

テクニカル指標を用いた LSTM モデルによる株価予測

Prediction of Stock Price using LSTM model with Technical Indicators

19529 中作 眞仁
指導教員 佐藤 豊

1. 緒 言

最近何かと話題になっている「機械学習」と私生活で感じた「資産運用」の重要性から今回の研究は「機械学習の理解を深め、株価の予測プログラムを開発する。」ことにした。具体的には

- 作成したプログラムが長期投資と短期投資でどちらに向いているかを検証
- Keras による各パラメータの動きを検証
- それぞれの結果をもとに銘柄や景気動向に対する自分の知識と比較し、今回自分が作成したプログラムの総合評価を行う。

2. 方 法

まず、株データを `yfinance` から習得し、このデータを元に 8 つのテクニカル指標を計算する。計算後それぞれのカラムごとに正規化を行い、終値データを加えてリストにする。このリストはテストデータとトレーニングデータに分けられる。この分けられたテストデータを使用して LSTM モデルを作成する。作成後、Keras による予測を実行する。実行後、テストデータと予測結果を使用してモデルの精度評価を行う。評価を行うアルゴリズムは決定係数 (R^2_SCORE) と平均二乗誤差の平方根 ($RSME$) を使用する。最後にグラフ化をする。

これを 10 個の銘柄ごとに、長期投資と短期投資のケースに分けて検証する。また、Keras の各パラメータを変更した場合の検証も行う。

3. 結 果

以下の図 1 と図 2 は銘柄 3M の長期データ (2021/1/1~2023/1/1) を使用し、ウィンドウサイズの検証をした際の結果である。左の図は実際の価格と予測価格のグラフを示しており、右の図はその散布図を示している。この散布図は黒色の直線に予測プロットが集まっていれば予測精度は高いとされ、逆に散らばっていると予測精度は低いとされる。さらに、精度の良し悪しは決定係数 (R^2_SCORE) と平均二乗誤差の平方根 ($RSME$) で数値化している。決定係数 (R^2_SCORE) は 1 に近ければ予測精度が高いとされ、逆に平均二乗誤差の平方根 ($RSME$) は 1 に近ければ予測精度が低いとされる。図 1 ではウィンドウサイズを 60 に設定しており、比較的予測精度が良い結果が出ている。図 2 ではウィンドウサイズを 40 に設定しており、予測精度は低い結果が出ている。

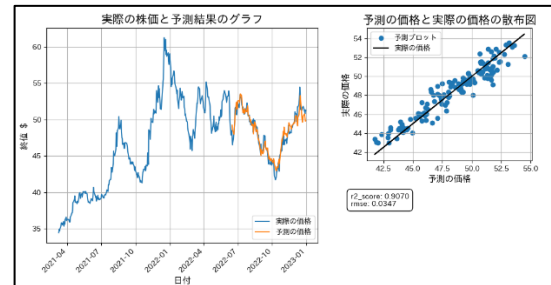


図 1 予測精度が高い場合の予測結果と散布図

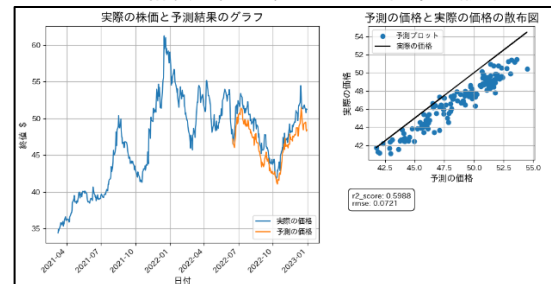


図 2 予測精度が低い場合の予測結果と散布図

4. 結 言

今回の研究では長期投資のほうが、非常に高い精度を示した。これは学習するデータの量が関係していると考えられる。

本研究では Keras の様々なパラメータを検証したが本稿ではウィンドウサイズを検証した結果のみを示しているため、ウィンドウサイズについて書き記す。ウィンドウサイズにはこれといった固定値はなく、銘柄ごとに最適解は異なると考えられる。

機械学習全体に対する自分の評価は、銘柄を限定する必要はあるが、各パラメータを調整することによって長期投資では有用であると考えられる。しかし、コロナのような世界的で大きな変動があった際には、予測は大きく外れると考えられる。

5. 今後の予定

今回の機械学習では思いつくテクニカル指標すべてを入力値とした。これらのテクニカル指標すべてが相乗効果をもたらしているとは考えづらい。そこで今後の課題として有用性の高いテクニカル指標を選定することがあげられる。

文 献

- [1] 片寄諒亮、吉岡真治 (2020)、「機械学習によるテクニカル分析の影響の調査」、『人工知能学会研究資料』、FIN-024
- [2] 松本健、牧本直樹 (2019)、「LSTM による時系列予測と株式投資戦略への応用」、『人工知能学会研究資料』、FIN-022