

株式会社データ分析 をやってみた

オペレーション基盤開発部プロセス統合開発G

西塔 明

01



自己紹介

02

基本情報

- 名前 : 西塔明
- 所属部署 : オペレーション基盤開発部
- 在籍年数 : 入社2年目
- 主な担当業務 : プロセスマイニング技術
を用いた業務プロセス分析

趣味

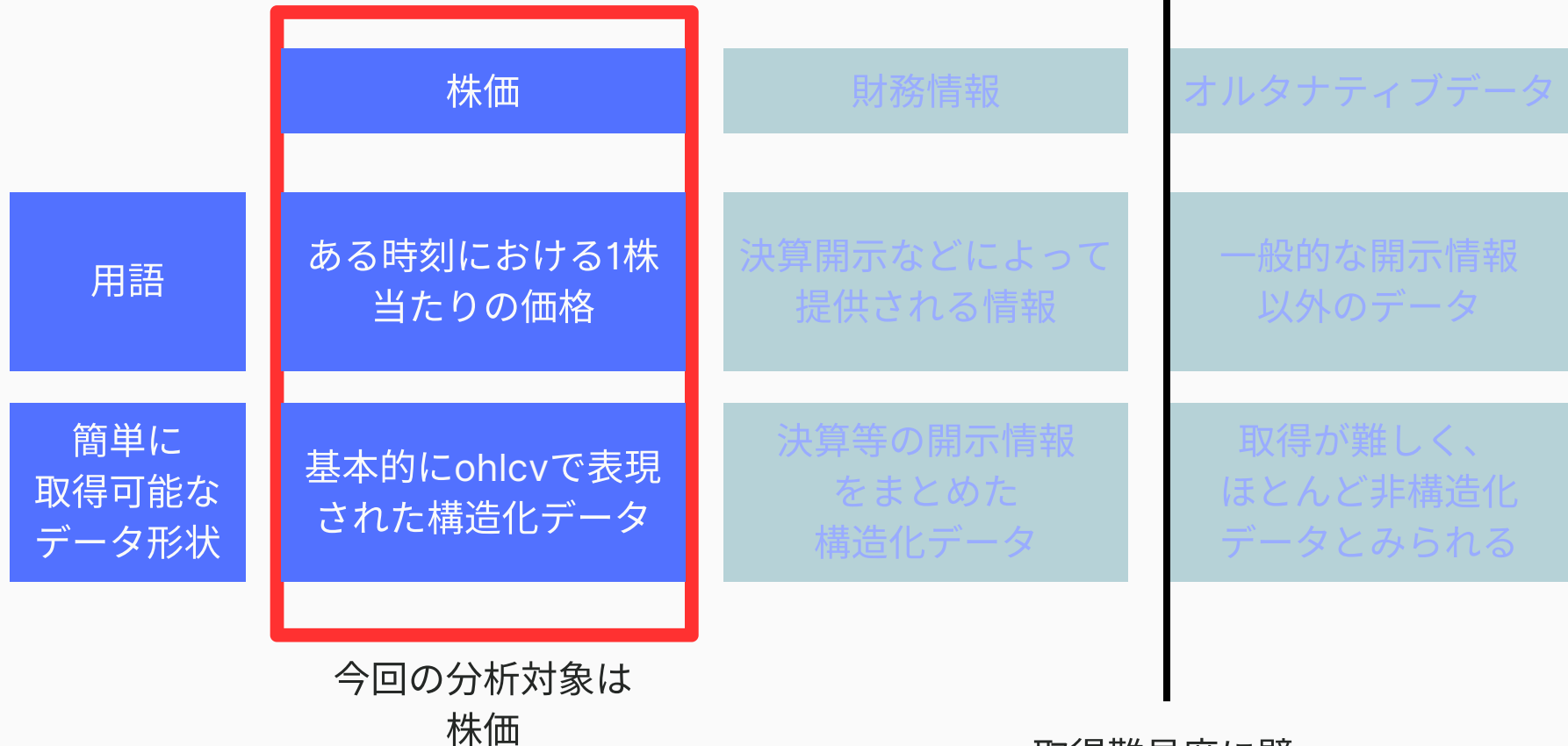
- ゲーム、バラエティ鑑賞、ラジオ、**金融取引全般**

今日の目的

- 株式データ分析の面白さを感じてもらう



株式データの種類 03



取得難易度に壁

※ChatGPTに構造化データ化させるなどの工夫で
ここも実は変革期が来てる？

Pythonライブラリ04 Dashで可視化

グラフ参照

- Flask, React, Plotlyが組み合わさって作られたライブラリ
- 株価はオープンソースAPIを使用して取得
- OHLCVの4本値＋取引高がデフォルト

Q.そもそも株価のデータ
分析して何をするの？

A.分析結果を元に
売買の判断をする



Q.何が面白いの？

06

A.資産を増やすという

同じ目的に対して

アプローチが多数存在するところ

- 例:テクニカル分析とファンダメンタル分析
- クロスセクション分析と時系列分析(1)
- 取引根拠も手法による投資期間もばらばら

今回の取組み

07

株価(指数)の過去データを用いた
売買シミュレーション

単純な売買シミュレーション

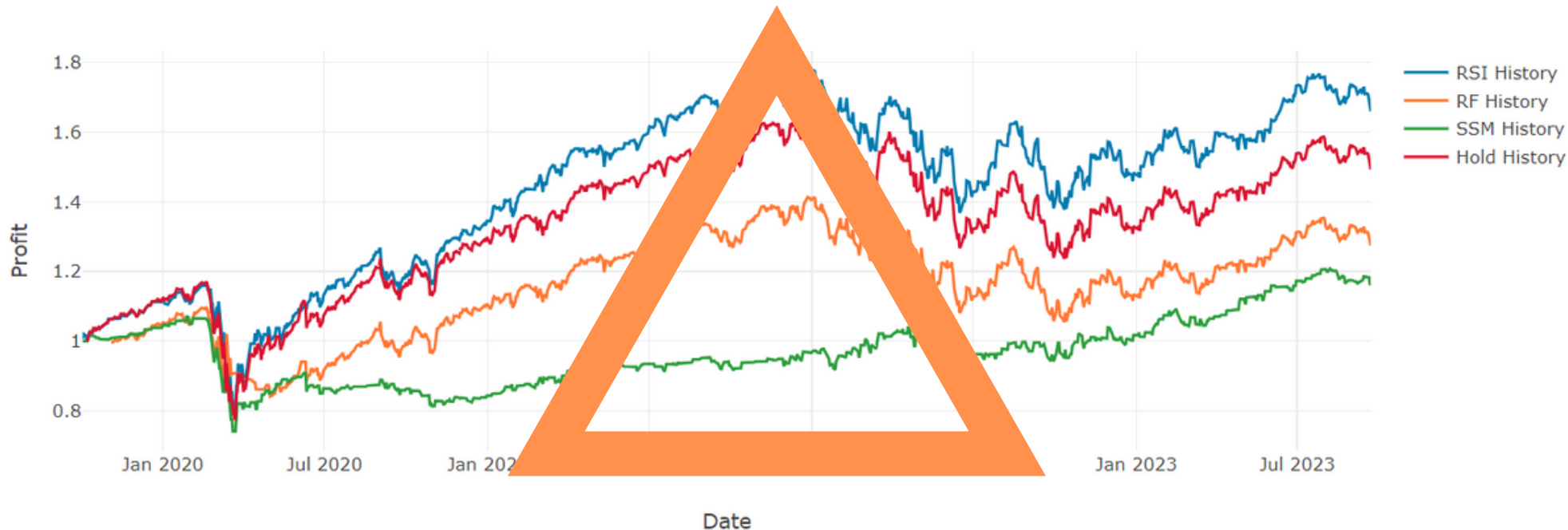
Adjusted Profit Histories for RSI, RF, SSM, and Hold Strategies



2019年10月~2023年9月までのS&P500指数を各取引戦略に基づいて売買シミュレーションを行った際の累積リターンの推移
ルールベースモデル、機械学習モデル、統計モデルを作成し、初日に買っただけのBuy&Holdモデルと比較

単純な売買シミュレーション

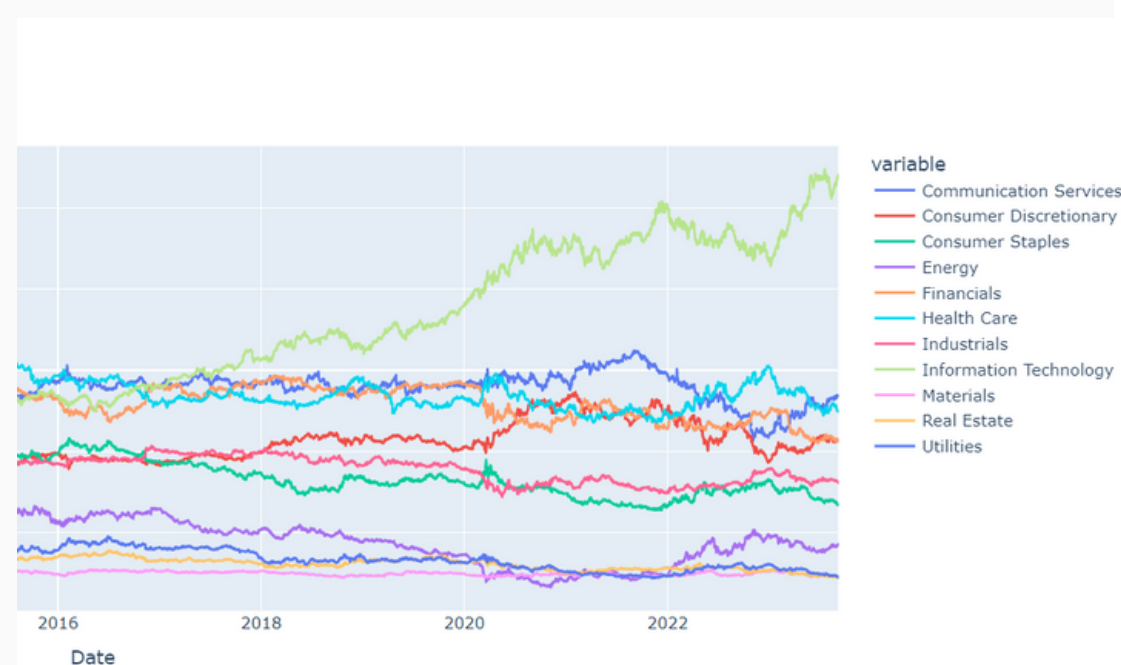
Adjusted Profit Histories for RSI, RF, SSM, and Hold Strategies



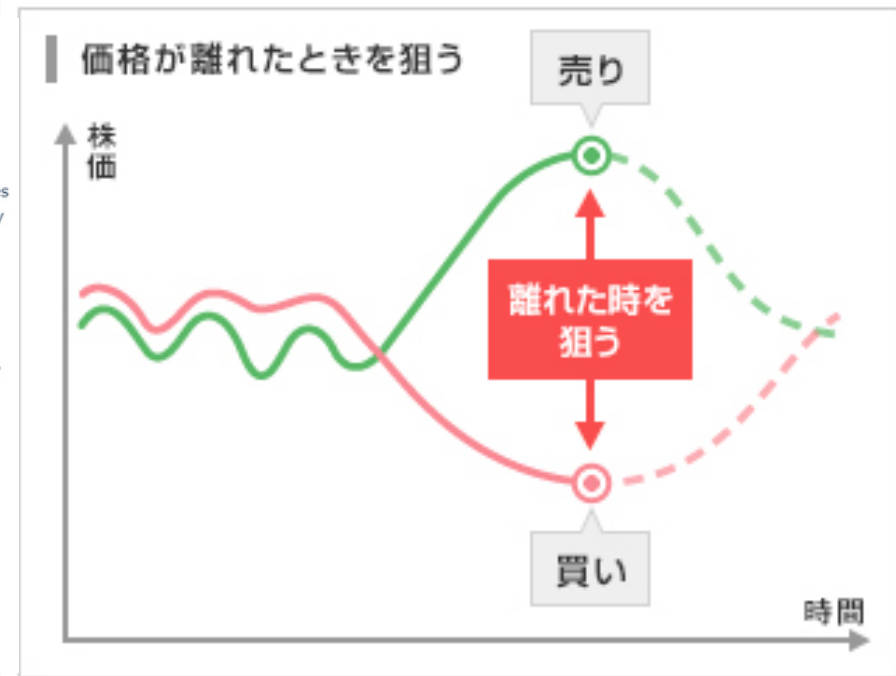
- 単純な株価予測に基づいたトレードではトレンド変化やノイズの影響により結局**Buy&Hold戦略に劣る**結果
- 実は、過去の時系列データを用いて予測対象(目的変数)に株価そのもののリターンを設定しても良い結果が得られない、と一般的に言われている

取引戦略を考える

10



S&P500の全体の時価総額に占める
各セクター時価総額の割合推移



ペアトレードの概念図(2)

似たような推移をしているセクターが複数存在したので
ペアトレード戦略をとれるのではないかと仮定

取引戦略結果

11

グラフ参照

- ペアの選定は共和分検定、売買判定はテクニカル指標
- 累積リターンの推移が改善された
- 各種シミュレーション評価指標の数値も実は良くなっている
- 最も優れたペアは**4年で原資産が2.5倍**(年利約25%※1※2)に
(参考:Buy&Hold戦略4年で1.4倍)
- 乖離幅の予測や乖離にかかっている時間などを予測(3)することで
更なるリターンの改善余地あり

※1:取引手数料、空売り期間の金利手数料未考慮

※2:仮想的なセクターポートフォリオのため、現実にはETFで代替が必要

まとめ

12

**過去データのみで株価を
当てるのは難しい
取引戦略が大事**

本発表は特定の商品の勧誘や売買の推奨等を目的としたものではありません

参考文献

- (1)中川 慧,価格変動パターンを用いた 株価予測手法の実証研究
- (2)信用取引の活用方法
- (3)東出 卓朗,初到達時間を用いた ペアポートフォリオ最適化問題の新定式化
- (4)ペアトレード(証券用語解説集野村證券)
- 以下本文中引用以外の文献
- (5)JINAN ZOU,Stock Market Prediction via Deep Learning Techniques: A Survey
- (6)伊藤 克哉,ABCD-Forecast:機密金融時系列予測のためのデータ拡張バギング 手法
- (7)吉川大介,データ駆動型ファイナンス,共立出版P7-10,101-108

用語集

用語	意味
RSI	相場の相対的な強弱を示す指数 テクニカル指標
MACD	長短の指数平滑移動平均を元に売買を 判断する指標
SSM	状態空間モデル、観測値の推定ではな く状態(確率変数)を推定する方法
ペイトレード	相関関係の高い二つの金融商品の差額 の変動を利用した取引方法。(4)

金融取引とは

03

金融商品を売買すること



株式

今回の分析対象



債権



外国為替



コモディティ



暗号資産

なぜ株式データなのか？

※個人の感想です

	株	債権	外国為替	コモディティ	暗号資産
データ 取得難易 度	○	△	○	△	◎
データ種 類	◎	○	△	△	○
参考情報 の豊富さ	◎	△	△	△	○

おまけ考察

13

機関投資家に優位性はあるのか？

機関投資家は時間的制約(顧客の存在)、戦略のリスク許容度の制約(日本ではDDが10%を超えたらその戦略は止めるそう)、取り扱い金額から生じる制約(自分の取引によって予測した相場から外れる)などがあり、取れない戦略が多々あるのでその点ではあると思われる。

