International Journal of Innovative Computing, Information and Control
Volume **16**, Number **3**, June **2020**

ICIC International ©2020 ISSN 1349-4198
pp. 1093–1099

RECURRENT NEURAL NETWORK BASED STOCK PRICE PREDICTION USING MULTIPLE STOCK BRANDS

Shota Rikukawa, Hiroki Mori and Taku Harada

Department of Industrial Administration
Tokyo University of Science
2641 Yamazaki, Noda, Chiba 278-8510, Japan
{ 7415115; 7414147 }@alumni.tus.ac.jp; harada@rs.noda.tus.ac.jp

Received November 2019; revised February 2020

Abstract. *Stock price is difficult to predict because its fluctuation factor is complicated. We herein propose a stock price predicting method using the stock price of the predicted stock brand as well as those of other stock brands. A recurrent neural network is applied as a learning module. Furthermore, we use dynamic time warping as a similarity measure of stock price as time-series data. Subsequently, we conduct an experiment using the stock price data of companies listed in the first section of the Tokyo Stock Exchange and demonstrate the effectiveness of the proposed method.*

Keywords: Stock price prediction, Stock brand, Recurrent neural networks, Dynamic time warping

1. **Introduction.** Financial time-series data are extremely difficult to predict because their fluctuation factor is complicated. Many studies have been conducted on the prediction of stock prices, which is one of financial time-series data [1-7]. [8] reported that it is necessary to combine multiple data to improve the prediction accuracy of economic time-series. [9] used seven time-series data such as the exchange rate of Yen for Dollar in addition to the Nikkei average stock price data for economic time-series prediction. [10,11] estimated the Nikkei average stock price using other indicators such as the New York Dow.

In this study, the stock price is predicted for each stock brand. For the use of stock price data, many studies predicted the stock price for each stock brand using only the stock price data of the stock brand to be predicted, and few studies used stock price data of stock brands other than the predicted stock brand.

Dynamic Time Warping (DTW) [12-14] is an evaluation index to measure the degree of similarity between time-series data. Research on stock price prediction using DTW is currently being conducted [15-17].

Therefore, we herein propose a method to use the stock price data of stock brands other than the predicted stock brand and demonstrate its effectiveness by evaluation experiments. Our proposed method generates a learner using a recurrent neural network. Stock price is expressed as the time-series data. Therefore, we used the recurrent neural network as a learner.

In the proposed method, a learner is generated using the stock price data of only one stock brand as training data. A plurality of learners is generated by generating such a learner for a plurality of stock brands. Subsequently, using DTW as an index, a learner is selected to be applied in the prediction from among a plurality of learners. Since DTW can be employed to estimate the similarities between the time-series datasets, the differences

DOI: 10.24507/ijicic.16.03.1093

in the stock price fluctuation patterns between the training period and the test period can be considered.

In this study, we conduct experiments using the stock price data of companies listed in the first section of the Tokyo Stock Exchange and demonstrate the effectiveness of the proposed method. Using stock price data of different stock brands, various price fluctuation patterns can be handled and the prediction accuracy is expected to be improved. The effectiveness of using stock price data of stock brands other than the stock brand to be predicted, which has not been shown much in previous studies, is demonstrated.

The structure of this paper is as follows: Chapter 2 describes the generation of a learner using a recurrent neural network, Chapter 3 describes a selection of stock brand to use by DTW, Chapter 4 describes the experimental setting, Chapter 5 describes the experimental results, and Chapter 6 summarizes the work.

2. **Generation of Neural Network Learner.** Recurrent Neural Network (RNN) is used as a learner in the stock price prediction process. Figure 1 shows the RNN model used in this study. The closing price of stock data for 30 days is inputted into a node in an input layer. Subsequently, the stock price on the 31st day is predicted after passing through the middle layer, and learning is performed using this predicted value as an output. In the training period, the 30 days stock price data to be used are shifted for each day and used for learning. In the test period, the closing price of stock for 30 days from the day before to the day to predict is used as input. In the proposed method, one learner is generated using only one stock brand. That is, n learners are generated using the stock data of n stock brands.

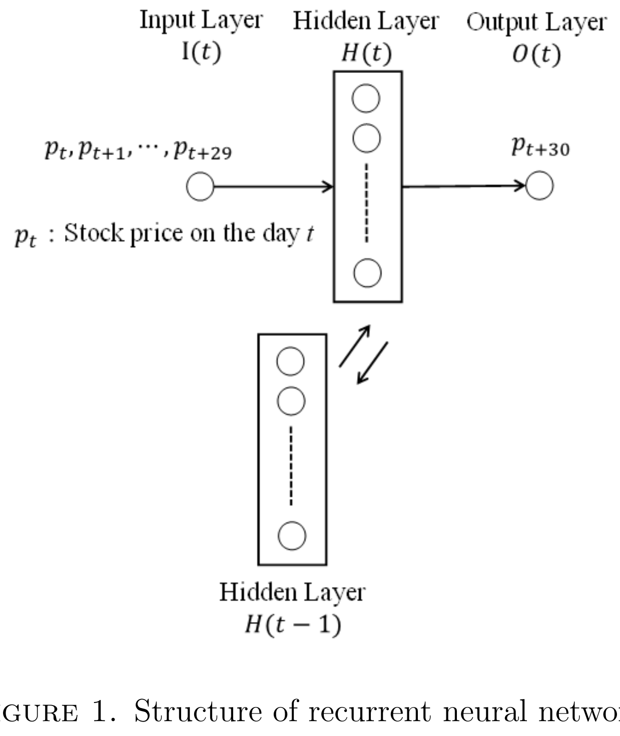

FIGURE 1. Structure of recurrent neural network

3. **Selection of Stock Brand by DTW.**

3.1. **DTW distance.** DTW [12-14] is an index to evaluate the similarity of time-series data. In DTW, it is possible to handle nonlinear expansion and contraction in the time direction by associating one point in the time-series data with multiple points in the other time-series data for two time-series data. Therefore, DTW distance can be defined even if the periodicity and length of two time-series data are different.

Let two waveform data be X, Y, the kth data of waveform X be x_k, and the kth data of waveform Y be y_k. Let the distance between x_{k_1} and y_{k_1} be $w_{k_1,k_2} = (x_{k_1}, y_{k_1})$. The set W of K correspondences of data is expressed by Equations (1) and (2).

$$W_{k_1} = \{w_{k_1,1}, w_{k_1,2}, \dots, w_{k_1,K}\} \tag{1}$$

$$W_{k_2} = \{w_{1,k_2}, w_{2,k_2}, \dots, w_{K,k_2}\} \tag{2}$$

W_k satisfies the following condition.

1) $w_{k_1-1,k_2} - w_{k_1,k_2} \leq 1$ and $w_{k_1,k_2-1} - w_{k_1,k_2} \leq 1$

2) $w_{k_1-1,k_2} - w_{k_1,k_2} \geq 0$ and $w_{k_1,k_2-1} - w_{k_1,k_2} \geq 0$

3) $w_{1,1} = (1,1)$ and $w_{K,K} = (K,K)$

The distance between waveforms by DTW is defined as follows:

$$DTW(x,y) = \min \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} |x_{k_1} - y_{k_2}| \tag{3}$$

Herein, the value obtained using Equation (3) is called the DTW distance.

3.2. **Application of DTW.** In this study, we use DTW distance as a selection index of a learner to improve the accuracy of stock price prediction. The stock price data with a smaller DTW distance from the stock price data of the predicted stock brand can be assumed as resembling the stock price data of the predicted stock brand. Therefore, a more accurate stock price prediction may be performed using a learner with stock price data similar to the stock price data of the predicted stock brand.

1) The stock price data of each used stock brand exhibit various price ranges. Therefore, stock price data are normalized to use data focusing on the fluctuation of stock price.
2) Figure 2 shows the calculation method of the DTW distance in the proposed method. The learner is generated by learning in 164 days per stock brand. The DTW distance is calculated between the stock data for the past 30 days of the predicted day in the test period of the predicted stock brand and 30 days of the 164 days of training period in the learner of each stock brand. Therefore, 135 DTW distances were calculated to one learner for predicting the day with the predicted stock brand. At that time, we

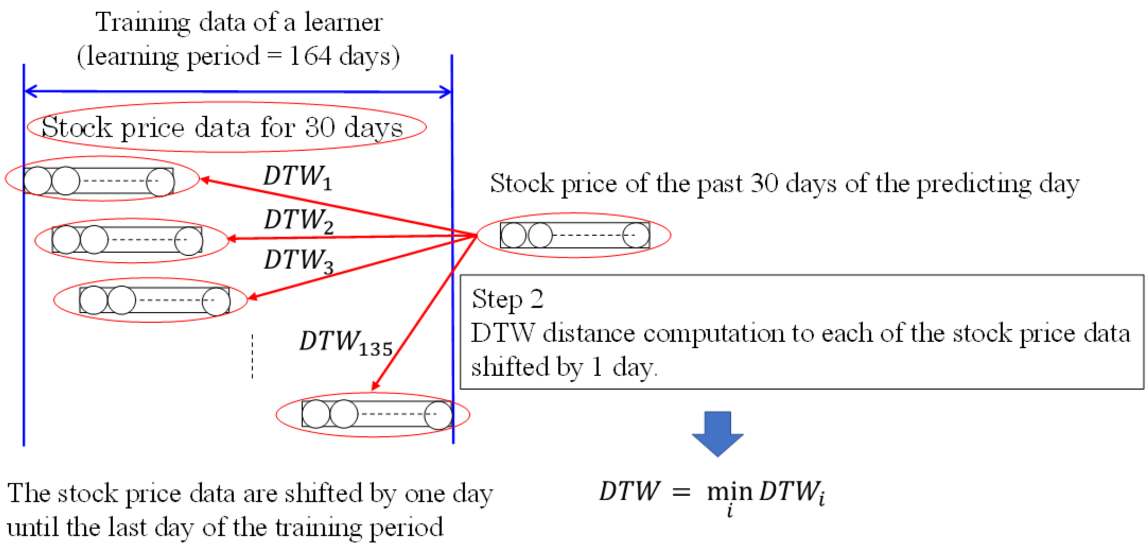

FIGURE 2. Calculation of DTW distance with each stock brand

calculate for the entire training period while shifting one day from the beginning of the training period. Subsequently, the minimum DTW distance in DTW_1 to DTW_{135} of each stock brand is set as the DTW distance between the predicted day and the learner.

3) As our proposed method generates a learner for each stock brand, the number of learners generated was the same as the number of stock brands that could be used for prediction. Step 2) is performed between the predicted stock brand and all learners for each prediction day. Subsequently, the learner with the smallest DTW distance in the DTW distances calculated in Step 2) is selected, and prediction is performed using this learner. This step is shown in Figure 3.
4) Step 2) and Step 3) are repeated for 30 days in the test period while shifting the day to predict by one day.

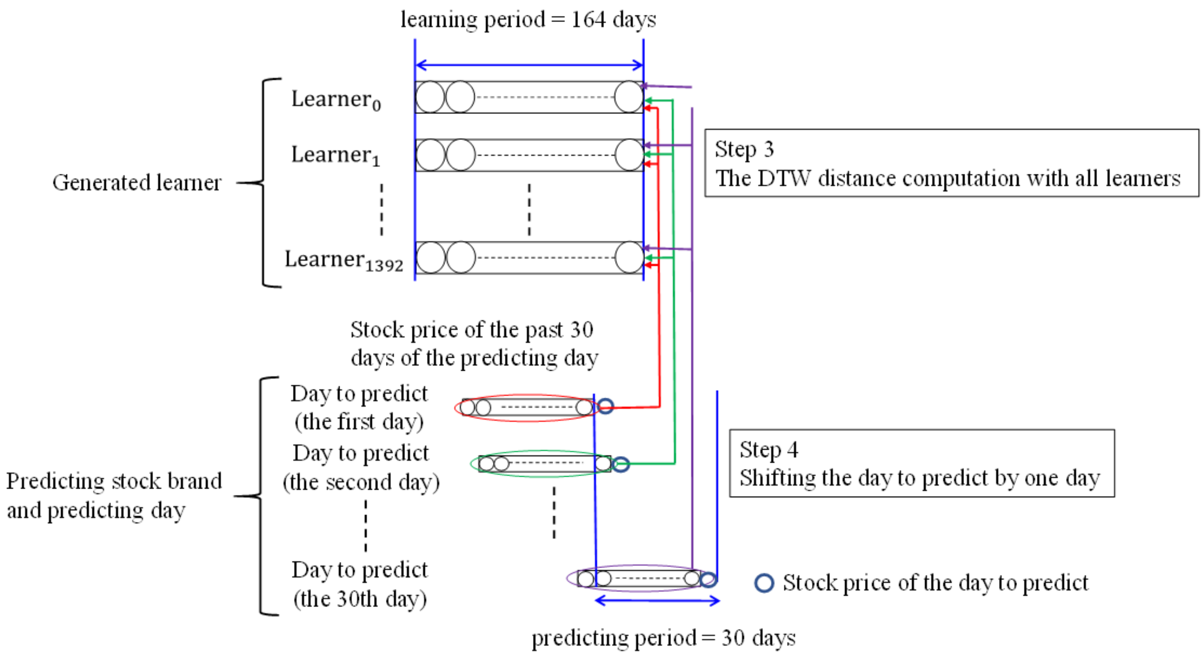

FIGURE 3. Selection of learners

4. Experimental Setting.

4.1. **Index to evaluate prediction accuracy.** The Root Mean Square Percentage Error (RMSPE) and Mean Absolute Percentage Error (MAPE) are used as evaluation indices for prediction accuracy. The RMSPE and MAPE are defined as follows, respectively. n is the number of data, y_i' is predicted value, and y_i is observation value.

$$RMSPE = \sqrt{\frac{100}{n}\sum_{i=1}^{n}\left(\frac{y_i' - y_i}{y_i}\right)^2} \tag{4}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i' - y_i}{y_i}\right| \tag{5}$$

As these values become smaller, it is possible to predict a value close to the true value in the prediction.

4.2. **Data for use.** The stock brands used for the experiment are 1393 stock brands that continue to be listed in the first section of the Tokyo Stock Exchange from 2004 to 2017 [18]. Among them, 50 stock brands randomly selected across industries were considered as prediction stock brands. These 50 stock brands are brands whose MAPE and RMSPE values are within 100 in their prediction results. If the MAPE and RMSPE are evaluated at an average of 50 stock brands, they will be significantly affected by the stock brands of which their values are significantly higher, and thus, it will be difficult to perform a legitimate evaluation. Therefore, herein, stock brands with an MAPE and RMSPE of 100 or less are considered as prediction stock brands.

4.3. **Training period and test period.** The training period and test period were set as follows:

- Training period: January 4 – August 31, 2016 (164 days)
- Test period: October 18 – November 30, 2016 (30 days)

4.4. **Parameters of RNN.** The values of the RNN parameters were set by preliminary experiments and are shown in Table 1. In the preliminary experiment, 10 stock brands randomly selected were predicted. Regarding the learning rate and number of epochs, the optimum parameter values differed for each stock brand, and thus the value of such parameters could not be determined to one value. Therefore, the prediction was performed by all combinations, and the value of the parameter with the best prediction accuracy was used. To set more appropriate value for the recurrent neural network parameters, further discussion on the selection method for employing the stock brands in the preliminary experiment will be needed.

TABLE 1. Recurrent neural networks parameters

Parameter	Setting value
Learning rate	$(0.01, 0.001)$
Batch size	20
Number of layers in the middle layer	1
Number of nodes in the middle layer	100
Activation function	Linear
Number of epochs	$(2000, 4000, 8000)$

5. **Experimental Results.** The prediction was performed in the 30 days of the test period using the proposed method. In the experiment, the proposed method is compared with the method using a single stock brand. In the proposed method, prediction is performed by switching to the learner of the stock brand with the shortest DTW distance daily. However, in the method using a single stock brand, prediction is performed using only the learner of the prediction stock brand. In each method, the MAPE and RMSPE are used to compare their prediction accuracies.

Table 2 compares the proposed method and the method using a single stock brand for the MAPE and RMSPE of each stock brand, and shows the number of stock brands whose values are smaller for each of MAPE and RMSPE. The smaller the values of MAPE and RMSPE are, the closer they are to the true values. As shown in Table 2, the proposed method has more stock brands in terms of good prediction accuracy than the method using a single stock brand. In addition, Table 3 shows that the proposed method is more accurate than the method using a single stock brand for each average value of MAPE and RMSPE. From the results above, it can be concluded that the proposed method is

Table 2. Number of stock brands with relatively small values of MAPE and RMSPE

	MAPE	RMSPE
Single stock brand method	16	19
Proposed method	34	31

Table 3. Average values of MAPE and RMSPE

	MAPE	RMSPE
Single stock brand method	9.2589	1.5889
Proposed method	5.5983	1.1784

Table 4. Average values of DTW distance

	Average DTW distance
Single stock brand method	3.0996
Proposed method	0.5546

effective for the method using a single stock brand. Table 4 shows the average values of DTW distance for 50 stock brands. For each stock brand, the average value of the shortest DTW distance obtained by a day during the test period for 30 days is evaluated. Shorter DTW distances were obtained with the proposed method compared to the methods that employed single stock brand.

6. **Conclusions.** We proposed a method for stock price prediction using stock price data of multiple stock brands. In the proposed method, a learner was generated using a recurrent neural network. Then, the learner to be applied was basically switched every day to be predicted. The DTW distance was used as a selection criterion for a learner. Furthermore, the effectiveness of stock price prediction using not only stock price data of stock brand to be predicted but also stock price data of other stock brands was shown experimentally.

Here, the DTW distance was calculated using 30 days stock price data. Since various price fluctuation patterns are possible, the future work should set the DTW distance calculation period with a greater flexibility.

REFERENCES

[1] P.-F. Pai and C.-S. Lin, A hybrid ARIMA and support vector machines model in stock price forecasting, *Omega – The International Journal of Management Science*, vol.33, pp.497-505, 2005.

[2] P.-C. Chang and C.-H. Liu, A TSK type fuzzy rule based system for stock price prediction, *Expert Systems with Applications*, vol.34, pp.135-144, 2008.

[3] G. S. Atsalakis and K. P. Valavanis, Surveying stock market forecasting techniques – Part II: Soft computing methods, *Expert Systems with Applications*, vol.36, pp.5932-5941, 2009.

[4] F. A. de Oliveira, C. N. Nobre and L. E. Zarate, Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index – Case study of PETR4, Petrobras, Brazil, *Expert Systems with Applications*, vol.40, pp.7596-7606, 2013.

[5] M. Ballings, D. V. den Poel, N. Hespeels and R. Gryp, Evaluating multiple classifiers for stock price direction prediction, *Expert Systems with Applications*, vol.42, pp.7046-7056, 2015.

[6] M. Göçken, M. Özçalici, A. Boru and A. T. Dosdoğru, Integrating metaheuristics and artificial neural networks for improved stock price prediction, *Expert Systems with Applications*, vol.44, pp.320-331, 2016.

[7] I. Ibidapo, A. Adebiyi and O. Okesola, Soft computing techniques for stock market prediction: A literature survey, *Covenant Journal of Informatics & Communication Technology*, vol.5, no.2, pp.1-28, 2017.
[8] S. Komaki and S. Shirayama, Forecasting the Nikkei stock average using deep belief network, *The Japanese Society for Artificial Intelligence, Special Interest Group on Financial Informatics*, SIG-FIN-012-08, 2014 (in Japanese).
[9] T. Nakata and T. Koseki, Prediction of short term economic time series focused on the direction of price changes by support vector machine, *Proc. of the Institute of Electronics, Information and Communication Engineers General Conference 2008*, p.264, 2008 (in Japanese).
[10] S. Kobayashi and S. Shirayama, Comparing the forecast of Nikkei225 using deep belief networks and Bayesian network, *The Japanese Society for Artificial Intelligence, Special Interest Group on Data Oriented Constructive Mining and Simulation*, SIG-DOCMAS, 2015 (in Japanese).
[11] S. Kobayashi and S. Shirayama, Time series forecasting with multiple deep learners: Selection from a Bayesian network, *Journal of Data Analysis and Information Processing*, vol.5, pp.115-130, 2017.
[12] H. Sakoe and S. Chiba, Dynamic programming algorithm optimization for spoken word recognition, *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol.ASSP-26, no.1, pp.43-49, 1978.
[13] E. J. Keogh and M. J. Pazzani, Scaling up dynamic time warping to massive datasets, in *Lecture Notes in Computer Science*, J. M. Żytkow and J. Rauch (eds.), Berlin, Heidelberg, Springer, 1999.
[14] E. J. Keogh and M. J. Pazzani, Scaling up dynamic time warping for datamining applications, *Proc. of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.285-289, 2000.
[15] K. Nakagawa, M. Imamura and K. Yoshida, Stock price prediction using k-medoids clustering with indexing dynamic time warping, *IEEJ Trans. Electronics, Information and Systems*, vol.138, no.8, pp.986-991, 2018 (in Japanese).
[16] S. Jeon, B. Hong and V. Chang, Pattern graph tracking-based stock price prediction using big data, *Future Generation Computer Systems*, vol.80, pp.171-187, 2018.
[17] X. Yao and H.-L. Wei, Short-term stock price forecasting based on similar historical patterns extraction, *Proc. of the 23rd International Conference on Automation & Computing*, 2017.
[18] *Tokyo Stock Exchange*, https://www.jpx.co.jp.