

# 深層学習を用いた日経平均株価の予測手法の改良～畳み込みの導入～ Improvement for the Forecasting Method of the Nikkei Stock Average by Using Deep Learning ~ Adopting for the CNN~

1841073 関口 晴紀

Haruki SEKIGUCHI

指導教員 秋葉 知昭

Deep Learning techniques are rapidly being developed in many fields, for example, stock price forecasting. In this study, I considered Nikkei Stock Average forecast using deep learning with CNN. From this I improved forecast idea of the Nikkei Stock Average forecast.

## 1. 緒言

日経平均株価は日本の代表的な指標であり、日本経済を表している。新聞やニュースで値動きが発信され、さらには投資の基準や水準になっており、重要性の高さが伺える。本研究は将来の日本経済を予測するために、東京証券取引所第一部に上場している 225 銘柄から算出される日経平均株価終値を予測する方法を提案する[1]。

本研究では株価や指標の予測精度向上の要因の一つである LSTM を用いた石毛の研究[1]に新たにデータと最適化関数を追加し、さらには畳み込みを導入して予測手法の改良を行い、最適化関数 5 つに加え、畳み込みを導入し、結果を比較する。

## 2. 先行研究

### 2.1 検知手法の概要

石毛の研究[2]では、LSTM を使用し、1 日の終値 50 日分のデータを入力し 51 日目の終値を予測した。最適化関数は SGD, RMSprop, Adam, RAdam を用いた結果を比較し考察した。

### 2.2 収集データ

石毛の研究[2]では SBI 証券が提供している取引ツールの HYPER SBI からダウンロードできる 1991 年から 2019 年までの 28 年間の 7123 日分の使用する終値のデータを収集した。

### 2.3 パラメータ及びモジュールの設定

最適化関数は SGD, RMSprop, Adam, RAdam の 4 つを用いた。誤差関数は MSELoss を使用し、学習率は、SGD のみ 0.1、他の 3 つは 0.01 とした。epoch は 2000 と 3000 で、バッチサイズは 10 とした。

### 2.4 出力データ

出力データは epoch, loss, 正解率, training\_accuracy, test\_accuracy を出力した。loss のグラフにおいて SGD は学習曲線が見られず、他 3 つの最適化関数は学習曲線が確認でき、最も高い正解率で Adam

と Radam の 7% で、正解条件は誤差 10 円未満だった。

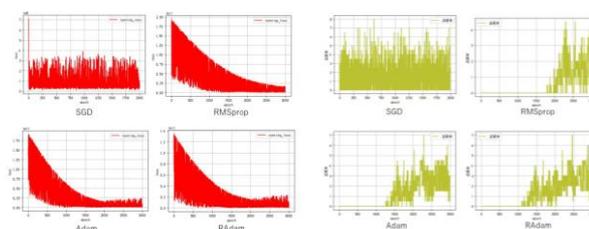


図 1 各最適化関数における loss と正解率の推移

## 3. 検知手法の提案

### 3.1 先行研究の問題点と改善

石毛の研究[2]の問題点は SGD がうまく学習できなかったことと、正解率が最も高く 7% だったことにある。

本研究では、新たに利用データの追加と最適化関数の追加及び畳み込みの導入を行う。

### 3.2 検知手法の概要

本研究では、LSTM を使用し、1 日の予測終値 10 日分のデータを入力し 11 日目の終値を予測する。最適化関数は SGD, RMSprop, Adam, Radam, Adadelta を用いた結果を比較し考察する。また畳み込みは 11 個目の 1 つの値を元データに追加していき、新たにデータセットを行い、トレーニングを二回行う。

### 3.3 収集データ

本研究では SBI 証券が提供している取引ツールの HYPER SBI からダウンロードできる 2001 年から 2021 年までの 20 年間の 4899 日分の日経平均株価・TOPIX・為替対円米ドル・日経平均株価先物のそれぞれ始値、高値、安値、終値を収集し、①日経平均株価、②日経平均株価+TOPIX、③日経平均株価+為替対円米ドル、④日経平均株価+日経平均株価先物の 4 つの組み合わせで比較した。

### 3.4 パラメータ及びモジュールの設定

最適化関数は SGD, RMSprop, Adam, RAdam に加え単位を合わせてくれる Adadelta の 5 つで比較する。誤差関数は MSELoss を使用した。学習率は SGD が 0.1, Adadelta は 1 から 1.5 他の 3 つは 0.01 から 0.05 をそれぞれ初期値に設定した。epoch は 500 に設定した。バッチサイズは 10 とした。

## 4. 結果及び考察

### 4.1 出力データ

出力データは epoch, loss, 正解率, training\_accuracy, test\_accuracy を出力し, 左上に利用データを示し, 上図は通常の, 下図は畳み込みを導入したグラフである。

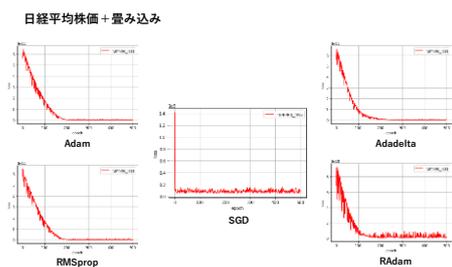
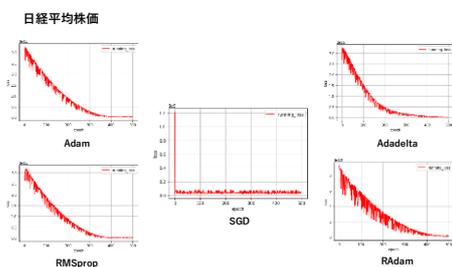


図 2 各最適化関数における loss の推移

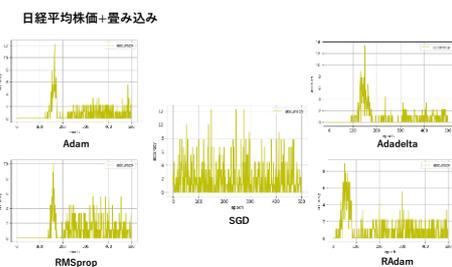
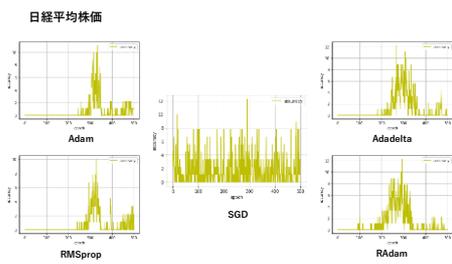


図 3 各最適化関数における正解率の推移

### 4.2 最適化関数の比較

学習に使う損失を数値化した loss を比較する。SGD では畳み込み導入による変化はなく, loss の学習曲線が見られなかった。他 4 つにおいては学習のスピードが速くなるのが確認できた。また畳み込みの導入で Adadelta, RAdam の TOPIX, 為替対円米ドルを追加した場合に分散傾向にあったが, 学習曲線が改善していることが分かった。

### 4.3 考察

本研究の提案手法では, SGD を使用した場合うまく学習ができず, その他 4 つの手法は学習ができた。SGD は極小値から抜け出せていないと考えられる。しかし他の RMSprop, Adam, RAdam, Adadelta は極小値から抜け出し, 学習をしている様子が見られる。本研究で最も高い正解は畳み込みを導入した Adadelta の 13% となった。学習の後半で 1% から 4% の間で落ち着いているところを見ると十分に予測できているとはとても言えない。最低でも 50% の正解率が必要であると考えられる。しかし畳み込みの導入で学習速度が速くなることは確認できた。

## 5. 結言

本報告では Pytorch を用いた日経平均株価の予測を行った。正解条件を誤差 10 円以内とし, 最も良い結果は Adadelta の 13% となった。SGD は学習がうまくできなかつたが, 畳み込みは学習速度向上と制度の微増という結果となった。

最も高い正解率が 13% ではうまく予測できているとは言えないため前述のような予測の正解条件を変更するなどの根本的な変更をすることによって十分な精度を得る必要がある。

これらの結果から LSTM で Adam, RAdam, RMSprop, Adadelta を用いた日経平均株価の予測はかろうじてできていることが確認された。よって予測の難しいと言われる株価の予測には LSTM と Adam, RAdam, RMSprop, Adadelta を用いた手法が有効であると言える。本研究では Adadelta の有用性と畳み込みの精度向上が確認できたが一次元データしか読み込めない弱点が分かった。今後は更新によってさらに精度の高い予測が可能になると考える。

## 文献

- [1] SBI 証券関東財務局長(金融)第 44 号: HYPER SBI
- [2] 石毛康太郎: 深層学習を用いた日経平均株価の予測手法の提案, 2020 年度千葉工業大学卒業研究(2020)