

株価時系列画像を用いた畳み込みニューラルネットワークによる株価予測 Stock Price Prediction Using Convolutional Neural Network with Stock Price Time Series Image

遠藤 博人[†] 穴田 一[†]
Hiroto Endo Hajime Anada

1. はじめに

近年、人工知能技術の発展が目覚ましく、様々な分野での適用が進んでいる。金融分野においても人工知能の適用が進んでおり、株価予測[1][2][3]や投資戦略[4]に関する研究が盛んに行われている。

宮崎らは、東証株価指数(TOPIX)と TOPIX 主要構成銘柄である Topix Core30 の株価 100 ステップ分(1 ステップ: 30 分)を畳み込みニューラルネットワーク(CNN)の入力として学習を行い、30 分後の株価の上昇下落を正解率 52% で予測することを確認している[1]。株価をそのまま数値データとして用いる利点としては、画像データよりも入力データのサイズが減ることにより少ない計算量で済む点や余計なノイズが入らない点などが挙げられる。また、池田は日経平均株価の 1 分足データを用いて 20 分間分のローソク足画像を作成し、その画像データを CNN の入力として学習を行い、3 分後の株価の上昇下落を予測している[2]。その結果、正解率約 60% の結果を出すことを確認している。株価を画像データに変換してから用いる利点としては、実際に人間が行っているようにチャートの形状を用いてトレンドを把握することで、今後の動向を予測するといった行動が学習できる点が考えられる。しかし、数値データでも同様の学習が行われているのではないかと考えられ、どちらのアプローチが有効か検証されていない。

我々は、数値データと画像データの有効性を確認するために、終値とローソク足のデータを用いて比較実験を行った。その結果、終値とローソク足ともに画像データの方が良い予測精度を出すことを確認した[5]。しかし、画像データでの正解率は 55% ほどであり、決して高い精度で予測ができていたとは言えない。

本研究では、チャート画像に表示する情報量を増やすことで、性能がどのように変化するかを確認していく。

2. 関連研究

宮崎らは、東証株価指数(TOPIX)と TOPIX 主要構成銘柄である Topix Core30 の株価 100 ステップ分(1 ステップ: 30 分)を CNN の入力データとして学習を行い、30 分後の株価の上昇下落を予測する研究を行った[1]。その結果、約 52% の正解率を出すことができおり、時系列データを CNN の入力データとして用いる有効性を示している。

池田は、日経平均株価の 1 分足データを用いて 20 分間分のローソク足画像を作成し、その画像データを CNN の入力として学習を行い、3 分後の株価の上昇下落を予測している[2]。その結果、約 60% の正解率を出すことに成功している。

以上のように数値データと画像データを用いた研究が行われているが、どちらのアプローチが有効か検証されていなかった。そこで我々は、数値データと画像データの有効性を確認するために、終値とローソク足のデータを用いて比較実験を行った。予備実験の結果が良かったため、数値データは 200 日分、画像データは 60 日分、それぞれ CNN の入力として学習を行い、5 営業日後の株価の上昇下落を予測した。その結果、終値とローソク足ともに画像データの方が良い予測精度を出すことを確認した[5]。しかし、画像データでの正解率は 55% ほどであり、決して高い精度で予測ができていたとは言えない。

3. 提案手法

我々は、画像データを用いる利点として、人間と同様にチャートの形状からトレンドを予測するといった行動が学習できると考えている。また、CNN 自体が人間の視覚野をモデルにした手法であるため、人間が見るデータ(時系列画像)を入力とする方が適している可能性がある。そこで本研究では、より人間が見ているような時系列画像に近づけるため、チャートに表示する情報量(テクニカル指標)を増やした画像を用いて株価の上昇下落の予測を行う。

提案手法では、終値とローソク足のそれぞれの画像データに代表的なテクニカル指標である移動平均線加えた画像を用いる。移動平均線とは、一定期間の株価の動きを平均化した値を線で結んだものである。これを用いることで、トレンドの方向感や売買のタイミングを探ることができる。また、2 本の異なる移動平均線を用いることで、株価の上昇傾向を示すゴールデンクロスや下落傾向を示すデッドクロスなどのサインを発見できる。本研究では、短期間での予測に用いられる 5 日移動平均線と 25 日移動平均線を用いた。

4. 評価実験

4.1 使用するデータ

本研究では、終値とローソク足を用いた画像とそれぞれに移動平均線を追加した 4 つの画像で性能比較を行う。使用するデータは、日経平均株価の 2004 年 1 月 1 日~2021 年 12 月 31 日とし、2004 年 1 月 1 日~2017 年 12 月 31 日を学

[†] 東京都市大学大学院 総合理工学研究科
Graduate School of Integrative Science and Engineering,
Tokyo City University

習期間, 2018 年 1 月 1 日~2021 年 12 月 31 日をテスト期間として用いた。

評価実験に使用する画像データは, 終値を表示した時系列画像とローソク足を表示した時系列画像, 終値と移動平均線を表示した時系列画像, ローソク足と移動平均線を表示した時系列画像の 4 種類を用いた(図 1, 2)。

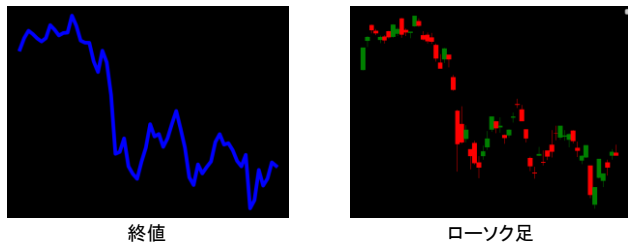


図 1 使用するデータ①

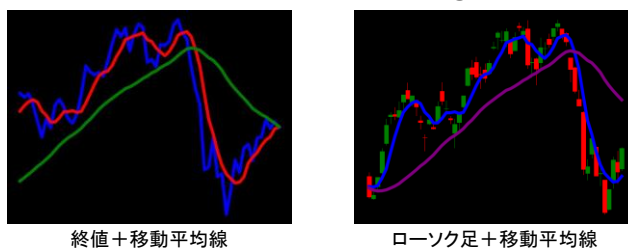


図 2 使用するデータ②

予備実験の結果が良かったため, 時系列画像に表示する日数は 60 日, 画像の入力サイズは 128 ピクセル×128 ピクセルとした。

4.2 使用する手法

評価実験で使用する手法は, 人工知能技術を支えている深層学習の中でも画像データに強みを持つ CNN を用いた。予測対象日については, 1 日後, 5 日後, 10 日後と間隔を空けて実験を行い, 最も精度の良かった 5 日後と設定した。したがって評価実験では, 図 1 及び図 2 のデータを CNN の入力とし, 5 日後の株価の上昇下落の予測を行った。

5. 実験結果

評価指標には, 正解率(Accuracy), 再現率(Recall), 適合率(Precision), F 値(F-measure)を用いた。正解率は, 全データに対してどれだけ正解できたかの割合を表す。再現率は, 実際に株価が上昇したうち, どの程度予測できたかの割合を表す。適合率は, 株価が上昇すると予測したうち実際に上昇した割合を表す。再現率と適合率はトレードオフな関係であるため, 両者をバランス良く評価できる F 値も用いる。図 1 及び図 2 を CNN の入力とし, 5 営業日後の株価の上昇下落を予測した実験結果を以下の表 2 に示す。

表 1 実験結果

データ	正解率	再現率	適合率	F 値
終値	56.04%	68.31%	59.22%	63.44%
ローソク足	55.16%	62.20%	59.40%	60.77%
終値+移動平均線	58.13%	68.90%	61.08%	64.75%
ローソク足+移動平均線	56.37%	68.90%	59.42%	63.81%

終値と終値に移動平均線を加えた際の結果を比較すると, 終値のみの正解率は 56.04%, 終値に移動平均線を加えた場合の正解率は 58.13% となり, 移動平均線を加えた方が高い性能を出している。また, ローソク足とローソク足に移動平均線を加えた際も同様の結果を確認できた。再現率や適合率, F 値といった指標についても移動平均線を加えた方が良い結果を出している。さらに, 終値とローソク足に移動平均線を加えた場合を比較すると, 終値に移動平均線を加えた画像の方が良い結果を出すことも確認できた。

6. おわりに

本研究では, 時系列画像を用いた株価予測の性能向上を目指し, チャート画像に表示する情報量を増やすことで, 性能がどのように変化するかを確認した。実験結果から, 終値やローソク足だけを用いるよりも移動平均線を加えた方が良い結果を出すことが確認できた。これは, チャートに表示されている情報量が増加したことによって CNN が上昇下落のパターンをより学習しやすくなり, 性能向上に繋がったと考えられる。さらに, ローソク足に移動平均線を加えるよりも, 終値に移動平均線を加える方が良い結果を出すことが確認できた。これは, 終値や移動平均線のように値を線で結んだような形状の方が比較的シンプルであるため, CNN が特徴を捉えやすかったのではないかと考えられる。

提案手法では, 時系列画像から特徴的なチャート形状を学習できたことが性能向上に繋がったと考えている。しかし, 深層学習モデルはブラックボックスであるため, 何を学習しているのかは定かではない。今後は, CNN が何を学習しているのかを解析することで, 画像データを用いる有効性の源泉を確かめていきたい。

参考文献

- [1] 宮崎邦洋, 松尾豊. 深層学習を用いた株価予測. 第 31 回人工知能学会全国大会. 2017, 2D3-OS-19a-3.
- [2] 池田欽一. 株価ローソク足チャート画像を用いた畳み込みニューラルネットワークによる株価変動予測. 北九州市立大学『商経論集』. 2019, Vol. 54, No. 1,2,3,4. p. 1-18.
- [3] 松井藤五郎, 和泉潔. 新聞記事の時系列テキスト分析による株式市場の動向予測. 第 30 回人工知能学会全国大会. 2016, 3L3-OS-16a-6.
- [4] 松井藤五郎, 後藤卓, 和泉潔, 陳ユ. 複利型強化学習における投資比率の最適化. 人工知能学会論文誌. 2013, vol. 28, no. 3, p. 267-272.
- [5] 遠藤博人, 穴田一. 株価時系列画像を用いた株価予測の有効性検証. 情報処理学会研究報告. 2022, vol. 2022-MPS-138, no. 35, p. 1-2.