Stock Price Prediction Based on LSTM in Python

Zhuojin Zhang*

Department of Mathematics, Xiamen University Malaysia, Jalan Sunsuria, Bandar Sunsuria, 43900, Sepang, Selangor, Malaysia

*Corresponding author: MAT2109464@xmu.edu.my

Abstract. To reduce the risk of people buying and increase the rate of return, applying a suitable model to predict stock prices is necessary. The research applies Long Short Term Memory (LSTM), which is a kind of artificial neural network in machine learning and is also a special kind of RNN. Through studying this model, the memory of LSTM is like human memory, which is composed of short-term memory and long-term accumulation. In this paper, the current data of 2020 and 2021 is to train, and the data from 2017 to 2019 is for testing. The data comes from three companies, GOOGLE, APPLE and AMAZON in the S&P500 index. Meanwhile, Mean Squared Error (MSE) is used for evaluation. Finally, in comparison with the different training times, LSTM presents good applicability, achieving negligible MSE.

Keywords: Stock price prediction; data analysis; LSTM.

1. Introduction

With the improvement of people's living standard, many people choose to increase their wealth by buying stocks. Exploiting machine learning is a way of predicting the future value of a stock, which allows us to apply data to accurately describe its future changes. Therefore, it reduces the risk of people buying and increases the rate of return.

In our real life, stock data is dynamic, nonlinear and nonparametric [1], while there exists multiple factors that affect stock prices, like many people buying all of a sudden or listed companies going bankrupt and so on. It is challenging for us to predict stock prices. However, many scholars still used many methods in machine learning to predict stock prices, such as Back Propagation (BP) Neural Network, Recurrent Neural Networks (RNN), ARIMA model and Long Short-Term Memory (LSTM) method. Some people chose BP neural network and ARIMA model to obtain results, which showed the prediction results of the former were better than the latter in short, medium and long term. BP neural network can solve the problem of nonlinear stock data but has limitations when dealing with the temporal correlation of stock data [3]. Peng Yan et al. employed Recurrent Neural Networks (RNN) to predict stock prices and trained and tested data sets by changing the number of Neural network layers and the number of Neural units in hidden layers. The results showed the prediction accuracy was 30% higher than that using traditional machine learning methods [2]. However, the gradient disappearance and explosion issues with the Naive RNN. Therefore, LSTM is a good choice to deal with the problems highly related to financial time series and has higher accuracy about the data problem of dealing with nonlinear trend and sequence correlation [3]. This essay mainly puts forward to a method using LSTM in python to predict stock prices. Based on this, this paper mainly applies S&P 500 stock index, regards the three company, GOOGLE, APPLE and AMAZON, as examples and employ the data from 01/01/2020 to 12/31/2021 for training and the data from 01/01/2017 to 12/31/2019 for testing. Meanwhile, Mean Squared Error (MSE) will be used for evaluation.

2. Method

2.1 Introduction of RNN

A recurrent neural network (RNN) is a class of artificial neural networks with internal self-connection, and it can deal with sequential changes in data. Fig. 1 shows the structure of RNN.

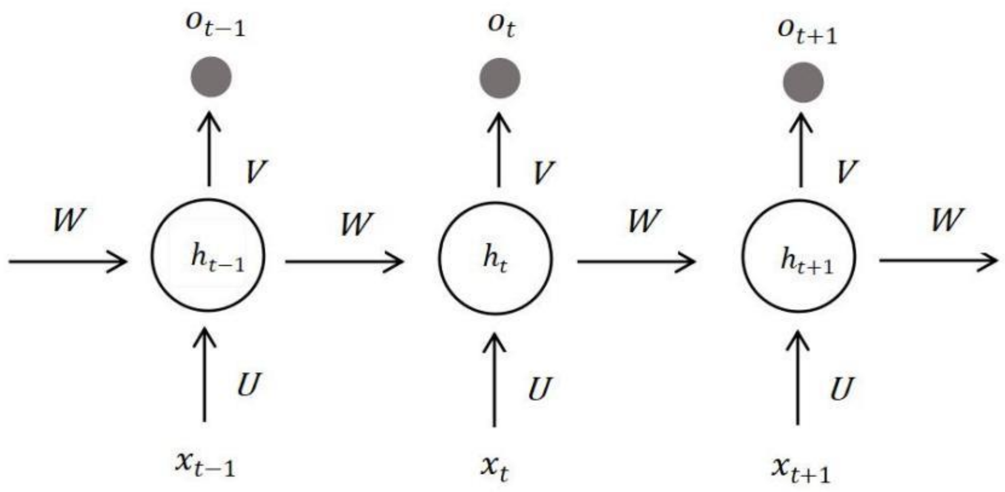

Fig. 1 The Structure of RNN

x_t: The input vector at time *t* (the input layer);
o_t: The output at time *t* (the output layer);
h_t: The memory at time *t* (the hidden layer);
W: Weight of input;
U: The weight of the input sample at this time;
V: The sample weight of the output

The RNN contains a feedback structure in the hidden layer. The hidden layer state of the previous moment can be passed to the hidden layer of the next moment through this structure, while it can combine with the external input variables as the input variables of the hidden layer at the current moment [4]. h_t is a memory unit to store the state of the hidden layer (i.e., the state of the previous moment). The basic formula is [5]

$$h_t = tanh(Wh_{t-1} + Ux_t) \tag{1}$$

where $tanh$ is the activation function in neural networks. The function of tanh is to filter information, while it is a nonlinear mapping. The output o_t is the result of the prediction with the memory of the current moment h_t. The equation is

$$o_t = f(Vh_t) \tag{2}$$

However, when training long sequences, RNN encounters the issue of gradient explosion and disappearance. [6]. To deal with this problem, Hochreiter & Schmidhuber put forward LSTM in 1997, which is a special kind of RNN [7]. LSTM, which has been applied in many fields, is suitable for analyzing and forecasting significant events with lengthy intervals and delays in time series. It can explain the relationship between the current data and the prior input data, save the state information from before the input network in its memory, and employ the previous state information to affect anticipated data and development trends [8]. LSTM and RNN have many differences:

RNN does not have cell state, whereas LSTM stores information in cell states.

$tanh$ is the only activation function in RNN. The $sigmoid$ and $tanh$ functions are LSTM activation functions.

RNN only deal with short-term dependencies, while LSTM can solve both short-term and long-term dependencies.

2.2 Implementation

Through applying LSTM method, stock price prediction can be solved better. The structure of LSTM is showed as Fig.2.

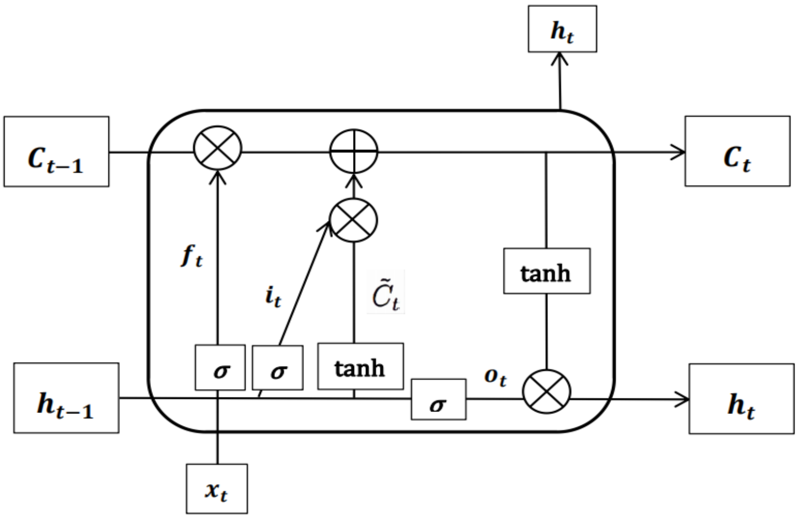

Fig. 2 The Structure of LSTM

LSTM is composed of a cell, an input gate, an output gate and a forget gate. What information should be filtered out is determined by the forget gate f_t. The input gate i_t determines which values go through the input gate to update memory status. The output gate o_t decides which components of the input and cellular memory are output [9]. The output of neurons belonging to the same t, which varies greatly, is stored in h_t. σ and $tanh$ are two activation functions, while σ outputs values in the range 0 to 1 and $tanh$ outputs values in the range -1 to 1. The equations and working process are described below [10, 11]:

Input gate can be formulated as,

$$i_t = \sigma(W_i * (x_t, h_{t-1}) + b_i) \tag{3}$$

$$\widetilde{C}_t = tanh(W_C * (x_t, h_{t-1}) + b_C) \tag{4}$$

Through the forget gate, some unimportant information is forgotten, and the other information will entry the input gate to determine what information need to be updated (i.e., i_t) and get the current state ($\widetilde{C}_t$, store the information in cell state).

$$f_t = \sigma\left(W_f * (x_t, h_{t-1}) + b_f\right) \tag{5}$$

$$C_t = i_t * \widetilde{C}_t + f_t * C_{t-1} \tag{6}$$

In output gate, after finishing the forget gate and input gate, long-term information (C_t) and short-term information (h_t) will be obtained. The last step is the storage operation and to input to the next neuron.

$$o_t = \sigma(W_o * (x_t, h_{t-1}) + b_o) \tag{7}$$

$$h_t = o_t * tanh(C_t) \tag{8}$$

The following flowchart (Fig.3) is the process based on LSTM in python, which can implement the stock price prediction. First, the data is processed into a tensor and then sends to the built LSTM model for training. The author sets the batch size to 16 and the size of hidden state is to 128. The model is trained using RMSprop optimizer with the learning rate of 0.001. The code is implemented using the popular deep learning framework, Pytorch, which achieves efficient learning using the GPU acceleration. After training, the author processes the test data with the same manner and output the future price of the stock. Finally, the author plots the predictions using matplotlib.

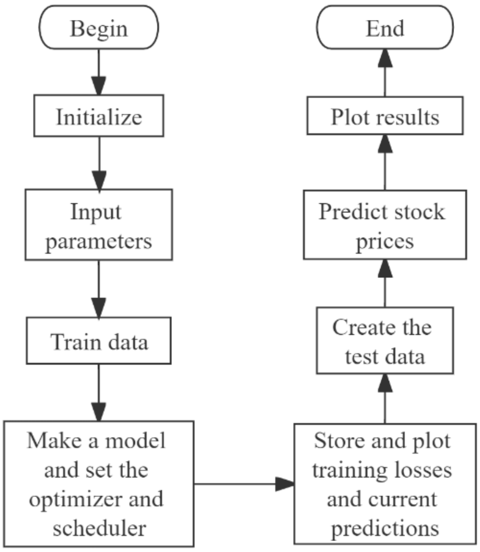

Fig. 3 A Flowchart Based on LSTM in Python

3. Results

The data from 01/01/2020 to 12/31/2021 is to train, and the data from 01/01/2017 to 12/31/2019 is to test. According to the process of the implementation above, "epoch" means the times of training data. MSE is the average of the squares of the errors. The table below are the results about MSE and epoch.

Table 1. The results about MSE and epoch

Epoch	MSE
1	0.84
50	0.10
100	0.08

In the meantime, the images of MSE and epochs are depicted as shown below. MSE is convergent as epoch changes.

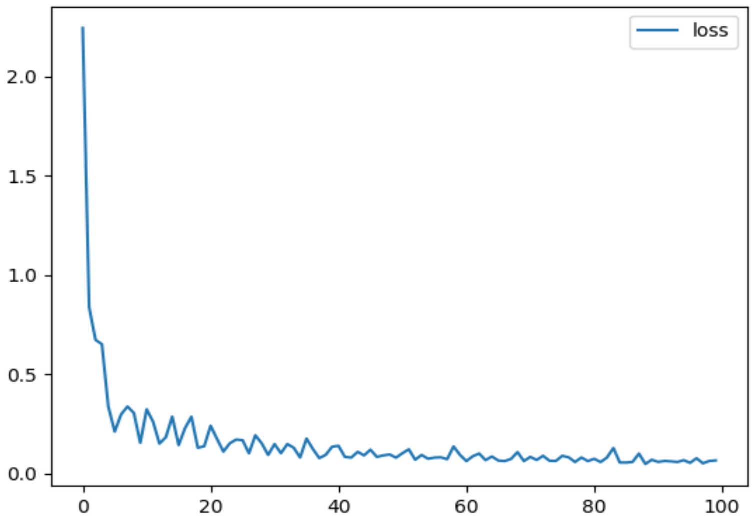

Fig. 4 MSE loss during training

Additionally, the expected value trained with epoch=50 is closer to the true value than with epoch=1 as Fig.5 and Fig.6.

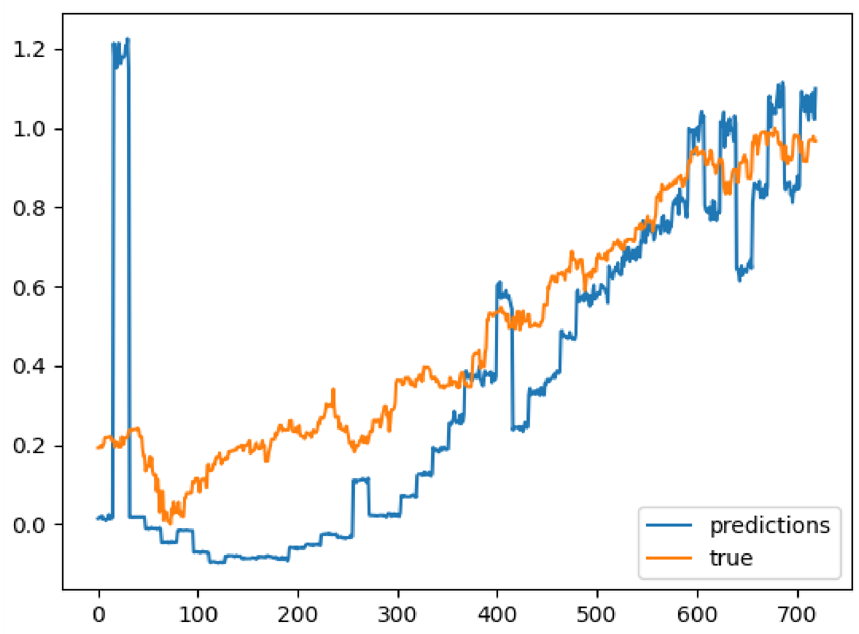

Fig. 5 Prediction results when epoch=1

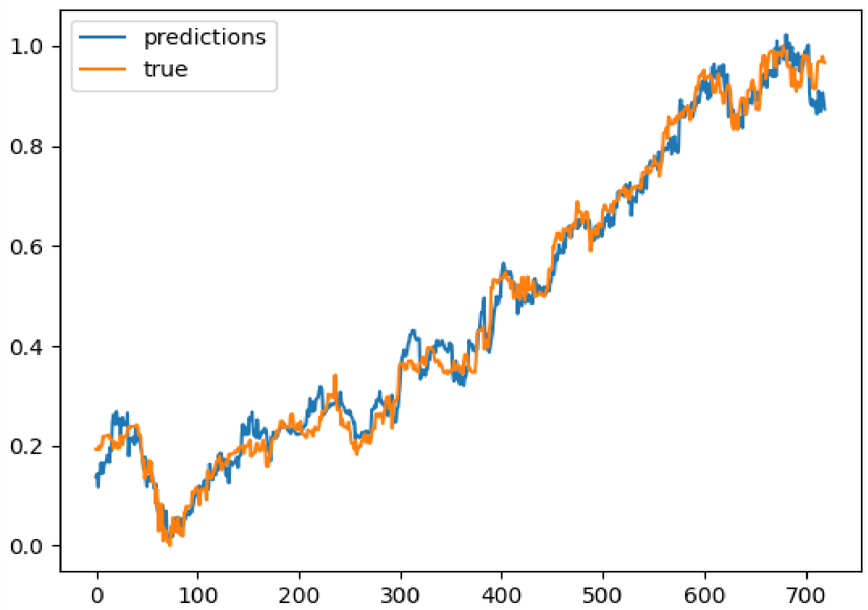

Fig. 6 Prediction results when epoch=50

As MSE is convergent and when epoch=100, MSE goes to 0, the graphs of stock price prediction can be shown directly as Fig.7, Fig.8 and Fig.9.

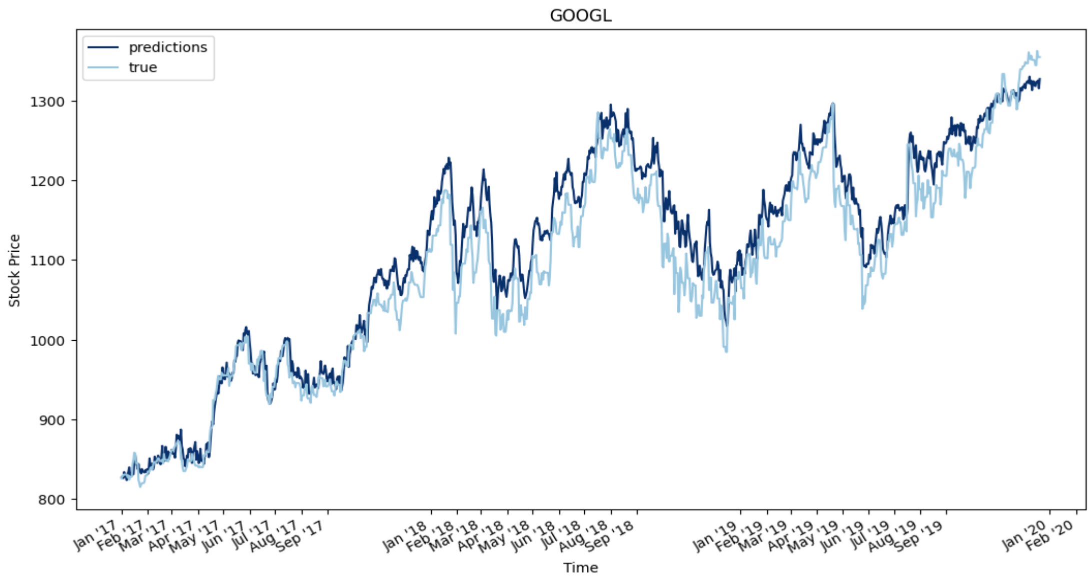

Fig. 7 Test results of GOOGLE

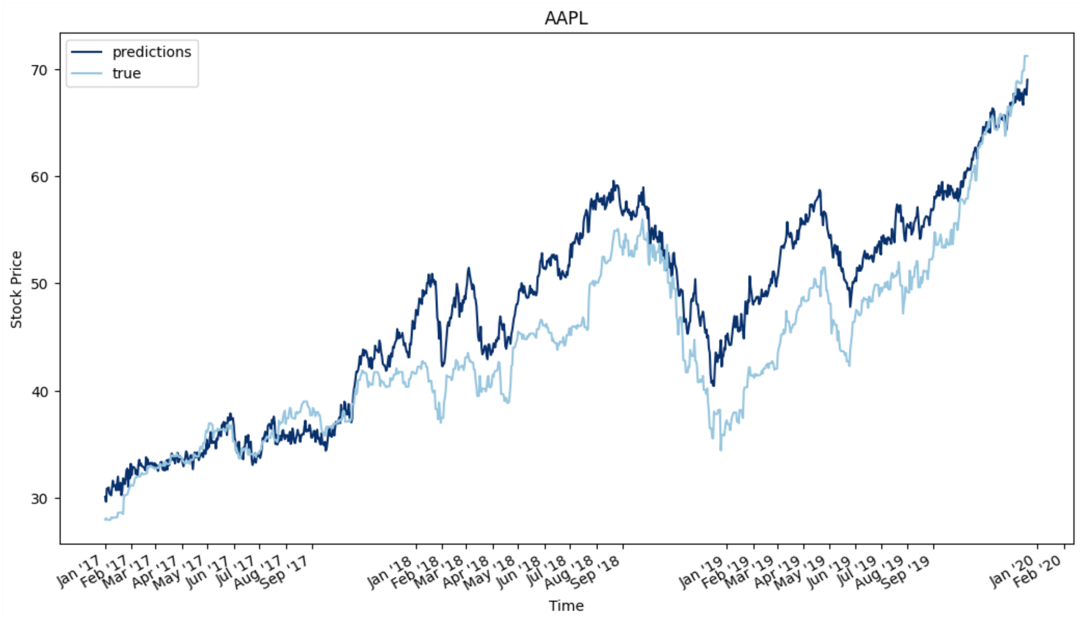

Fig. 8 Test results of APPLE

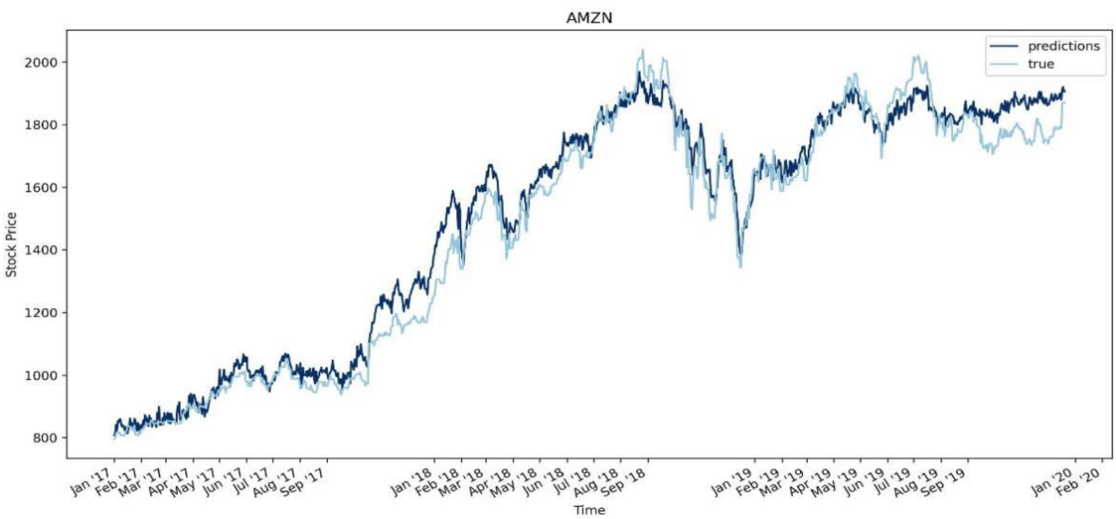

Fig. 9 Test results of AMAZON

Employing this approach, it can be seen that while stock prices occasionally match, the predicted value and the real value move in nearly the same way. However, there are still some problems during the implementation.

In the last three figures, the stock prices show a large error in some parts even though the trends are similar.

Every time of training has different results, like Fig.7 and Fig.10.

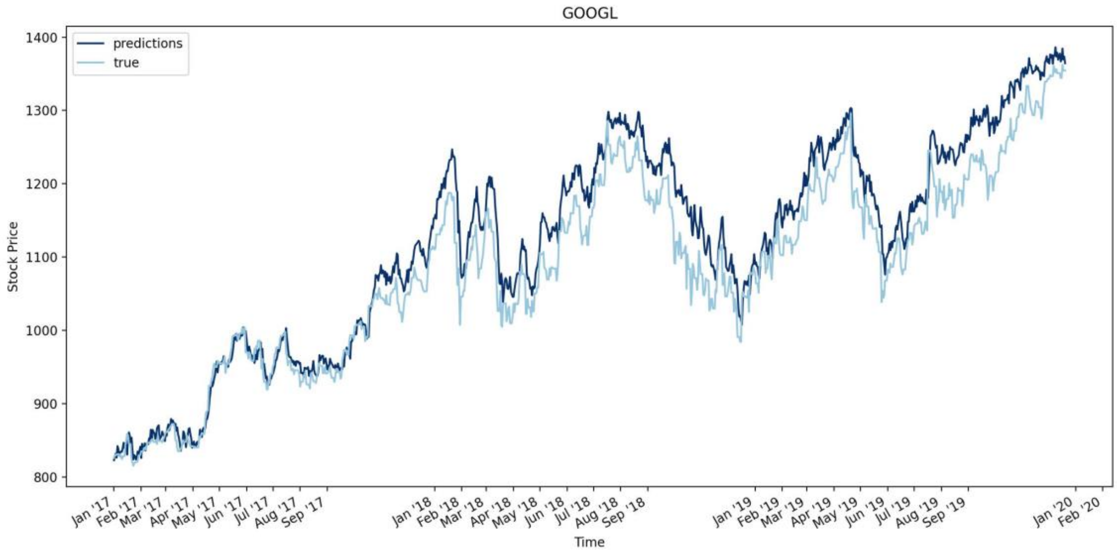

Fig. 10 The second time of prediction of GOOGLE

Therefore, the optimizer and scheduler need to be reoptimized to solve the first and second problem (i.e., reduce errors). Since the true value is fixed, errors should be minimized, which are caused by the forget gate during LSTM training. The last problem still needs to deal with. As C_t and h_t are derived from the C_{t-1} and h_{t-1}, the training slows down. Thus, simplifying some steps may accelerate the training speed of LSTM, which is worthy of further research.

4. Conclusion

In this paper, LSTM presents good applicability to predict stock prices. This research revolves around 5 perspectives. (1) The first part is to describe the working process and the difference of RNN and LSTM. (2) The second part is to use 8 steps in python to implement the process. (3) Through using data from S&P500 of the three companies and the MSE decreases to 0.08, according to the study, the real value and the anticipated value generally follow the same pattern. That means LSTM model is suitable for solving this kind of problem. (4) The final part is to find some problems during the implementation and give solutions. However, the author finds the program runs very slow. Since C_t and h_t are derived from the C_{t-1} and h_{t-1}, the training slows down. Thus, simplifying some steps may accelerate the training speed of LSTM, which is worthy of further research.

References

[1] Fu, K., Yin, X., & Wang, Q. (2020). Stock Trend Prediction of Multi-Feature Based on LSTM. Journal of Beijing University of Posts and Telecommunications(Social Sciences Edition), 22(5), 17–27.

[2] Cao, A., & Wu, M. (n.d.). Daily Stock Price Prediction Based on Long and Short Term Memory Neural Network. China Computer & Communication, 2022(1), 59–61.

[3] Li , L., Zeng, L., Jiang, S., & He, W. (2022). Stock Price Prediction Based on LSTM Neural Network. Journal of Yunnan University Nationalities(Natural Sciences Edition), 3.

[4] Bai, Y., Cheng, H., & Ye, Z. (2022). Application of Underactuated Process Control Method Based on RNN Compensator. Control Engineering of China, 1–10.

[5] Ding, F. (2022). Price Trend Prediction of an Engineering Product Based on RNN Model. Science and Technology Innovation, 07, 77–80.

[6] Jiang, B., Yue, J., & Lan, S. (2021). Research on Stock Prediction System Based on LSTM Neural Network. JOURNAL OF HEBEI INSTITUTE OF ARCHITECTURE AND CIVIL ENGINEERING, 39(4), 166–170.

[7] Lin, X., & Zhu, X. (2022). Attention-Mechanism-Based LSTM Model for Stock Price Predicting. J Chongqing Technol & Business Univ(Nat Sci Ed), 39(2), 76–82.

[8] Peng, Y., Liu, Y., & Zhang, R. (2019). Modeling and Analysis of Stock Price Forecast Based on LSTM. Computer Engineering and Applications, 55(11), 209–212.

[9] Zhang, X., Xu, X., & Wei,S.. (2020). Stock Prediction with LSTM Neural Networks. Computer Knowledge and Technology, 16(28), 39–43.

[10] Huang, Y., & Fang, W. (2021). Research on Stock Price Prediction Based on LSTM Network. Modern Computer, 27(34), 51–60.

[11] Zhang, K., Ye, C., Li, Z., & Wang, J. (2021). Analysis and Forecast of Stock Price Based on LSTM Model of Pytorch. COMPUTER TECHNOLOGY AND DEVELOPMENT, 31(1), 162–167.