# 機械学習エンジニアコース卒業課題

# 株式トレードにおける銘柄分析と 株価シミュレーション

第2101期 遠藤 伸

# アジェンダ

- 0. はじめに
- 1. 株価のテクニカル分析における課題
- 2. 課題解決のためのソリューション
- 3.機能一覧
- 4. 分析に使用するデータ(株価・投資指標・財務)
- 5. 分析に使用するデータ(情報サイトスクレイピング)
- 6. シミュレーション(投資指標)
- 7. シミュレーション(株価)
- 8. 学習と推測
- 9. 機能および技術要素

2021/5/14 **2** 

### 0. はじめに

名前 :遠藤 伸

出身校:青山学院大学経済学部

経歴:日本ユニシス株式会社 勤続32年(2020年6月早期退職)

データミックス、DIVE INTO CODEで機械学習を学ぶ

趣味 : ウインドサーフィン、ダイビング、写真撮影

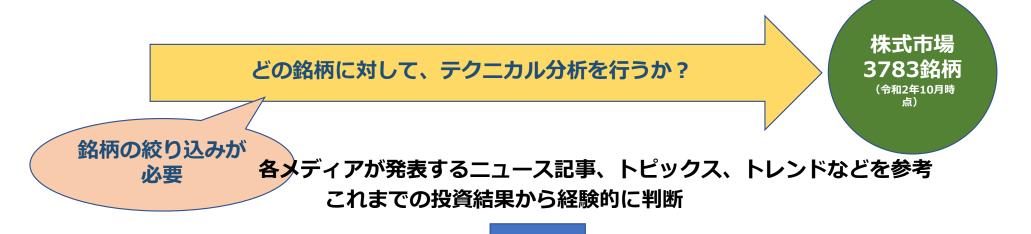




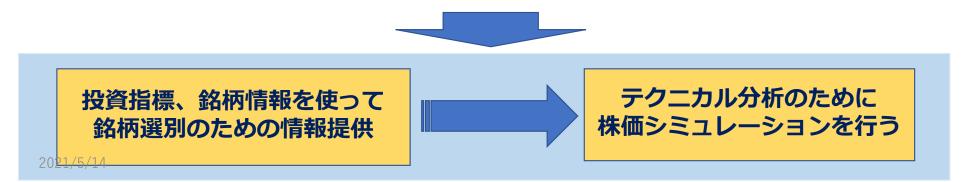


### 1. 株価のテクニカル分析における課題

課題: テクニカル分析する対象銘柄の選別が難しい



どの銘柄を絞り込むか、明確な基準がなく、投資初心者にとって銘柄選別が難しい。



### 2. 課題解決のためのソリューション

課題解決のために、「銘柄情報表記機能」,「投資情報シミュレーション機能」,「株価シミュレーション機能」 を提供する。

- 銘柄情報表示機能
- ・投資情報シミュレーション機能

投資指標、銘柄情報を使って銘柄選別のための情報提供



情報サイトをスクレイピングし、 銘柄情報を提供する。

投資指標、財務データ、国債金利 情報の過去データを使い、シミュ レーションする。

・株価シミュレーション機能

テクニカル分析のために 株価シミュレーションを行う

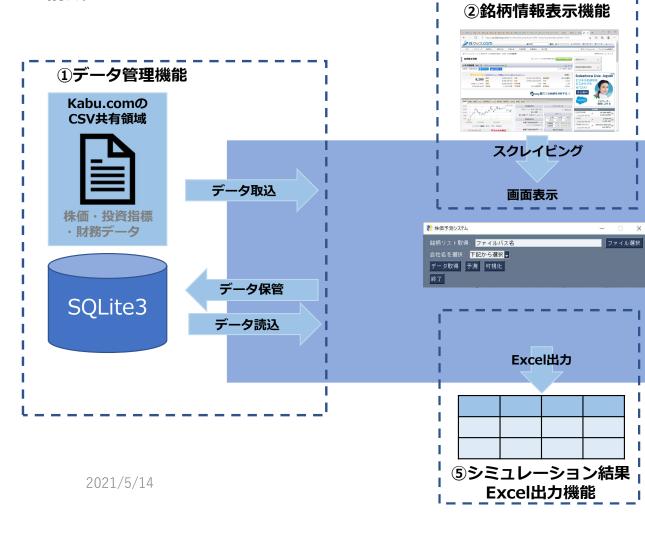


株価シミュレーションを行う。

# 3. 機能一覧

本機能は、「データ管理機能」,「銘柄情報表示機能」,「Excel出力機能」,「シミュレーション機能」により

構成されている。







6

### 4. 分析に使用するデータ(株価・投資指標・財務)

個人投資家向けに企業の株価、投資、財務指標などのデータを提供するZEROBYTE社のサービス「KABU+」および、財務省が提供する固定利付国債の金利情報(<a href="https://www.mof.go.jp/jgbs/reference/interest\_rate/">https://www.mof.go.jp/jgbs/reference/interest\_rate/</a>)を利用。

#### 株価一覧表

NO	項目名	項目説明	項目値(例)
1	SC	証券コード (StockCode)	130
3	名称	会社名	板洋
3	市場	株式市場	東証一郎
4	常程	会社の業種	水産・農林
- 5	日時	取引日時	2018/1/4 15:01
6	株価	本日取引終了後の株価	4,32
- 7	前日比	前日取引終了後の株価からの差額	25
8	前日比(%)	前日取引終了後の株価に対する差額の割合	0.51
9	初日終値	前日取引終了後の株価	4,29
10	始值	本日取引の株価給値	4,27
11	高値	本日取引の株価高値	4,33
12	安値	本日取引の株価安値	4,22
13	出来高	本日売買された取引量	61,50
14	売買代金(千円)	本日売買が成立した金額(与株価×出来高)	263,44
15	特価総額(百万円)	株価×発行済株式数	47,21
16	值模下段	値幅制限の上限値(ストップ高の特価)	3,59
17	值幅上限	価幅制限の下限値(ストップ安の株価)	4.99

#### 投資指標データ

NO	項目名	項目説明	項目値(例)
***	SC	証券コード (StockCode)	1301
2	日時	取引日時	2018/1/4 15:00
3	名称	会社名	板洋
- 4	市場	株式市場	東延一郎
- 5	章程	会社の業種	水産・農林
6	時值総額(百万円)	将值×発行済株式数	47,210
7	発行済株式数	会社が約款で示した株式数の範囲内で、実際に発行した株式の総数	10,928,283
8	配当利回り(%)	1株当たりの年間配当金/現在株価	1.16
9	1株配当	1株当たりの年間配当金	50
10	PER (予想)	将価収益率:1株当たり当期執利益/現在株価	16.81
11	PBR (実績)	林価純資産信率:株価 / 1株純資産 (総資産-負債)	1.72
12	EPS (予想)	1株当たり当期純利益:当期純利益 / 発行済株式数	257.06
13	BPS (実績)	1株当たり純資産:純資産/発行済株式数	2504.43
14	最低購入額	現在核価 × 単元株	432,000
15	単元株	最低売買単位	100
16	高値日付	年初来高値のついた日付	2017/12/26
17	年初来高值	年初から現在までの最高値	4,415
18	安値日付	年初来安値のついた日付	2017/1/18
19	年初来安值	年初から現在までの最安値	2.681

#### 10年償還国債金利情報

NO	明日名	項目投明	項目値(例)
1基準	8	金利の計上基準日	2020/8/3
219		1年で信速する国債の年率	-0.151
32.4		2年で債遇する関債の年率	-0.14
43.9		3年で個選する国債の年率	-0.141
54.9		4年で信息する国債の年率	-0.139
65%		5年で信道する団債の年率	-0.112
76年		6年で信選する国債の年率	-0.115
87年		7年で信遂する国債の年率	-0.113
98年		8年で債遇する国債の年率	-0.083
1097		9年で償還する団債の年率	-0.024
11109		10年で偏進する国債の年率 分析用データとして総合後。「安全資産利子 事」項目として使用する。	0.025
1215年		15年で信用する国債の年平	0.254
1320年		20年で保護する国債の年率	0.398
1425年		25年で慎速する国債の年率	0.485
15/30≆		30年で保護する国債の年率	0.554
16403		40年で信用する国債の年率	0.564

決算・財務・業績データ

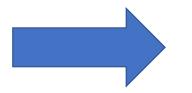
NO	項目名	項目説明	項目値(例)
1	SC	証券コード (StockCode)	130
2	名称	会社名	SEX
3	決算期	事業年度の区切り:月-年	Mer-1
4	决算竞员日(本決算)	決算日から45日以内に発表する日	2017/5/11
- 5	先上高(百万円)		236,56
6	定案利益(百万円)	売上高・売上原価・販売費および一般管理費	3,72
7	经常利益(百万円)	営業利益 + 営業外利益 - 営業外費用	3,70
. 8	出期利益(百万円)	超常彩益 + 特別彩益 - 特別損失	2,423
9	松資産(百万円)		97,39
10	自己資本(百万円)	純資産: 総資産・他人資本	24,97
11	資本金(百万円)		5,66
12	有利子負債(百万円)		50,915
13	自己資本比率	NAR / 19/10 / /	25.0
14	ROE	自己資本料益率 新期的制造/ (%的資 - 副除于纳州 - 少数株主約分)	10.19
15	ROA:	和資産利益率: 当期終刊益/総資産	2.53
16	免行消除式数		10.928.283

#### 分析用データ(銘柄別収益情報)

NO	項目名	内容
1	SC	銘柄を識別するキー
2	日時	時系列のシミュレーションを行う時点
3	収益率	時価総額における日次変化率
4	市場収益率	全銘柄の時価総額に対する平均日次変化率
5	超過収益率	日次国債金利に対する収益率の超過率
6	市場超過収益率	日次国債金利に対する市場収益率の超過率
7	25日平均乖離率	過去25日の収益率平均との差分
8	マーケットベータ	過去125日分の市場超過収益率の超過収益率への回帰係数
9	企業規模	時価総額(株価×発行済株式総数)
10	簿価時価比率	自己資本(百万円)/ 時価総額(百万円)
11	財務レバレッジ	総資産(百万円)/ 時価総額(百万円)
12	株価収益率	当期利益(百万円)/ 時価総額(百万円)

#### 分析用データ(銘柄別株価情報)

NO	項目名	
1	SC	
2	日時	
3	始値	
4	高値	
5	安値	
6	終値	



対象銘柄の選別 (シミュレーション)

過去の投資指標 シミュレーション

(指定日付)

上位、下位の ランキング

対象銘柄の株価分析

株価 シミュレーション

## 5. 分析に使用するデータ(情報サイトスクレイピング)

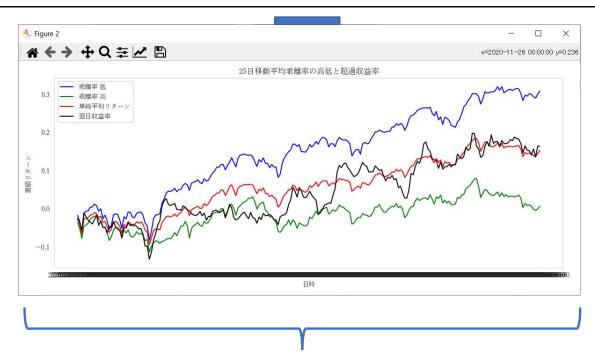
情報サイトからスクレイピングにより、銘柄情報を取得する。



### 6. シミュレーション(投資指標)

銘柄ごとに各投資指標によるランキング分けを行い、上下20%の銘柄の収益率をプロットする。

企業規模、簿価時価比率、財務レバレッジ、株価収益率、25日移動平均乖離率、マーケットベータ



- ・銘柄ごとに投資指標(超過収益率、株価収益率、マーケットベータ等)を上下ランクに分け。
- ・マーケット平均の収益率を表示。
- ・選択した銘柄の収益率を表示。

### 7. シミュレーション (株価)

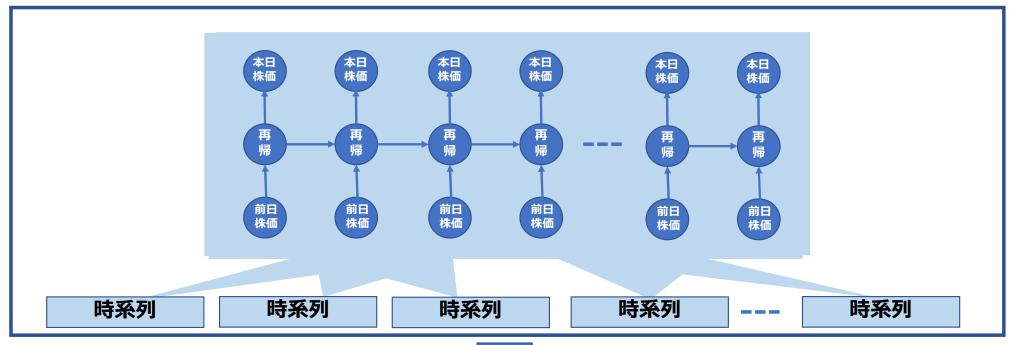
過去の株価(終値)を時系列データとして、学習用、検証用に分割し学習、検証を行い分析モデルを作成する。 その分析モデルを用いて、未来の株価シミュレーションを行う。

ブラック・ショールズモデル  $dSt = \mu t*St + at*St*Wt$  Mtとatを重みとし更新する。T分布からWt抽出



### 8. 学習と推測

時系列データの学習のため、再帰型ニューラルネットワーク(RNN)モデル(LSTM)を使用する。





時系列データを母集団からランダムに何度も学習させることにより、重みを最適に更新する。

ブラック・ショールズモデル  $dSt = \mu t * St + a t * St * Wt$ 

μtとatの重みを更新する。T分布からWt抽出

2021/5/14 **11** 

# 9. 機能および技術要素

NO	機能	技術要素	用途	デモ
1		COL ita?	・データベース構築 ・CSVから差分データ追加	実装
1	データ管理	SQLite3	・CSVから左がナータ追加 ・データベース読込	実装  実装
		D. O. com ( colonium	・HTML読込	実装
2	銘柄情報表示 	PyQuery, selenium	・JavaScript読込	実装
3	投資指標シミュレーション	一部線形回帰分析	・マーケットベータ	実装
4	株価シミュレーション	TensorFlow、RNN	・時系列データの学習・予測	検証済
5	シミュレーション結果Excel出力	openpyexcel	・銘柄情報・分析結果の出力	未検証
			・メニュー画面表示	実装
	GUI	PySimpleGui	・リストボックス ・銘柄情報表示 ・シミュレーショングラフ表示	実装 
6				実装
				実装
			・pyファイルの実行	実装

**2021**/5/14 **12**