

## 深層学習による株価予測に関する実験的考察

### Experimental Study on Stock Price Prediction using Deep Learning

小松 隆行\*

Takayuki Komatsu

#### 概要

本稿では、時系列データを学習するニューラルネットワークの RNN (リカレントネットワーク) の発展形である LSTM (Long Short-Term Memory) を用いたニューラルネットワークを、深層学習のアルゴリズムを用いて学習し、株価を予測する計算機シミュレーションを行い、実験的な考察を行ったことを報告する。訓練データとして用いたのは、過去 18 年間の日経平均株価のデータであり、学習のための訓練データの営業日の期間、株価を予測する営業日の期間、訓練データ数、訓練データ系列の長さ、LSTM のユニット数などの条件を変えながら多数のシミュレーションを行い、それぞれの予測における平均絶対誤差率 (MAPE) によって評価した。ネットワークの学習曲線の特徴についても考察する。

#### 1. はじめに

近年、コンピュータプログラムによる株の自動取引、いわゆるシステムトレードが現実のマーケットにおいてその割合を増加させている。米国の株取引においては、全体の 7 割がシステムトレードによって行われ、日本のでも 5 割を超えているといわれている。フラッシュクラッシュなどの急激な株価の変化 (暴落) の原因が、このようなコンピュータプログラムにあるとも言われている。さらに、人工知能の急速な進歩により、ビッグデータを利用したアルゴリズムが発達し実用化されてきている。特に、強力な学習方式の深層学習 (Deep Learning) <sup>(1)(2)(3)</sup> を用いたものが出現している。証券会社から人工知能による株取引の商品も増えて続けているが、その詳しい仕組みやアルゴリズムは公開されていないことが多い。世界中の twitter のツイートの内容を参照して、特定の用語によって株価の騰落が起きるとして自動取引アルゴリズムに反映させる研究と事例もある。

株価予測は、時系列分析と予測の研究のアプリケーションとして、これまで多く研究されてきた。ARIMA モデルや、GARCH モデルによって、精度の高いモデリングと予測が可能であることが報告されている。数年前から世界を席捲している深層学習のアルゴリズムは、当初画像認識の分野や時系列データ

分析と予測でのブレークスルーとなり、年々その応用範囲を拡大している。ノーフリーランチ定理<sup>(4)</sup>が示すように、適切なアルゴリズムを適用する前に、適用する問題領域についての知識や知見を十分理解し、それらを可能な限り利用することを考える必要があると考えられる。したがって、株価予測のために、深層学習のアルゴリズムを従来の株価データの解析手法や分析手法を用いず、株価データを入力データと出力データとして使用し学習させることは、それに反するかもしれない。しかしながら、従来の知識や知見を用いず、この強力な手法を如何に適用すれば有用なモデリングと予測が可能となるかを調べて知見を得ることは、大きな意義があると考えられる。深層学習での成功事例は、問題領域の知識や知見をまったく持たない場合でも絶大な成果をあげているからである。このようなことから、深層学習を用いた株価予測のシミュレーションを行い、数値的な検証を行うことは意義あることと考えられる。

本報告では、日経平均株価を訓練データとして用いて、RNN (Recurrent Neural Network) <sup>(5)(6)</sup> の発展形であり、長い時系列データの学習を可能とし非常に性能の良い LSTM (Long Short-Term Memory) <sup>(7)</sup> を用いたニューラルネットワークを、深層学習のアルゴリズムで学習させた後に、株価予測を行う。株取引

\* 北海道科学大学未来デザイン学部メディアデザイン学科

には、短期的トレード（当日から数週間）、中期的トレード（1か月から2年）、長期的トレード（2年以上）などがあり、保有期間によって売買方針が異なることが一般的である。したがって、どのような方針で取引するかによって予測モデルの構築方針や方法も異なると考えられるが、本報告では短期的な予測を前提に予測モデルを計算機シミュレーションで学習して構築し、株価予測を行い検証して考察を試みる。

## 2. これまでの関連する研究や事例

これまで、深層学習を用いた株価予測に関する研究は多数行われてきている。Takeuchiらは、積層RBMからなるディープラーニングモデルを使って、過去の株価データのみから総利益を大幅に増やしている<sup>(8)</sup>。Chenらは、中国の株式市場の株収益予測にLSTMが有効であることを示している<sup>(9)</sup>。Xiongらは、LSTMネットワークでS&P500のボラティリティを予測するモデルを構築し、Google Domestic Trendという外部情報を追加入力として予測精度（平均絶対誤差率）を改善させた<sup>(10)</sup>。Dingらは、ニュース記事からイベントを利用し、長期と短期の影響を考慮したCNNでS&P500の株価予測の精度を改善させている<sup>(11)</sup>。Jiaは、株価予測において、様々なサイズのLSTMを1層から3層の中間層として用いて、その有効性を示している<sup>(12)</sup>。Persioらは、LSTMをはじめとする様々な構造のモデルでS&P500などの株式市場の株価予測を試みている<sup>(13)</sup>。Akitaらは、新聞記事を分散表現ベクトルにし、過去のイベントの影響をLSTMによってモデリングし東証の株価で検証している<sup>(14)</sup>。宮崎らは、CNNを株価（日経平均など）予測に応用し予測力の改善とCNNの可能性を示した<sup>(15)</sup>。田村らは、テクニカル分析の機械学習手法とファンダメンタル分析の機械学習のアンサンブル学習モデルで株主価値の推定モデルを提案している<sup>(16)</sup>。また、人工知能（AI）による株価予測モデルを用いた銘柄情報の提供を開始した事例もある<sup>(17)</sup>。これは過去の決算パターン毎の決算発表後の値動きを学習し、財務情報を入力として緩やかに上昇する銘柄を予測するものである。

## 3. 計算機シミュレーションの目的

本報告では、日足の株価（始値、終値）を訓練データとし、中間層としてLSTMを用い、入力層と出力層を持つニューラルネットワークを、様々な条件を

変えて深層学習によって学習させたモデルによって株価予測を行い、予測誤差をMAPE（平均絶対誤差率）で評価し、その結果を考察することを目的とする。

## 4. 計算機シミュレーションの内容

### 4.1 訓練データ

訓練データとして使用する株価のデータは日経平均株価の過去のデータである。今回は、2000年3月13日から2018年3月19日までの日経平均株価データを、Yahoo Finance<sup>(18)</sup>からダウンロードし取得し使用した。これらの株価のうち、日足の始値と終値を用いる。使用した株価のデータの一部をチャートで図1に示す。

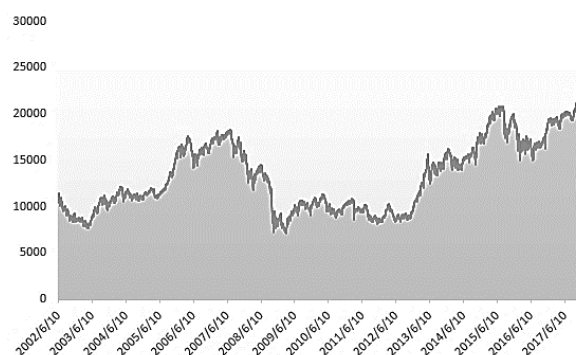


図1 使用した日経平均株価のチャートの一部

学習データとは、ネットワークの入力層から入力する株価のデータである。学習データは、連続する株価の値であり、ある  $t$  番目の営業日の株価（始値または終値）を  $x_t$  ( $t=1, 2, 3, \dots$ ) と表し、長さ  $L$  の連続したそれらの系列を

$$X_t = \{ x_t, x_{t+1}, x_{t+2}, \dots, x_{t+L-1} \}$$

と表す。また、これらを使って予測する、ある  $t$  番目の営業日の株価を  $y_t$  ( $t=1, 2, 3, \dots$ ) と表し、長さ  $L$  の連続したそれらの系列を、

$$Y_t = \{ y_t, y_{t+1}, y_{t+2}, \dots, y_{t+L-1} \}$$

と表す。さらに、それぞれの連続する株価の系列の長さ  $L$  をそれぞれ、

$$L_{\text{train}} = |X_t|, \quad L_{\text{pred}} = |Y_t|$$

と表す。ここで  $|\cdot|$  は集合の要素数を表す。

また、 $N$  個の連続した  $X_t$  の系列の組と、 $Y_t$  の系列の組を

$$X_t = \{X_t, X_{t+1}, X_{t+2}, \dots, X_{t+N-1}\}$$

$$Y_t = \{Y_t, Y_{t+1}, Y_{t+2}, \dots, Y_{t+N-1}\}$$

と表す。すなわち、 $X_t$ の連続する各要素は、株価の長い系列において、ある  $t$  番目の営業日時点から 1 営業日ずつスライドしたデータ系列であり、 $X_t$  はそれらの集合である。 $Y_t$  も同様である。

またここで、 $X_t$  と  $Y_t$  の要素数  $N$  をそれぞれ次のように表す。

$$N_{\text{train}} = |X_t|, \quad N_{\text{pred}} = |Y_t|$$

本報告での学習と予測のシミュレーションは、ローリング方式で行う。具体的には、2000 年から 2018 年までのある特定日の株価を予測する場合、学習データの組として、その日の前日から遡って、 $L_{\text{train}}$  日間の連続する株価系列  $X_t$  を  $N_{\text{train}}$  個用意し  $X_t$  を構成する。以下では、 $Y_{t+1} = X_{t+1}$  として、株価系列  $X_t$  から  $Y_{t+1}$  を予測する。 $X_t$  と  $Y_{t+1}$  の組  $(X_t, Y_{t+1})$  を訓練データと呼ぶ。また、以下のようにする。

$$L_{\text{train}} = L_{\text{pred}} \in \{10, 25, 52, 102\}$$

$$N_{\text{train}} = N_{\text{pred}} \in \{200, 300, 500\}$$

なお、学習と評価および予測においては、始値予測のネットワークは、始値の系列を学習データとして入力して学習させた後に、始値の予測に用い、終値予測のネットワークは、終値の系列の学習データとして入力して学習させ、終値の予測に用いる。

#### 4.2 深層学習のネットワークモデル

本報告で使用するネットワークモデルは、入力層、中間層、出力層から構成されるニューラルネットワークモデルであり、中間層が LSTM から構成されるものである。LSTM 内部には、入力ゲート、出力ゲート、および忘却ゲートなどがあり、RNN での短所である勾配消失問題を解決し、長期依存がうまく扱えない問題を緩和できる長所があり、時系列データの予測モデルとして非常に有用であることが知られている。中間層ではこのような構造の一つの LSTM のメモリユニットと呼び、その数をブロック数と呼ぶことにする。本報告では、中間層に LSTM のブロックを複数個使ったネットワークを用いる。LSTM のブロック数を  $s$  と表す。本報告では、

$$s \in \{52, 102, 152, 202\}$$

とする。

$N = N_{\text{train}}$  として、 $N$  個の訓練データの組が、 $((X_t, Y_{t+1}), (X_{t+1}, Y_{t+2}), (X_{t+2}, Y_{t+3}), \dots, (X_{t+N-1}, Y_{t+N}))$  として選ばれたとき、入力層から株価データ  $X_t$  を与え、出力層で  $Y_{t+1}$  を与えて、誤差関数として二乗誤差関数を用いてミニバッチ方式の勾配降下法を用いて二乗誤差が微小になるまで学習させる。言い換えると、ある  $t$  番目の営業日から長さ  $L_{\text{train}}$  の長さの連続する株価の系列から、 $t+1$  番目の営業日の株価から  $L_{\text{train}}$  の長さの連続する株価系列を予測するように学習させる。本報告では、LSTM ブロック数  $s$  の値も変えて、深層学習のアルゴリズムで学習後に予測を行い、その結果を考察することにする。株価の予測には、訓練データの最後のデータ  $X_{t+N-1}$  の先頭営業日の翌日からの連続した株価系列  $X_{t+N}, X_{t+N+1}, X_{t+N+2}, X_{t+N+3}, X_{t+N+4}$  の 5 個をそれぞれ入力としたとき出力される系列の最後の営業日の株価を予測値とする。

#### 4.3 平均絶対誤差率

学習終了後のネットワークを用いて、株価の予測を行った結果の評価には、以下の式で定義される平均絶対誤差率 (MAPE : Mean Absolute Percentage Error) を用いる。

$$\text{MAPE} = \frac{100}{n} \sum_{k=1}^n \left| \frac{y - x}{x} \right|$$

ここで  $n$  は予測されたデータ数、 $x$  は真の値、 $y$  は予測値である。

#### 4.4 計算機シミュレーションの内容と目的

コンピュータによる実際の学習と評価のシミュレーションでは、全期間 (全訓練データ) の株価の最大値で各株価を割り、株価の値を 0 から 1 の間の実数値になるようにして使用している。シミュレーションで使用した OS は Windows10Pro, CPU は Intel 製® Core™ i7-7700HQ (2.80GHz), RAM は 32GB, GPU は NVIDIA 製 GeForce GTX 1070 である。以下に、今回行ったシミュレーションをそれぞれ「実験 1~3」と呼び、それらの内容と目的について説明する。なお、全ての実験は、エポック数 2000, ミニバッチサイズ 5, パラメータ更新のための学習係数決定方式は Adam, 学習係数初期値は 0.001, LSTM レイヤ数 (中間層の数) は 1 個で行った。

実験 1) 訓練データ数  $N_{train}$  を変えた場合: 学習データ  $X_t$  の長さ  $L_{train}$  を 102, LSTM のブロック数 102 個にして, 予測期間毎に  $N_{train}$  を 200, 300, 500 と変えながら学習させて予測し, 訓練データ数による違いを調べる.

実験 2) 学習データ  $X_t$  の長さ  $L_{train}$  を変えた場合: LSTM のブロック数を 102 個とし,  $N_{train}$  を 500 にして, 一つの学習データにおける連続する株価データの数期間の長さ  $L_{train}$  を変えながら, それぞれ学習させて予測し, 学習データ長による違いを調べる.

実験 3) LSTM のブロック数  $s$  を変えた場合:  $N_{train}$  を 500, 学習データ  $X_t$  の長さ  $L_{train}$  を 102 にして, LSTM のブロック数を変えながら, それぞれ学習させて予測し, ブロック数の影響を調べることを目的とする. LSTM のブロック数  $s$  を 52 個, 102 個, 152 個, 202 個の 4 通りで実施した.

## 5. 結果と考察

以下の図 2 に本稿でのシミュレーションにおける株価予測の例を示す. 横軸に株価を予測する営業日を取り, 予測の起点日を 0 とし, その起点日以前の  $L_{train}$  の長さの  $N_{train}$  個の株価データを使い学習したネットワークで, 起点日の翌日から 5 日後まで (横軸で 1 から 5 までに相当) を予測したものである.

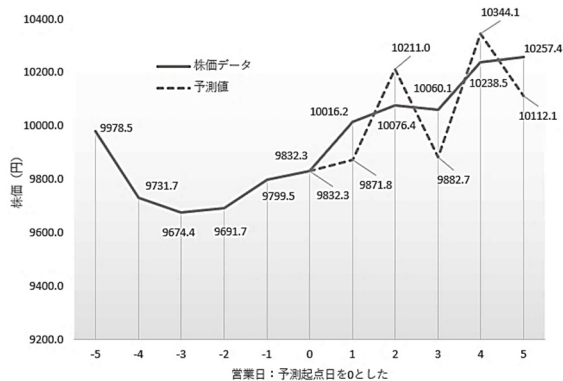


図 2 学習ネットワークによる株価予測の例

今回のシミュレーション環境において, 一つのネットワークの学習に要した時間は, 概ね 2 時間から 6 時間である.  $N_{train}$  の値や,  $L_{train}$  の値が大きいくほど大きくなる傾向にあった. なお以下の表では, それぞれの期間毎に学習評価と予測を 2 回ずつ行い, MAPE の値が概ね良いネットワークの値を記載している.

### 5.1 訓練データ数について

実験 1 では, 訓練データ数  $N_{train}$  を変えた場合のシミュレーションを行った. 予測したい異なる営業日の期間を 7 つ選んで, その予測初日の前日から遡って  $L_{train}$  が同じ値 ( $L_{train}=102$ ) で, 訓練データ数を  $N_{train}=200, 300, 500$  の 3 通りでシミュレーションを実行した. それらの結果を以下の表 1 から表 4 に示す. 表 1 と表 2 は, 予測期間が 1 日, すなわち学習データ期間の最終日の翌日のみの結果であり, 表 3 と表 4 は, 予測期間が 5 日間の結果である.

表 1 日経平均株価の始値予測の MAPE (翌日のみ)

	$N_{train}=200$	$N_{train}=300$	$N_{train}=500$
期間 1	0.08%	0.14%	0.04%
期間 2	2.07%	2.46%	2.57%
期間 3	0.90%	0.57%	0.51%
期間 4	3.24%	2.68%	2.93%
期間 5	2.13%	2.02%	1.99%
期間 6	0.37%	0.54%	0.54%
期間 7	2.06%	1.48%	1.77%

表 2 日経平均株価の終値予測の MAPE (翌日のみ)

	$N_{train}=200$	$N_{train}=300$	$N_{train}=500$
期間 1	1.54%	1.66%	1.47%
期間 2	1.15%	0.85%	0.20%
期間 3	2.43%	2.05%	1.50%
期間 4	1.60%	1.17%	1.06%
期間 5	1.21%	0.79%	1.37%
期間 6	0.46%	0.56%	0.90%
期間 7	1.78%	1.07%	1.53%

表 3 日経平均株価の始値予測の MAPE (5 日間)

	$N_{train}=200$	$N_{train}=300$	$N_{train}=500$
期間 1	1.13%	1.09%	0.74%
期間 2	1.12%	1.02%	1.08%
期間 3	1.06%	0.74%	0.61%
期間 4	1.46%	1.43%	1.43%
期間 5	1.35%	1.34%	1.46%
期間 6	0.48%	0.49%	0.51%
期間 7	1.38%	0.73%	1.08%

表 4 日経平均株価の終値予測の MAPE (5 日間)

	$N_{train}=200$	$N_{train}=300$	$N_{train}=500$
期間 1	1.43%	1.48%	1.15%
期間 2	1.04%	0.94%	0.73%

期間 3	1.32%	1.06%	0.85%
期間 4	1.76%	1.68%	1.65%
期間 5	1.41%	1.42%	1.31%
期間 6	0.47%	0.50%	0.70%
期間 7	1.37%	0.74%	1.19%

表 1 から表 4 までの全ての期間内において MAPE が最小なのは  $N_{\text{train}}=500$  の場合が最も多い (15 個)。  $N_{\text{train}}=200$  の場合が MAPE 最小なのは 5 個のみである。 また、MAPE が 1%未満のものは  $N_{\text{train}}=200$  の場合 6 個、  $N_{\text{train}}=300$  の場合 12 個、  $N_{\text{train}}=500$  の場合 11 個であり、訓練データ数が多いほうが良いものと推測される。一方で、期間 6 は全ての MAPE が 1%未満であり、最も良い結果となっているが、その中でも  $N_{\text{train}}=200$  の場合が最も MAPE が小さい。訓練データの内容によっては、そのデータ数が少ない場合のほうが良いこともあると言える。

## 5.2 時系列長 $L_{\text{train}}$ について

株価の推移は、規則性とランダム性の両方の要素を含んでいると考えられるため、学習のための連続した株価の系列の長さである  $L_{\text{train}}$  の数値を適切に指定するためには、訓練データの特徴、例えば変化量や変化率の大きさや、変動幅などを考慮して設定する必要があると考えられるが、その明確な決め方についてこれまでまだ知見は得られていない。したがって、実験 2 では、予測期間を固定し、学習データ系列長  $L_{\text{train}}$  を変えてシミュレーションを行った。学習データ系列長  $L_{\text{train}}=10, 25, 52, 102$  の 4 通りの場合について、それぞれ期間 1 から 5 において  $s=102, N_{\text{train}}=500$  として実行した。それらの結果の平均絶対誤差率の違いを以下の表 5 から表 8 に示す。

表 5 日経平均株価の始値予測の MAPE (翌日のみ)

	$L_{\text{train}}=10$	$L_{\text{train}}=25$	$L_{\text{train}}=52$	$L_{\text{train}}=102$
期間 1	0.11%	0.37%	0.11%	0.04%
期間 2	2.23%	2.26%	2.60%	2.57%
期間 3	0.54%	0.69%	0.88%	0.51%
期間 4	2.74%	2.92%	2.80%	2.93%
期間 5	2.14%	2.35%	2.27%	1.99%

表 6 日経平均株価の終値予測の MAPE (翌日のみ)

	$L_{\text{train}}=10$	$L_{\text{train}}=25$	$L_{\text{train}}=52$	$L_{\text{train}}=102$
期間 1	1.38%	1.18%	1.71%	1.47%

期間 2	0.31%	0.16%	0.18%	0.20%
期間 3	1.64%	2.02%	0.50%	1.50%
期間 4	0.92%	1.19%	1.42%	1.06%
期間 5	0.81%	1.21%	1.32%	1.37%

表 7 日経平均株価の始値予測の MAPE (5 日間)

	$L_{\text{train}}=10$	$L_{\text{train}}=25$	$L_{\text{train}}=52$	$L_{\text{train}}=102$
期間 1	0.79%	0.64%	0.70%	0.74%
期間 2	0.96%	0.97%	1.15%	1.08%
期間 3	0.68%	0.73%	0.73%	0.61%
期間 4	1.55%	1.60%	1.55%	1.43%
期間 5	1.40%	1.30%	1.29%	1.46%

表 8 日経平均株価の終値予測の MAPE (5 日間)

	$L_{\text{train}}=10$	$L_{\text{train}}=25$	$L_{\text{train}}=52$	$L_{\text{train}}=102$
期間 1	1.04%	0.98%	1.38%	1.15%
期間 2	0.78%	0.71%	0.75%	0.73%
期間 3	0.90%	1.04%	0.97%	0.85%
期間 4	1.59%	1.71%	1.60%	1.65%
期間 5	1.57%	1.36%	1.36%	1.31%

MAPE が 1%未満であるものが 3 個以上あるものは、始値においては期間 1 と期間 3、終値においては期間 2 と期間 3 である。これらの結果において、始値における期間 1 と期間 3 では 4 つの結果のうち 3 つが  $L_{\text{train}}=102$  の場合に MAPE 最少となっている。  $L_{\text{train}}$  が大きいほど計算時間が長くなる傾向であったので、計算環境に依存して現実的な計算時間で学習を終了できるような小さめの  $L_{\text{train}}$  の値でシミュレーションをした後、MAPE の値を考慮しながら、  $L_{\text{train}}$  の数値を増やしてゆくのの方がよいと考えられる。 LSTM の特徴から、  $L_{\text{train}}$  が大きいほどその能力を発揮できるものと考えられるので、  $L_{\text{train}}$  が 102 より大きい場合でのシミュレーションも必要であると考えられる。一方で  $L_{\text{train}}=102$  の場合よりも  $L_{\text{train}}=25$  のほうが MAPE が小さい場合もあることから、  $L_{\text{train}}$  の値を幅広く変えてた学習が必要であるとも考えられる。

## 5.3 LSTM のブロック数について

LSTM のブロック数の決め方に明確な知見はこれまで得られていないため、実験 3 では LSTM のブロック数を変えた場合の影響を調べるためにシミュレーションを行った。LSTM のブロック数を、  $s=52, 102, 152, 202$  の 4 通りの場合について、それぞれ

期間1から3において  $L_{train} = 102$ ,  $N_{train} = 500$  と  
して実行した。それらの結果の平均絶対誤差率の違いを以下の表9から表12に示す。

表9 日経平均株価の始値予測の MAPE (翌日のみ)

	s=52	s=102	s=152	s=202
期間1	0.04%	0.04%	0.33%	0.30%
期間2	2.50%	2.57%	2.47%	2.46%
期間3	0.66%	0.51%	0.91%	0.92%

表10 日経平均株価の終値予測の MAPE (翌日のみ)

	s=52	s=102	s=152	s=202
期間1	1.38%	1.47%	1.75%	1.94%
期間2	0.03%	0.20%	0.50%	0.28%
期間3	2.19%	1.50%	1.44%	1.67%

表11 日経平均株価の始値予測の MAPE (5日間)

	s=52	s=102	s=152	s=202
期間1	0.78%	0.74%	1.03%	1.04%
期間2	1.06%	1.08%	1.05%	1.09%
期間3	0.72%	0.61%	0.73%	0.65%

表12 日経平均株価の終値予測の MAPE (5日間)

	s=52	s=102	s=152	s=202
期間1	1.13%	1.15%	1.41%	1.66%
期間2	0.64%	0.73%	0.81%	0.76%
期間3	1.04%	0.85%	1.40%	1.04%

MAPE が1%未満であるものが3個以上のものは、LSTMのブロック数  $s$  の値にかかわらず、始値では期間1と期間3、終値では期間2となっている。期間に依存して MAPE の良し悪しが表れているようであるが、始値と終値の良し悪しが対照的である。始値の MAPE が良い傾向の期間は終値の MAPE が悪い傾向がある。各期間での MAPE の最小値は、 $s = 52$  または  $s = 102$  である場合が多い。特に、MAPE の値が1%未満の期間では、始値の場合では全て  $s = 102$  の場合が最小であり、終値の場合ではすべて  $s = 52$  の場合が最小である。終値において、 $s = 102$  (実験1) での結果よりも  $s = 52$  の場合のほうが良い結果となっている。訓練データにおいて  $L_{train} = 102$  であることから、それに関して  $s$  が102以下の場合のほうがモデリングには適していることが推測される。ブロック数が多いということは表現能力が高いと言えるが、一方で冗長である可能性もあり、

多いほうがよいと言い切れない。学習として必要十分なブロック数であることが理想であると言えるが、そのような傾向が結果に表れていると考えられる。

#### 5.4 学習曲線について

学習曲線の例を以下の図3に示す。多くはこのような順調な誤差関数の減少と収束を示す、きれいな学習曲線となった。このような場合の訓練データの例を図4に示す。

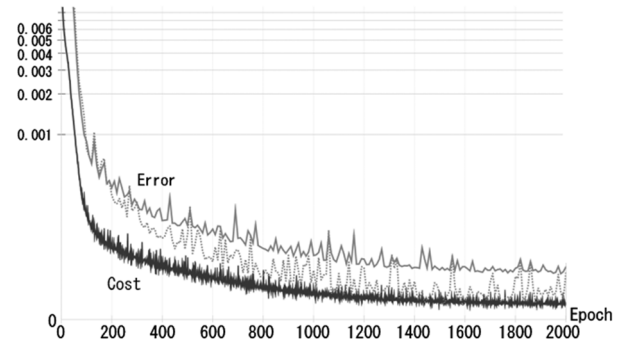


図3 順調な誤差関数減少の学習曲線の例

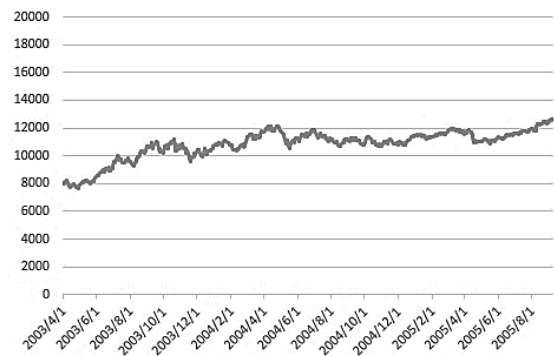


図4 順調な誤差関数減少の学習曲線の場合の学習データ(株価)の例:  $N_{train}=500$

一方、学習曲線が突然急激に変化し回復するという事象が発生することがあり、その例を以下の図5に示す。

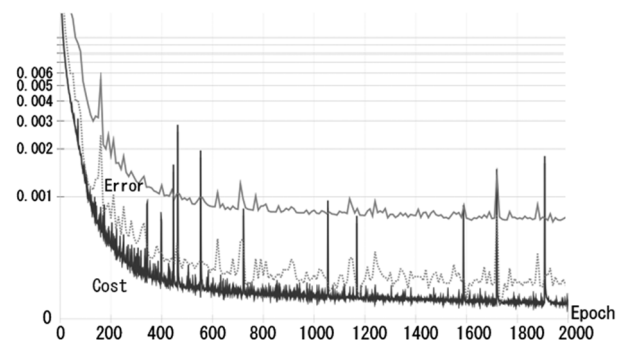


図5 学習曲線の急激な変化と回復の例

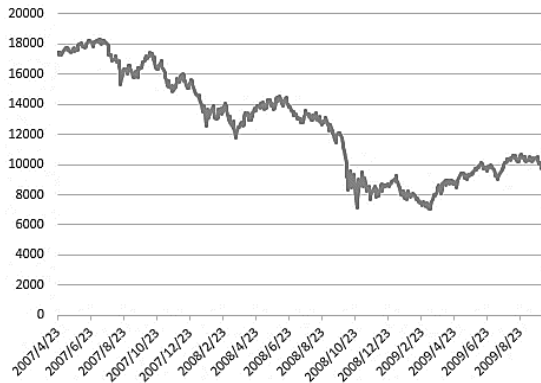


図6 学習曲線の急変化と回復があった場合の学習データの例： $N_{\text{train}}=500$

今回行った計算機実験での学習では、図3で示したような急激な変化が無いきれいな学習曲線を描く場合と、図5で示すような急激な変化と回復現象が発生する場合があった。学習における2000回のエポックにおいて、このような急激な変化と回復の現象の発生件数は、学習データによって異なる。図5の場合は約10回起きている。このような現象は、株価の急激な変化に起因する誤差関数の超曲面の急激な変化のために起こると推測されるが、学習が発散して破綻することは無かった。学習最終段階から終了時（2000エポック終了時）にかけて、このようなことが起きてても、最良の評価のパラメータの数値を保存しておいているので、予測はそれらを再現したネットワークで行っている。したがって、十分に長い学習エポック数を確保し、学習を通じて最良のネットワークを保存しておくことが必要である。

このような学習曲線になった場合の訓練データを検証すると、訓練データの中に株価が急激に大きく変化したものが含まれている場合に発生する傾向にあった。そのようなデータを図6にチャートで示した。急激な変化のない学習曲線の場合の訓練データ（図4）と比べると、急激な株価変動が多く、訓練データ内での株価の最大値と最小の差も大きい。こういったことが、誤差関数の超曲面における急激な変化になっていると考えられる。

## 6. まとめ

本報告では、過去18年間の日経平均株価を訓練データとして、LSTMを用いたニューラルネットワークを深層学習のアルゴリズムにより学習させ、予測を行い、平均絶対誤差率を用いて評価し、実験的な考察を行った。訓練データの内容に依存して、学習

と予測の善し悪しが左右される傾向があるため、単一の学習モデルを構築するのではなく、様々な条件で学習を行い、予測日や予測期間に適したモデルを構築し予測に利用することが重要であると考えられる。また、本報告でのシミュレーションにおける条件パターンの少なさや、同じ条件でのシミュレーション回数が少ないことから、予測精度の向上のためにシミュレーション回数を増やす必要がある。今回はエポック数を2000回としたため、十分な学習収束が行われていない場合もあると考えられるため、エポック数を十分大きくした場合の学習制度の向上や汎化能力の向上、およびネットワークの構成を変えた場合の検証も必要であると考えられる。また、テクニカル指標の導入による効果も興味深いテーマであると考えられる。TOPIXやJASDAQなどの他の国内市場の株価や、S&P 500やNYダウ平均株価などの海外のなどの他のマーケットの株価、および個別銘柄の株価などを対象とし、知見を得ることも重要と考えられる。

## 参考文献

- (1) G. E. Hinton and Simon Osindero, Yee-Whyeh Teh: A Fast Learning Algorithm for Deep Belief Nets, *Neural Computation*, 18:1527-1544, 2006.
- (2) Y. Bengio: Learning deep architectures for AI, *Foundations and Trends in Machine Learning*, Vol 2, No.1, pp.1-127, 2009.
- (3) 岡谷貴之: 深層学習, 講談社, 2015.
- (4) Wolpert, D.H. and Macready, W.G.: No Free Lunch Theorems for Search, Technical Report SFI-TR-95-02-010, Sanga Fe, NM, 1995.
- (5) J. L. Elman: Finding Structure in Time, *Cognitive science*, 14, pp.179-221, 1990.
- (6) K. J. Lang, A. H. Waibel, G. E. Hinton: A time-delay neural network architecture for isolated word recognition, *Neural Networks*, Volume 3, Issue 1, pp.23-43, 1990.
- (7) S. Hochreiter and J. Schmidhuber: Long Short-Term Memory, *Neural Computation*, Vol.9, 8, Nov.15, pp.1735-1780, 1997.
- (8) L. Takeuchi, Y. Y. Lee: Applying Deep Learning to Enhance Momentum Trading Strategies in Stocks, 2013.
- (9) K. Chen, Y. Zhou and F. Dai: A LSTM-based

- method for stock returns prediction: A case study of China stock market, 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, pp. 2823–2824, 2015.
- (10) R. Xiong, E. P. Nichols, Y. Shen : Deep Learning Stock Volatility with Google Domestic Trends : 2018年5月5日, <https://arxiv.org/abs/1512.04916>, 2015.
- (11) X. Ding, Y. Zhang, T. Liu, J. Duan : Deep Learning for Event-Driven Stock Prediction, Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI2015.
- (12) H. Jia: Investigation into the effectiveness of long short term memory networks for stock price prediction, <https://arxiv.org/abs/1603.07893>, 2016.
- (13) L. D. Persio, O. Honchar: Artificial neural networks architectures for stock price prediction: Comparisons and applications, International Journal of Circuits, Systems and Signal Processing, 2016.
- (14) R. Akita, A. Yoshihara, T. Matsubara and K. Uehara : Deep learning for stock prediction using numerical and textual information, 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), Okayama, pp. 1–6, 2016.
- (15) 宮崎邦洋, 松尾豊 : 深層学習を用いた株価予測の分析, The 31th Annual Conference of the Japanese Society for Artificial Intelligence, 2017.
- (16) 田村浩一郎, 上野山勝也, 飯塚修平, 松尾豊 : 深層学習を用いたアンサンブルモデルによる株主価値推定モデルの提案, 人工知能学会論文誌, 33 巻, 1 号, p. A-H51\_1–11, 2018.
- (17) 大和証券 : 人工知能 (AI) による株価予測モデルを用いた銘柄情報の提供について, 2018年3月18日, <http://www.daiwa.jp/doc/171026.html>.
- (18) Yahoo! Finance : 2018年5月2日, <https://finance.yahoo.com/>.