

テクニカル分析を取り入れた RNN による株価変動予測

松永 大[†] 森 康久仁[†] 松葉 育雄[†]

[†] 千葉大学大学院融合科学研究科

1. はじめに

テクニカル分析とは過去の価格データから将来の価格変動を予測する手法である。しかし、予測結果が個人の経験や直感に大きく依存しているため、科学的な分析とは言い難い。

一方で、計算機能力の向上や情報通信技術の発展から計算機を利用した価格変動の予測に注目が集まっている。中でも機械学習を用いた株価変動予測に関しては積極的に研究が行われている。しかし、投資家の経験や直感に大きく依存してしまうテクニカル分析を機械学習に取り入れた研究は少ない。そこで本研究では機械学習による株価変動予測に対するテクニカル分析の導入が有効であるか検証することを目的とする。

2. 再帰型ニューラルネットワーク

ニューラルネットワークは生物の神経回路網を模倣した学習アルゴリズムであり、機械学習の一手法である。中でも本研究では、中間層のユニットが持つ情報を次の時刻に引き継ぐことによって時系列データの特徴を捉えることに特化した再帰型ニューラルネットワーク(以下 RNN)を使用した。予測に用いた RNN の中間層を 3 層重ねることで深層構造を持つネットワークを形成した。中間層のユニットの数はそれぞれ 30, 30, 10 とし 1 つ目と 2 つ目の中間層は自己回帰機能を持つユニットで構成されている。

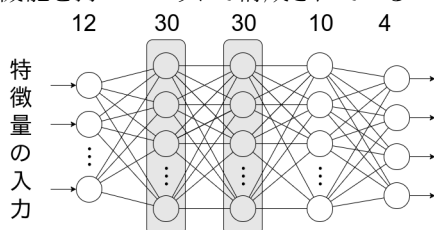


図 1. RNN の構造

3. 提案手法と使用データ

代表的なテクニカル分析としてグランビルの法則を用いて入力データを作成した。グランビルの法則とは移動平均線と現在の株価の位置関係から将来の株価の変動を予測する手法である。この手法を元に「実際の価格と移動平均線との差」と「移動平均線の状態(上昇傾向 or 横ばい or 下降傾向)」を加えた入力データを用いることで予測精度にどのような変化が生じるのかを確かめた。

本研究では東証の 2011 年度トヨタ株(1 月 17 日～7 月 29 日)を 2 分毎にサンプリングした価格データを使用する。またデータを 2 等分に分割し、前半を訓練データ、後半をテストデータとした。

学習に用いる特徴量の数は 12 である。表 1 に学習に使用した特徴量を示す。これらの特徴量を RNN の入力層へと入力する。

表 1. 使用特徴量一覧

$x_1 \sim x_{10}$	時刻 $t \sim t - 9$ の価格
x_{11}	移動平均線と時刻 t の価格の差
x_{12}	移動平均の状態

さらに、次の時刻との価格差を用いて被説明変数を設定する。価格の上昇下降だけでなく、価格が急激に変化した場合を考慮に入れて、4 つのクラスに分類した。データの価格に変動が生じなかった場合は除いて学習を行った。

表 2. 被説明変数

	次の時刻との価格差	データ数 (訓練データ)	データ数 (テストデータ)
急上昇	6 円以上	471	330
上昇	5 円以下	2023	2042
下降	-5 円以上	1952	1961
急下降	-6 円以下	496	357

株価データは上昇下降に関し、サンプル数に偏りがあるため、学習時に学習係数を調整することでサンプル数の少ないクラスの影響力を上げて学習を進めた。

4. 予測結果

テストデータを用いて RNN による株価変動予測の精度を求めた。表 3 に平均予測精度と最高予測精度を示す。

表 3. 予測精度の比較

	平均予測精度	最高予測精度
テクニカル分析あり	55.97%	59.21%
テクニカル分析なし	53.57%	57.72%

表 3 の「テクニカル分析あり」は提案手法における予測精度、「テクニカル分析なし」は単純に過去 12 時刻の価格を特徴量として同一のネットワークに入力した場合の予測精度である。RNN を用いることで株価変動を予測することができた。また、単純に過去の価格を RNN に入力するよりも、テクニカル分析を組み込むことによって予測精度が向上することが確認できた。

5. まとめ

本研究において入力層に与える特徴にテクニカル分析を組み込むことによって予測精度が向上することが確認できた。今後はその他のテクニカル分析の手法を取り入れることでさらなる精度の向上を目指したい。

参考文献

- [1] 鈴木 智也, 林 大賀, “決定論的非線形予測に基づいた金融テクニカル分析”, 電子情報通信学会論文誌 A Vol. J98-A No.2 pp.237-246, 2015/02/01