Hindawi Discrete Dynamics in Nature and Society Volume 2022, Article ID 7119678, 8 pages https://doi.org/10.1155/2022/7119678



Research Article

A Data Organization Method for LSTM and Transformer When Predicting Chinese Banking Stock Prices

Zong-Yu Peng and Pei-Chang Guo

School of Science, China University of Geosciences, Beijing 100083, China

Correspondence should be addressed to Pei-Chang Guo; gpeichang@126.com

Received 12 October 2021; Revised 3 December 2021; Accepted 7 December 2021; Published 4 January 2022

Academic Editor: Chun Wei

Copyright © 2022 Zong-Yu Peng and Pei-Chang Guo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The accurate prediction of stock prices is not an easy task. The long short-term memory (LSTM) neural network and the transformer are good machine learning models for times series forecasting. In this paper, we use LSTM and transformer to predict prices of banking stocks in China's A-share market. It is shown that organizing the input data can help get accurate outcomes of the models. In this paper, we first introduce some basic knowledge about LSTM and present prediction results using a standard LSTM model. Then, we show how to organize the input data during the training period and give the comparison results for not only LSTM but also the transformer model. The numerical results show that the prediction results of LSTM and transformer can be improved after the input data are organized when training.

1. Introduction

Predicting stock prices is a very important and difficult research topic for researchers from financial and academic fields. As we know, the price of a stock is influenced by many factors including company profiles, industry prospects, government policies, investors' behavior, news and social media, etc. We understand that the stock prices are affected by many uncontrolled factors, and we may need to study the annual share report, company performance, and so on to get the meaningful prediction for a stock. But we observe that there was no big fluctuation in Chinese banking stocks, and predicting price changes of Chinese banking stocks can be handled as a time series forecasting problem. Long shortterm memory (LSTM) neural networks [1] and the transformer model [2] are two powerful deep learning tools for times series forecasting. In this paper, we show how to organize the data to make the input have more periodic information when using LSTM and transformer to predict the prices of Chinese banking stocks.

There have been many research results about stock price predictions in recent years. In [3], two types of analysis are used. Fundamental analysis is suitable for long-term

forecasting, which uses the information about company profiles, political factors, and economic factors. Technical analysis uses the historical prices of a stock to predict the future price, which is preferred for short-term prediction. Technical analysis is widely used to indicate when to buy or sell stocks, where K-line, moving average, and relative strength index are common algorithms [4]. Islam et al. [5] utilized the EMD method for the decomposition of data sequences, in comparison with the wavelet decomposition method, finding that the EMD method is more effective.

In recent years, intelligent algorithms are designed and analyzed in many domains such as the attitude tracking and control of uncertain rigid spacecraft [6–8], the control of wind energy conversion systems [9, 10], and the analysis and design of high step-down converters [11]. With the rapid progress of artificial intelligence, more and more researchers use machine learning techniques for stock market forecast. Support vector regression (SVR) is a machine learning model that could identify the hyperplane in a high-dimensional space. In [12], SVR was used to predict the opening price in the following day given the historical time series data. Recurrent neural networks (RNNs) are a class of deep learning models for processing sequential data. In [13],

recurrent neural networks (RNN) were shown to have better accuracy than SVR models when predicting stock prices.

In practical applications, long short-term memory (LSTM) neural networks, which belong to gated RNNs, can learn long-term dependencies more easily than the simple recurrent architectures. LSTM networks are proved to be very successful for times series forecasting and have been applied successfully to many fields including speech recognition and machine translation [1]. Also, the transformer is the state-of-the-art model based on the attention mechanism [2, 14]. In this paper, we propose a data organization method and show that our method works for these two mainstream models to predict the Chinese banking stock prices.

Recently, LSTM has attracted many scholars to use it in predicting the stock prices. We introduce the following related works. In 2016, paragraph vector was applied to obtain distributed representations of newspaper articles, and a LSTM model that takes into account textual information was proposed to deal with the influence of time series [15]. In 2018, the Conv1D-LSTM model, a combination of 1-dimensional convolution neural network (CNN) and LSTM model, was proposed to predict stock prices [16]. This model can absorb strengths from two networks: CNN can effectively perform feature extraction and LSTM can handle sequential data well. A wavelet transform is used to denoise historical stock data based on LSTM and an attention mechanism in [17]. A normalized comparison of the performances of LSTM and GRU for stock market forecasting was performed in [18]. In [19], the authors proposed a method combining deep learning with empirical mode decomposition to predict financial trends accurately from financial data. After the transformer model was proposed for sequence modeling [2], new approaches based on the transformer were proposed to tackle the stock movement prediction task [14].

As for China's stock market prediction, there also have been many research papers in recent years. For example, in 2015, Chen et al. used a LSTM model to forecast stock returns in China's A-share market, and this model has higher accuracy compared with random prediction method [20]. In [21], Peng et al. improved stock price prediction through changing the number of LSTM layers and hidden neurons.

We can see that different LSTM-based and transformer-based models have been proposed for stock price predictions. LSTM and transformer use the characteristic information in the sequential series to predict the trend of future data. So, for every prediction, the input data matter a lot for the accuracy of outcome value. In this paper, we present a data organization method and show that this method can improve the accuracy of validation data for LSTM and the transformer. The motivation is as follows. According to the economic theory of "Kondratiev waves," capitalist economies experience major business cycles that recur at several years. Economic activity often follows a cyclical pattern. So, the prices of many stocks move up and down nearly in cycles too. So, we think the information in more distant days should be carried for the input data in

order for the LSTM and the transformer models to better learn from the historical prices. Then, we devise this data organization method. After the organization, the input data will carry information in more distant days. The experiments confirm that this method works. In Section 2, we present the structure of a LSTM model and present the prediction results of validation data for a bank stock. In Section 3, we show how to organize the data to make the input have more periodic information when training. Experiments illustrate the effectiveness of this method. At the end of the paper, we give some conclusions.

2. Experiments with LSTM

2.1. The Structure of LSTM. First, we give a brief introduction about LSTM. LSTM was first proposed by Hochreiter and Schmidhuber in 1997 [22]. The problem of exploding gradients and vanishing gradients in vanilla RNNs can be relieved by this architecture. The flow of information is regulated by using gates to selectively read, write, and forget information in LSTM networks. There are three gates: input gate, forget gate, and output gate, and two states: cell state and hidden state, in a LSTM block diagram. The key to LSTM is the cell state, which refers to the state of cell used for long propagation of information. A LSTM block diagram produces long-term outcome C_t and short-term outcome h_t and takes h_{t-1} , the short-term outcome at previous moment, and the input of x_t at this step as inputs. We present a LSTM block diagram in Figure 1.

First, we explain the notation involved in Figure 1. f_t denotes the "forget gate layer," i_t is the "input gate layer," o_t is the "output gate layer," \tilde{C}_t is a tanh layer creating new candidate value for C_t , W_* and b_* denote the corresponding weights and bias for each layer where * denotes f, i, C, o respectively. C_t is the new cell state, which is a weighted combination of old cell state C_{t-1} and candidate value \tilde{C}_t . h_t is the output, a filtered version based on C_t . We have the following equations about three gates and the cell state [23].

(1) Forget gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \tag{1}$$

(2) Input gate:

$$i_{t} = \sigma(W_{i}[h_{t-1}, x_{t}] + b_{i}),$$

$$\widetilde{C}_{t} = \tanh(W_{C} \cdot [h_{t-1}, x_{t}] + b_{C}).$$
(2)

(3) Cell state (long-term) equation at t moment:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t. \tag{3}$$

(4) Output gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o),$$

$$h_t = o_t \cdot \tanh(C_t).$$
(4)

We use Keras API in TensorFlow framework to build a LSTM model for stock price predictions. The model has two LSTM layers, and the first LSTM layer has 80 neurons and

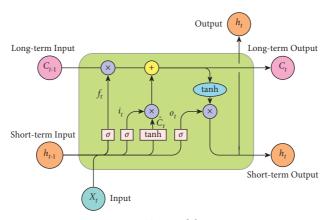


FIGURE 1: LSTM model structure.

the other LSTM layer has 100 neurons. The following are the details about the prediction for the daily closing stock prices of Industrial Commercial Bank of China (ICBC, sh601398) in China's A-share market.

- 2.2. Experiment Details. The aim is to predict the closing stock price on the 11-th trading day using daily closing stock price information in the past 10 trading days. First, the original stock price data are linearly scaled to be in the interval (0, 1]. After this normalization process, the LSTM model takes the stock price data in the last 10 continuous trading days as the input and outputs a prediction value about the stock price in the 11-th trading day.
 - (1) Data Preparation. We download the daily closing price data of ICBC stock via web crawler from a domestic financial data platform and store them in HDF files. We select the daily stock price data in 3300 trading days to train and test a LSTM model. During the training process, we use the data in the first 3100 days to generate 3090 prediction values to get the LSTM to learn its parameters. During the validation process, we use the stock price data in 210 days, i.e., the remaining 200 days and 10 days backward, to generate 200 prediction values to compare with the real stock price values in these remaining 200 days.
 - (2) *Model Training*. During the training process, we save the learning model that produces the highest validation accuracy, i.e., the model with the smallest loss of validation data. The epoch number is 25, and the mini-batch size is 16.
 - (3) Stock Price Prediction and Model Evaluation. Using the saved trained model to predict stock prices, we present the comparison results between prediction values and real values of validation data and then compute some evaluation indicators (error information) of the saved model.
- 2.3. Prediction Results. We use the above LSTM model to predict the stock price of Industrial and Commercial Bank of China (ICBC) and demonstrate the prediction results of validation data. The historical real stock price data are from

March 28, 2007, to November 18, 2020, and were bought from https://yucezhe.com/product?name=trading-data.

As we can see from Figure 2, the accuracy of prediction values of validation data is not very satisfactory. In the next section, we will organize the data to improve the accuracy.

3. Organizing Input Data to Improve the Accuracy

Daily prices of a stock are time series data and often have periodic characteristics. The theory of "Kondratiev waves" proposed by the Russian economist Kondratiev tells us that each commodity market could support an overarching long wave market cycle [24]. Inspired by "Kondratiev waves," we may improve the accuracy of prediction values by feeding the model with organized data from a longer historical period. For the model, we will not take the 10 closing price values in the past 10 days as the input; instead, we will take 10 specified closing price values chosen from data in the past n days ($n \gg 10$) as the input. We see that after the input data are organized, the LSTM model can gain benefits in terms of accuracy of validation data.

3.1. Details about Organizing Input Data. In our experiments, we use price values in the following 10 specific days to get the prediction value in the current day: 2 years ago, 1 year and a half ago, 1 year ago, 6 months ago, 3 months ago, 1 month ago, 2 weeks ago, 1 week ago, 2 days ago, and 1 day ago.

Noting that stock transactions may happen discontinuously because of holidays, suspension of trading in a stock, and other factors, we use the following equalities to choose the above specified days.

1 year
$$\approx 250$$
 trading days,
1 month ≈ 21 trading days, (5)
1 week ≈ 5 trading days.

So, the 10 characteristic days will be chosen as follows: 500 days ago, 375 days ago, 250 days ago, 125 days ago, 63 days ago, 21 days ago, 10 days ago, 5 days ago, 2 days ago, and 1 day ago.

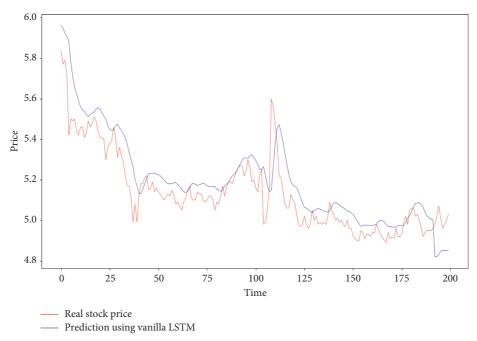


FIGURE 2: Daily line prediction results of validation data for ICBC stock.

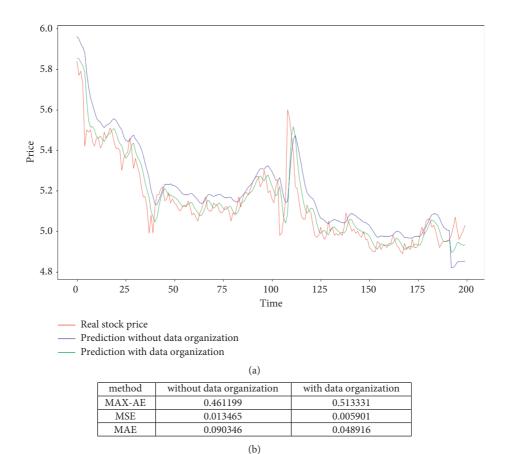


FIGURE 3: Comparison results of validation data for ICBC stock price. (a) Prediction results. (b) Model evaluation indicators.

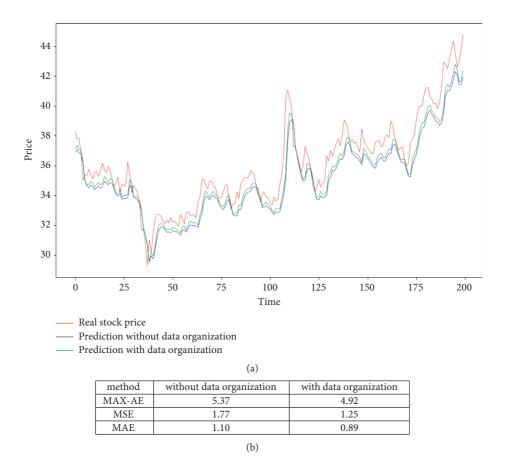


FIGURE 4: Comparison results of validation data for China Merchants Bank Co Ltd stock price. (a) Prediction results. (b) Model evaluation indicators.

We still use the data of ICBC (sh601398), i.e., the daily closing prices in 3300 trading days to do experiments. The historical real stock price data are from March 28, 2007, to November 18, 2020, and were bought from https://yucezhe.com/product?name=trading-data. But this time, the input of 10 values for the LSTM model to produce the prediction value is chosen to be the daily closing prices in the above 10 characteristic days. The code for organizing the input training data for the LSTM model to learn is reported as follows:

```
for i in range(period, len(train_set)):
    if i < 500:
        x_train.append(train_set[i - period:i, 0])
    else:
        x_train.append(train_set[
            [i - 500, i - 375, i - 250, i - 125,
            i - 63, i - 21, i - 10, i - 5, i - 2,
            i - 1], 0])</pre>
```

y_train.append(train_set[i, 0])

As we can see, now the prediction value on the current day needs the information 500 days ago. So, during the training process, if we use the data in the first 3100 days to train, we can only produce 2600 prediction values to get the LSTM to learn its parameters. To take full advantage of the data to train the LSTM model, we do as follows. For an epoch when training, there are two stages. First we use the data in the first 500 days to generate 490 prediction values to train the LSTM model, where the daily closing prices in the past 10 continuous trading days are taken as the input to train the LSTM model; then, we use the data in the first 3100 days to generate 2600 prediction values to train the LSTM model, where the daily closing prices in the 10 chosen characteristic days are taken as the input to train the model. Thus, we can totally produce 490 + 2600 = 3090prediction values to get the LSTM to learn its parameters. During the validation process, we use the stock price data in 210 days, i.e., the remaining 200 days in the total 3100 days and 10 days backward, to produce 200 prediction values to

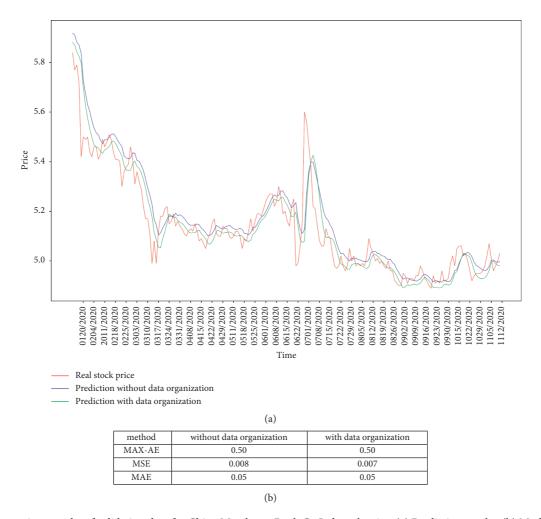


FIGURE 5: Comparison results of validation data for China Merchants Bank Co Ltd stock price. (a) Prediction results. (b) Model evaluation indicators.

compare with the real closing price values in the remaining 200 days.

We name this approach as "prediction with data organization" and name the approach in Section 2 as "prediction without data organization." We introduce some acronyms to help see the comparison results. MAX-AE denotes the maximum of absolute errors between each prediction value and real value of validation data, and MSE and MAE denote mean square error and mean absolute error of validation data, respectively. We demonstrate the prediction results for ICBC stock on validation data in Figure 3. We can see that the organization of data can improve the prediction accuracy of validation data.

Experiments were also conducted for other Chinese banking stocks. After introducing the data organization, all the prediction results can be improved more or less. In Figure 4, we present the results for China Merchants Bank Co Ltd (sh600036) in China' A-share market. The historical real stock price data are from March 28, 2007, to November 18, 2020, and were bought from https://yucezhe.com/product?name=trading-data. Again, we can see that the organization of data can improve the accuracy of validation

To further demonstrate the effectiveness of this data organization method, next we use the transformer model to conduct comparison experiments. Here we demonstrate the results for ICBC and China Merchants Bank Co Ltd for example. The historical real stock price data are from March 28, 2007, to November 18, 2020, and were bought from https://yucezhe.com/product?name=trading-data. The comparison results using transformer are shown in Figure 5 for ICBC stock and Figure 6 for China Merchants Bank Co Ltd. Again, we can see that the data organization approach can improve the accuracy of validation data.

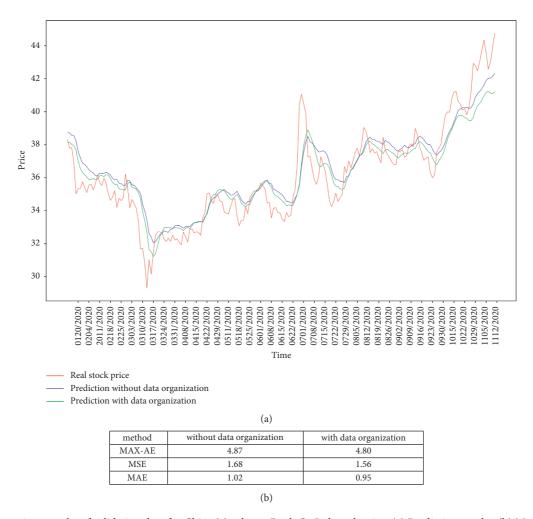


FIGURE 6: Comparison results of validation data for China Merchants Bank Co Ltd stock price. (a) Prediction results. (b) Model evaluation indicators.

4. Conclusion

In this paper, motivated by "Kondratiev waves," we report an interesting phenomenon when using deep learning to predict stock prices. The LSTM and the transformer are powerful tools for processing sequential data. But there is time lag for the outcomes of LSTM and transformer when predicting daily stock prices. In this paper, we show how to organize the data to make the input have more periodic information when training. We conducted some experiments for some banking stocks in China's A-share market. From the experimentation results, this data organization approach can help improve the accuracy of predictions about banking stocks on validation data.

While the idea leads to accuracy gains by incorporating long-range information through the use of carefully selected input sequences, there are many details that are needed to be investigated about this approach. In this paper, the input sequence of the LSTM model and the transformer model are chosen to be the price information 500 days ago, 375 days ago, 250 days ago, 125 days ago, 63 days ago, 21 days ago, 10 days ago, 5 days ago, 2 days ago,

and 1 day ago. But up to now, this choice has no strict theoretical analysis. We just find that this choice is fairly good for many Chinese banking stocks. We should not expect that these values are the best ones for all stocks. How could we choose the lags in practice? This is left to be investigated in the future.

Data Availability

The data used to support the findings of this study are available from the corresponding author (peichang@cugb.edu.cn) upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported in part by the National Natural Science Foundation of China (grant nos. 12001504 and 41874165) and the Fundamental Research Funds for the Central Universities (grant no. 2652019320).

References

- [1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge, UK, 2016.
- [2] A. Vaswani, "Attention is all you need," in *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, 2017.
- [3] D. Shah, H. Isah, and F. Zulkernine, "Stock market analysis: a review and taxonomy of prediction techniques," *International Journal of Financial Studies*, vol. 7, no. 2, pp. 1–22, 2019.
- [4] O. Bustos and A. Pomares-Quimbaya, "Stock market movement forecast: a systematic review," *Expert Systems with Applications*, vol. 156, pp. 1–15, 2020.
- [5] M. R. Islam, M. Rashed-Al-Mahfuz, S. Ahmad, and M. K. I. Molla, "Multiband prediction model for financial time series with multivariate empirical mode decomposition," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 593018, 21 pages, 2012.
- [6] Q. Chen, M. Tao, X. He, and L. Tao, "Fuzzy adaptive nonsingular fixed-time attitude tracking control of quadrotor UAVs," *IEEE Transactions on Aerospace and Electronic Sys*tems, vol. 57, no. 5, pp. 2864–2877, 2021.
- [7] Q. Chen, S. Xie, and X. He, "Neural-network-based adaptive singularity-free fixed-time attitude tracking control for spacecrafts," *IEEE Transactions on Cybernetics*, vol. 51, no. 10, pp. 5032–5045, 2021.
- [8] Q. Chen, Y. Ye, Z. Hu, J. Na, and S. Wang, "Finite-time approximation-free attitude control of quadrotors: theory and experiments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 3, pp. 1780–1792.
- [9] C. Wei, Z. Zhang, W. Qiao, and L. Qu, "Reinforcement-learning-based intelligent maximum power point tracking control for wind energy conversion systems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6360–6370, 2015.
- [10] C. Wei, Z. Zhang, W. Qiao, and L. Qu, "An adaptive network-based reinforcement learning method for MPPT control of PMSG wind energy conversion systems," *IEEE Transactions on Power Electronics*, vol. 31, no. 11, pp. 7837–7848, 2016.
- [11] C. Wei, Y. Zhao, Y. Zheng, L. Xie, and K. Smedley, "Analysis and design of a non-isolated high step-down converter with coupled inductor and ZVS operation," *IEEE Transactions on Industrial Electronics*, p. 1, 2021.
- [12] Y. Xia, Y. Liu, and Z. Chen, "Support Vector Regression for prediction of stock trend," in Proceedings of the 2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering, vol. 123-126, 2013.
- [13] Y. Liu, Z. Qin, P. Li, and T. Wan, "Stock volatility prediction using recurrent neural networks with sentiment analysis," in Advances in Artificial Intelligence: From Theory to Practice, S. Benferhat, K. Tabia, and M. Ali, Eds., Springer, Berlin, Germany, 2017.
- [14] Q. Ding, S. Wu, H. Sun, J. Guo, and J. Guo, "Hierarchical multi-scale Gaussian transformer for stock movement prediction," in Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence Special Track on AI in FinTech, pp. 4640–4646, Yokohama, Japan, January 2020.
- [15] R. Akita, A. Yoshihara, and T. Matsubara, "Deep learning for stock prediction using numerical and textual information," in Proceedings of the 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), June 2016.

- [16] S. Jain, R. Gupta, and A. Moghe, "Stock price prediction on daily stock data using deep neural networks," in *Proceedings of the 2018 International Conference on Advanced Computation and Telecommunication (ICACAT)*, Bhopal, India, December 2018
- [17] J. Qiu, B. Wang, and C. Zhou, "Forecasting stock prices with long-short term memory neural network based on attention mechanism," *PLoS One*, vol. 15, no. 1, Article ID e0227222, 2020.
- [18] T. B. Shahi, A. Shrestha, A. Neupane, W. Guo, and S. Price, "Stock price forecasting with deep learning: a comparative study," *Mathematics*, vol. 8, no. 9, p. 1441, 2020.
- [19] S.-L. Lin and H.-W. Huang, "Improving deep learning for forecasting accuracy in financial data," *Discrete Dynamics in Nature and Society*, vol. 2020, pp. 1–12, 2020.
- [20] K. Chen, Y. Zhou, and F. Dai, "A LSTM-based method for stock returns prediction: a case study of China stock market," in *Proceedings of the 2015 IEEE International Conference on Big Data (Big Data)*, Santa Clara, CA, USA, October-November 2015.
- [21] Y. Peng, Y. Liu, and R. Zhang, "Modeling and analysis of stock price forecasting based on lstm," *Computer Engineering and Applications (in Chinese)*, vol. 55, no. 11, pp. 209–212, 2019.
- [22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [23] A. Graves, Supervised Sequence Labelling, Springer, Berlin, Germany, 2012.
- [24] N. Leo, The Sixth Kondratieff: The New Long Wave in the Global Economy, CreateSpace Independent Publishing Platform, Scotts Valley, CA, USA, 2017.