…WEBで落ちてるもの使おうぜ：今回はこちら

株価データサイト k-db.com

トップ

先物

指数

個別株

統計



転職や 就活に。



デジタル
マーケティングの
知識やスキルを
無料で学べるe
ラーニング。



採用コストを抑えたいあなたへ

完全無料の採用支援ツール engage by en

採用サイト作成

求人掲載

応募者管理ツール

完全無料! 詳しくはこちら >

当サイトは2017年12月末でサービス終了です。

■先物

日経平均先物やTOPIX先物などのデータ 4本値+出来高+売買代金 がCSVダウンロードできます。
セッションごとデータ 日中・夜間
各銘柄ごとの時系列データ 日夜・4時間足・1時間足・30分足・15分足・5分足

■指数

日経平均やマザーズ指数など株価指数の4本値データ がCSVダウンロードできます。
セッションごと 日足・前場・後場
各指数ごとの時系列データ 日足・前後場・1時間足・30分足・15分足・5分足

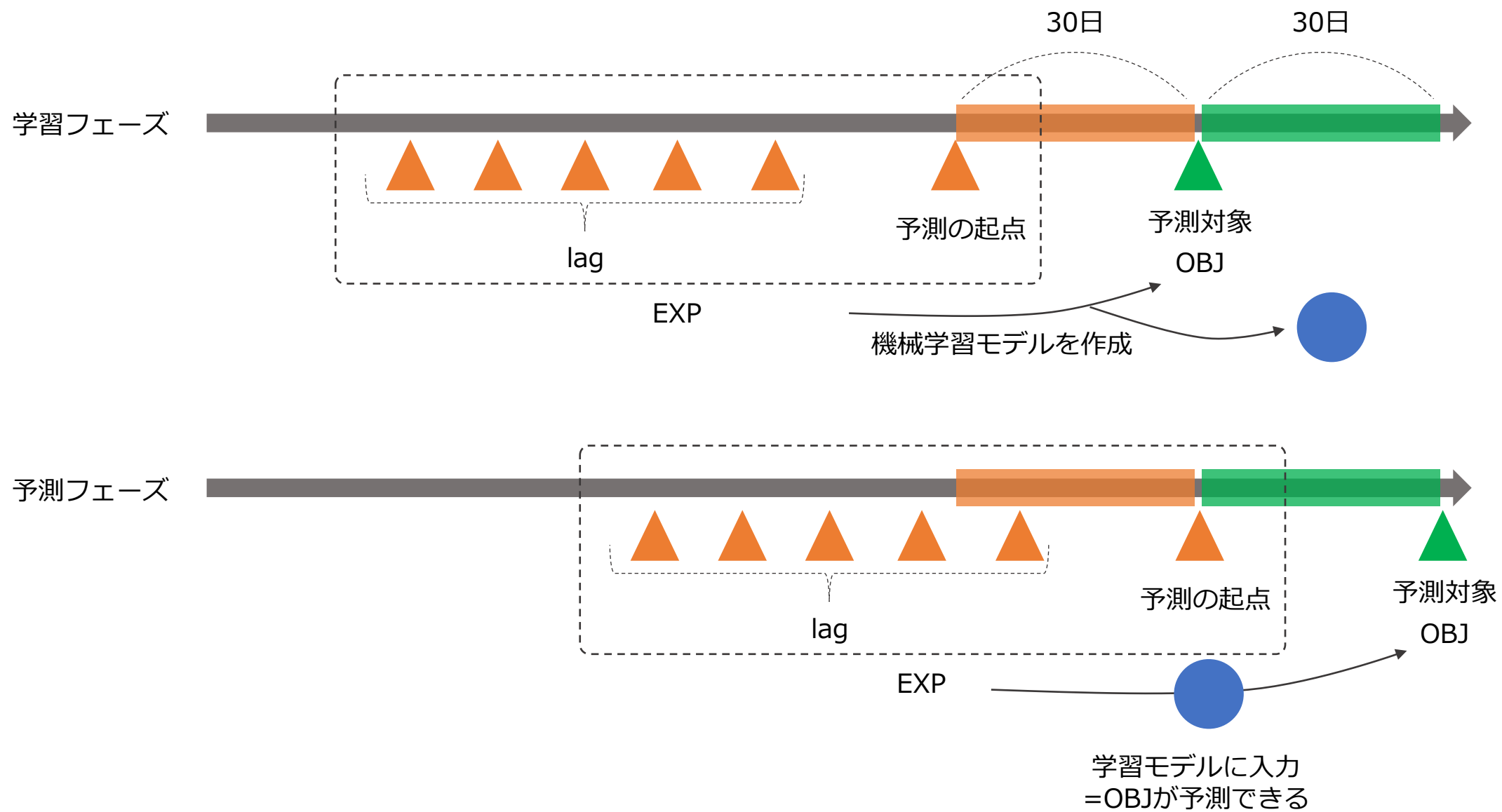
■個別株

個別株のデータ 4本値+出来高+売買代金 がCSVダウンロードできます。
セッションごとデータ 日足・前場・後場
各銘柄ごとの時系列データ 日足・前後場・1時間足・30分足・15分足・5分足

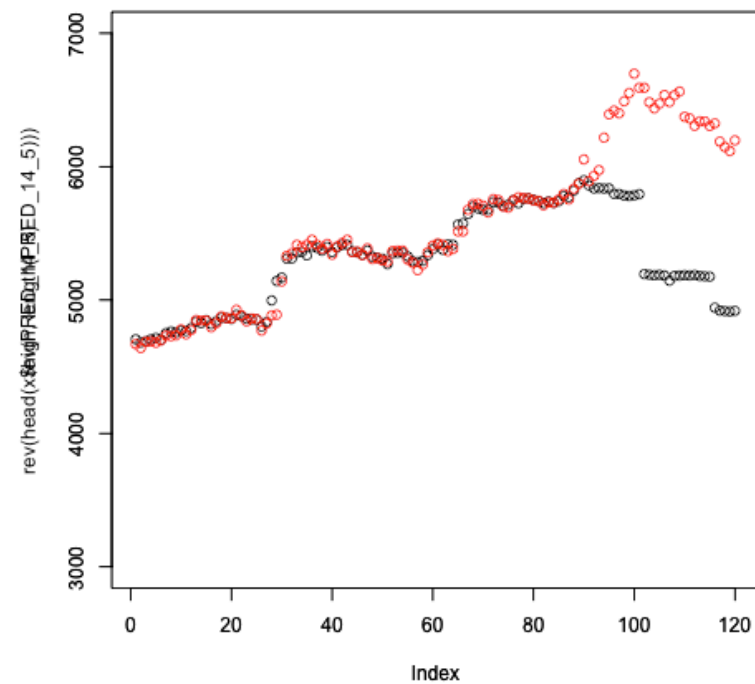
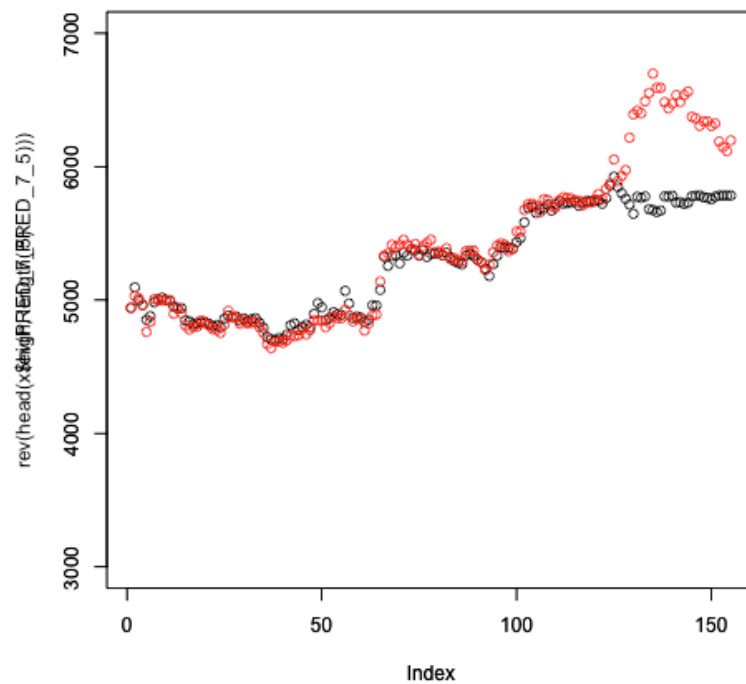
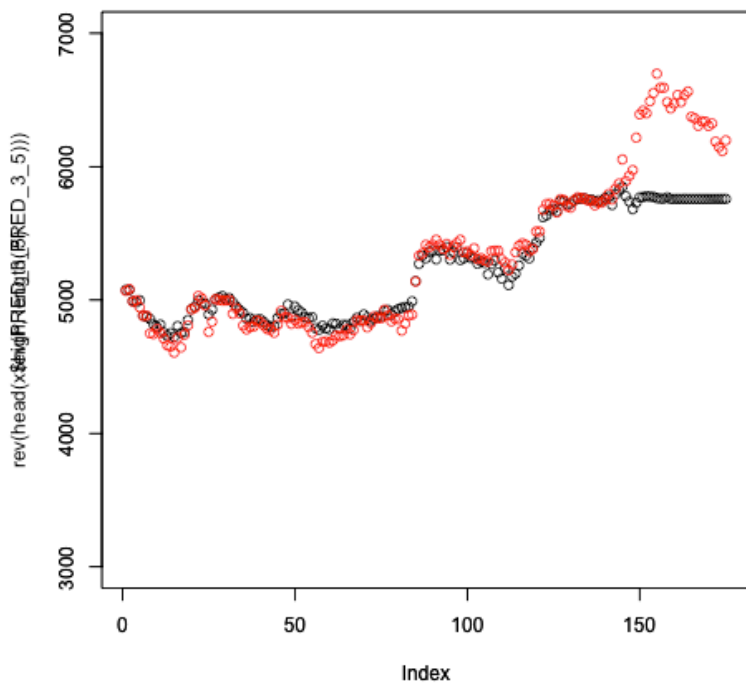
■市場統計

市場ごと・業種ごと の売買代金の集計や値上がり/値下がり銘柄数など日々の統計のデータです。
市場・業種を選択して、各市場業種ごとの 時系列表示 も可能です。

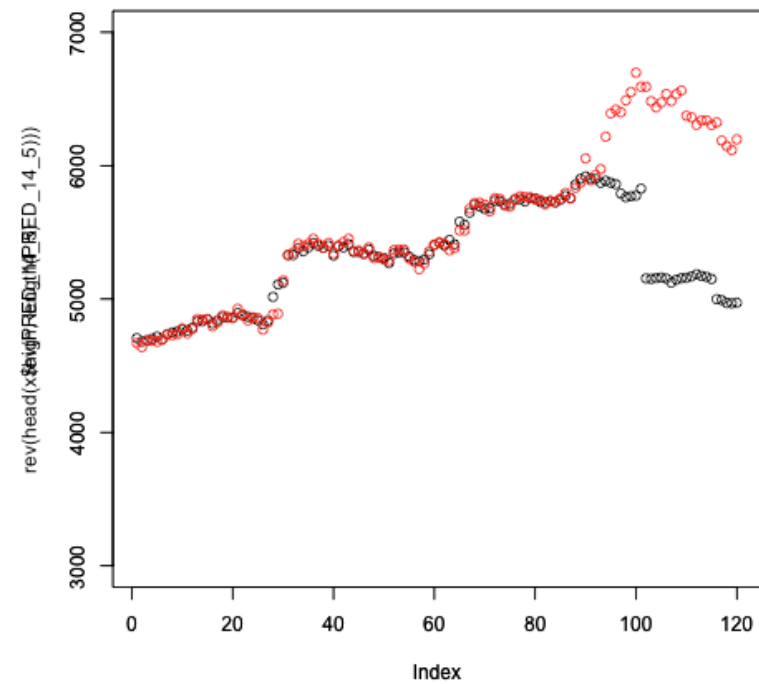
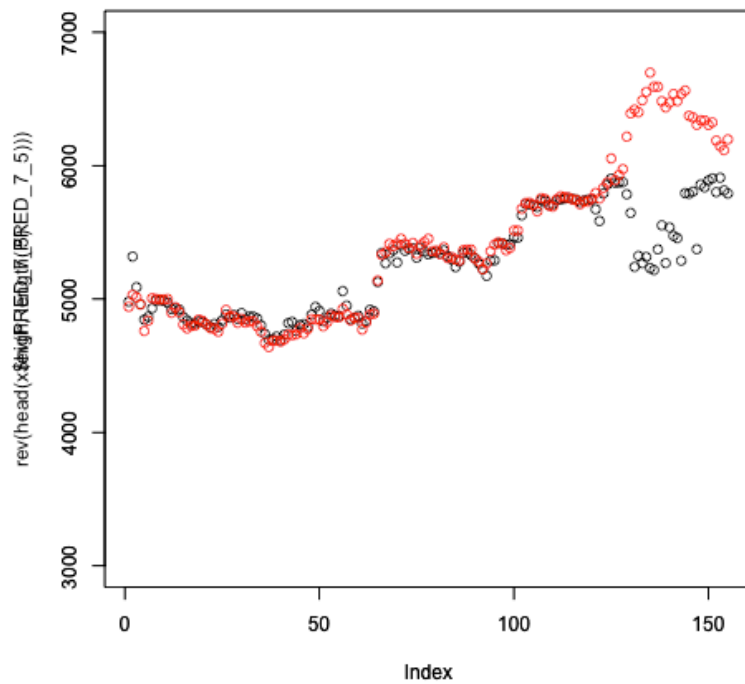
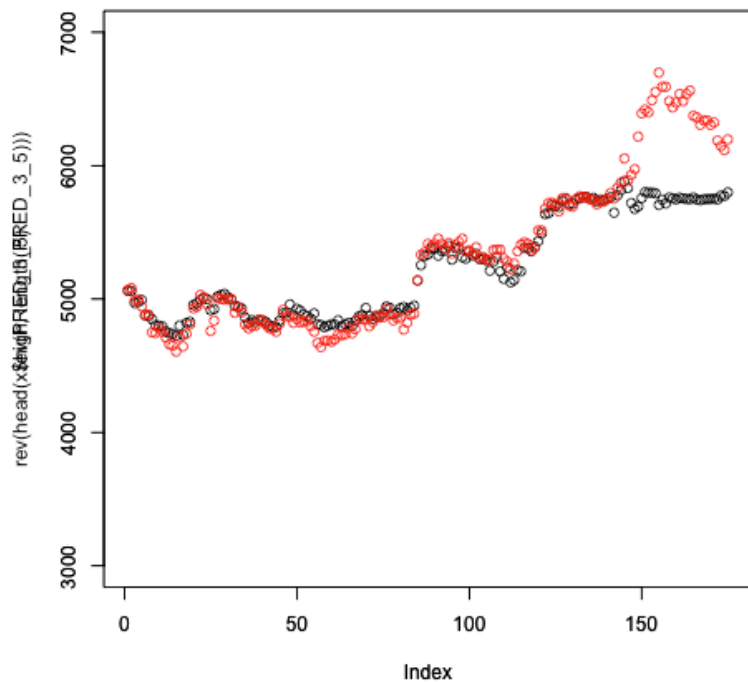
- 何やってるか



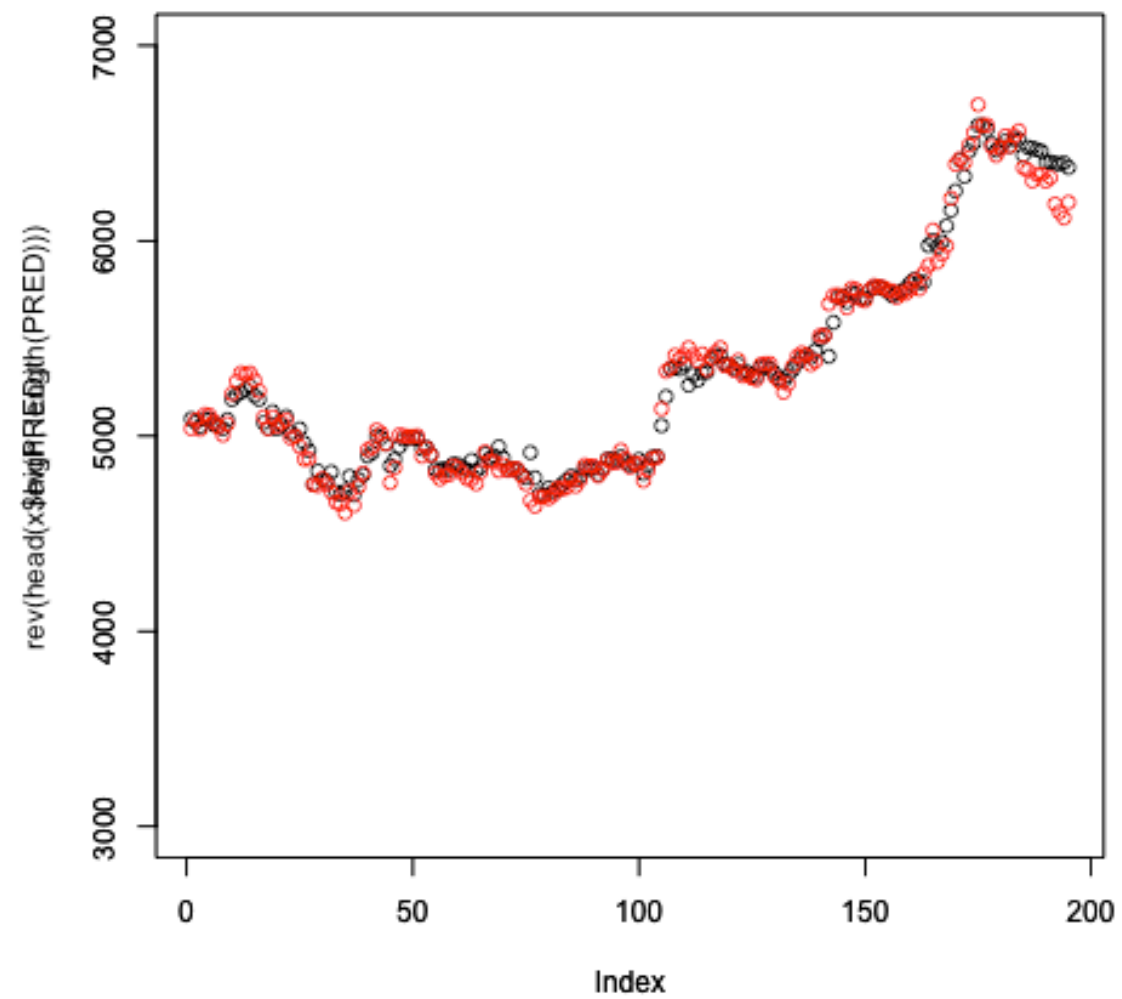
- 結果(単純に過去のデータを一緒に学習, 30日予測)



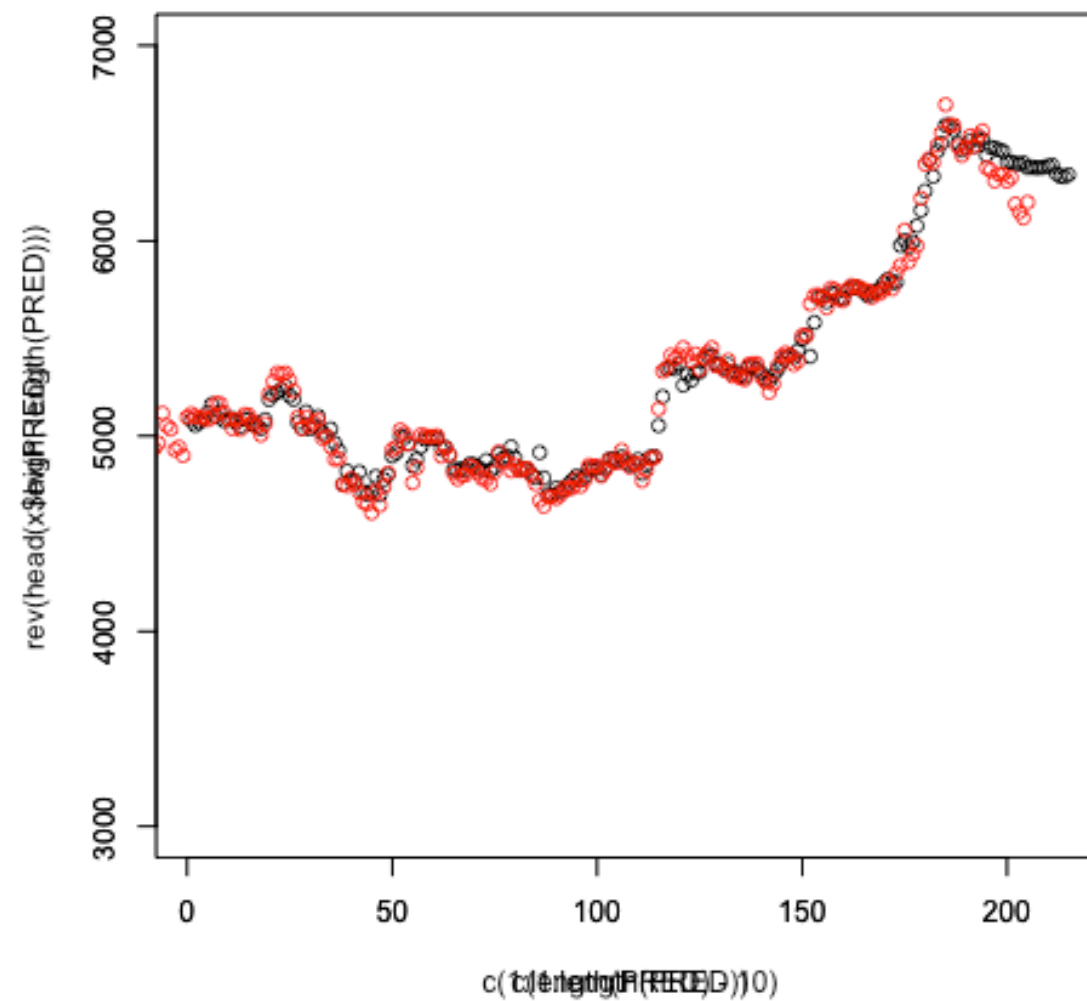
- 結果(過去データ+変化量, 30日予測)



- 結果(過去データ+変化量, 10日予測)



- ちなみに、未来予測：上がりそう？



アジェンダ

- イン트로ダクション
 - ラスト2回！エンディングに向けて…！
 - 回帰の精度向上のためのデータハンドリング
- 早速やってみよう
 - データスクレイピング…株価の取得
 - 株価の取得～1ヶ月分の予測
 - 生データの回帰における精度・データ前処理後の精度比較
- 最後の宿題

- 宿題

- デンソーの株価について、取引量も説明変数に入れて学習して評価
- パラメータを変えて、RMSEの小ささを競おう
- 図面も見せること