

株価予測におけるリカレントニューラルネットワークと Support Vector Regression の性能比較

渡邊 祐太郎[†] 平原 誠[†]

[†] 法政大学理工学部

1 はじめに

株式投資は資産運用の有効な手段の一つである。現在機械学習を用いた株価予測が活発化している [1]。機械学習には様々なモデルが存在するが、本研究ではリカレントニューラルネットワーク、Support Vector Regression を用いて株価の予測を行い、その性能を比較する。

2 研究の概要

本研究で予測する日経平均株価は、TOPIX と並び、日本株の重要な指標である。これを予測することは、株価売買をする投資家の手助けとなるほか、企業の営業方針を決める指標の一つになる。

本研究ではテクニカル分析で使用される「始値」、「安値」、「高値」、「終値」の 4 つの値を入力とする。ここで予測する値は、日経平均株価の翌日の「終値」とする。

3 本研究で使った学習アルゴリズム

3.1 リカレントニューラルネットワーク (RNN)

RNN は前の時刻の中間層の状態と次の時刻の入力とを統合することで、時系列情報を扱えるようにしたネットワーク構造となっている。時間軸展開すると多層 NN を時間方向につなげたネットワークとみなすことができる。そのため学習に誤差逆伝播法を用いることができる。この学習方法は Back Propagation Through Time (BPTT) と呼ばれる。

3.2 Support Vector Regression (SVR)

SVR は Support Vector Machine (SVM) を回帰分析に応用したモデルである。教師あり学習の一つであり、SVM と同様に高い汎化能力を持つことで知られている。

SVR では与えられた学習データを教師信号 y と出力 $h(x)$ との誤差 $(y - h(x))$ が $\pm \epsilon$ に収まるように学習を行う。このとき関数の複雑さを決める γ と学習データが誤差 $(y - h(x))$ の外に出た場合のペナルティの大きさを決める C のパラメータ調節が重要となる。

4 実験

入力「始値」、「安値」、「高値」、「終値」の 4 つの値とし、出力(教師信号)は翌日の終値とする。学習データ 400 日分(2015/1/15～2016/8/30)で学習を行い、予測データ 300 日分(2016/8/31～2016/11/16)を用いて予測を行う。

4.1 RNN による予測

中間層の出力を中間層の入力として回帰させる 3 層の

RNN を構築した。中間層のユニット数を変化させ実験を行

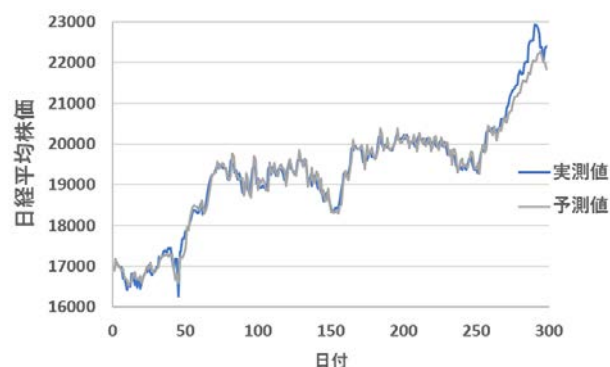


図 1 RNN での予測結果

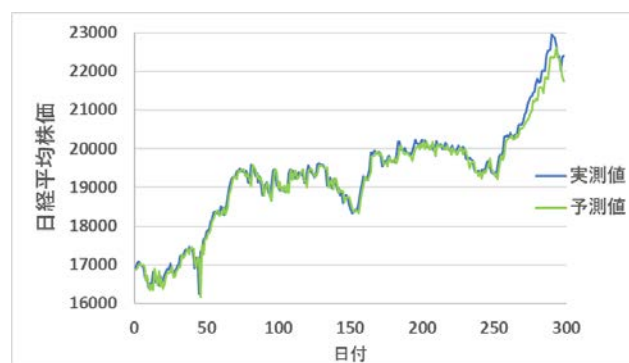


図 2 SVR での予測結果

い、もっとも予測精度の高い結果となった中間層ユニット数 10 の場合の結果を図 1 に示す。

4.2 SVR による予測

SVR では C , γ の値の調節が重要である。このパラメータを変化させ実験を行った。もっとも予測精度の高かった $C=500000$, $\gamma=0.003$ の場合の結果を図 2 に示す。

5 まとめ

SVR は株価予測においても高い精度を出すことができた。しかし予測データの 270～295 あたりの株価の急激な変化には 2 つの手法で対応できないことが分かった。このような変動にも対応できるような学習モデルの構築が必要である。

参考文献

[1] Jie Wangn, Jun Wangn: “Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks,” Neural Networks, Vol.91, pp.8-20, 2017.