

Stock Price Prediction using Deep Recurrent Model

Liang Zhen

Evans, University of Washington, Seattle, Washington, 98195, the United States

*Corresponding author's e-mail: liangzhensama@gmail.com

Abstract. In this paper, we aim to predict the stock trend using previous stock data. In specific, we perform LSTM training on the stock data of 2017 and 2018 by using the Sandp500 data as the base. The data includes the stock date, opening price, closing price, the highest price of the day, the lowest price of the day, and volume. Through these data we conducted 100 epoch training. After training, our LSTM can correctly predict the learning curve without any delay influence, with a satisfying performance. And we got the visual graph and chart to prove that the result is effective. The MSE value decreased from the original 3.026 to 0.238. This shows that the error of the result is small and the reliability is high. After LSTM training, it can help stock investment companies and related institutions and individuals to better avoid risks and provide individuals with high-credibility investment advice.

Keywords: DataSeries.cs, opening price, closing price, learning curve

1. Introduction

The stock market raises huge amounts of money for today's various businesses to help them grow while rewarding ordinary investors handsomely. The stock market in the traditional sense is that investors judge the future prospects of a company based on their own knowledge and experience to decide whether to invest in the company. With the emerging development of deep learning based method[1,2,3,4], many method with novel backbones are proposed, such as AlexNet[5], VGG[6], ResNet[7], DenseNet[8] and inception[9] network. However, with the development of computer technology and the continuous improvement of prediction programs such as LSTM, we can make some predictions about the future through historical data to better help us make various decisions. LSTM is long short-Term Memory, which is a kind of Recurrent Neural Network. It is a Neural Network designed to solve a series of time-based problems. We can use the LSTM network to find various types of certain stocks. Information, such as opening price, closing price, etc., to predict future stock movements. LSTM variants have also made extremely important contributions in other fields, such as impressive developments in language modeling. Use context to predict the next word through Neural networks. In this way, the performance of English grammar can be better improved through long-term training. in theory it allows the exact modeling of the probability of a word sequence. And compared to other networks, the training time will be reduced to achieve better results[10]. LSTM variants also play an important role in intelligent transportation, as cities grow and vehicles increase. Traffic problems in big cities have increasingly become a major factor that plagues contemporary people. On the one hand, the solution to this problem can be widened and increased roads by increasing investment, and on the other hand, intelligent transportation also has a broad enough application prospect. LSTM can analyze road conditions through a large amount of data to predict the best travel time and route to help users travel better[11]. LSTM is also a powerful tool in predicting Covid-19. The number of cases in the future can be predicted by using the large number of cases and time and the number of deaths in the past ten countries to achieve effective prevention. Stability and enhanced forecasting accuracy have been demonstrated after training and can be used for Covid-19 forecasting and effective planning and management.

2. Method[12]

2.1. Dataset Description

This article will obtain the data of the stock market in 2017 and 2018 through the Internet, including the opening and closing prices of each day, as well as the highest prices and volume. The data sample is listed in Table 1.

Table 1. Sample of the utilized data in the public dataset, which contains 6 items.

Date	Open	High	Low	Close	Volume
1980-12-12	0.10017843544483200	0.10061396968821300	0.10017843544483200	0.10017843544483200	469033600
1980-12-15	0.09538761096937120	0.09538761096937120	0.09495207667350770	0.09495207667350770	175884800
1980-12-16	0.08841833210513140	0.08841833210513140	0.08798280358314510	0.08798280358314510	105728000
1980-12-17	0.09016046673059460	0.0905960011109503	0.09016046673059460	0.09016046673059460	86441600
1980-12-18	0.09277442097663880	0.09320994944934960	0.09277442097663880	0.0927744209766388	73449600
1980-12-19	0.09843632578849790	0.09887186010128430	0.09843632578849790	0.09843632578849790	48630400
1980-12-22	0.1032271534204480	0.10366346687701400	0.1032271534204480	0.1032271534204480	37363200
1980-12-23	0.10758327692747100	0.1080188112785870	0.10758327692747100	0.10758327692747100	46950400
1980-12-24	0.11324515193700800	0.11368068614712000	0.11324515193700800	0.1132451519370080	48003200
1980-12-26	0.12369873374700500	0.12413426813772600	0.12369873374700500	0.12369873374700500	55574400

2.2. Introduction of LSTM

Long short-term memory (LSTM), is a special kind of RNN, which has better performance in longer sequences than traditional RNN LSTM. Compared with RNN with the single memory state H_t , LSTM use both cell state and hidden state to restore the previous memory, which can achieve a longer temporal correlation.

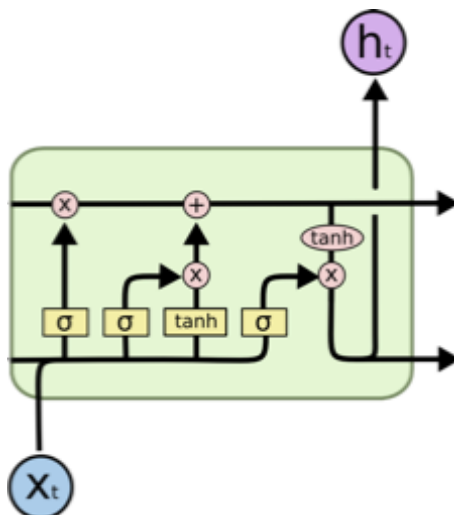


Figure 1. Framework of long-short term memory (LSTM) network. Resource from <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>.

The operation process of LSTM is divided into three parts, which are the forget process, the selection and memory process, and the output process. We will detail these three steps below.

Forget process: It is not important to decide what data is by controlling the previous formula C^{t-1} by the calculated z^f . The process of forgetting these unimportant data is called the Forget process

Memory selection process: Selective memory through x^t , important information will be remembered, and unimportant information will be remembered less. Z is obtained by calculation. Control z^i as gate.

Memory generation process: After generating the selected input and memory data. In memory generation process, the LSTM aims to generate the current hidden state and cell state by combining them, which can be summarized as

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (1)$$

3. Results and discussion

3.1. Comparison results

In the experiments, we train the LSTM with RMSprop optimizer and train during 100 epochs with the learning rate of 0.001. The results are shown in Table 2. Before training, the model has inferior performance, which only obtains 3.026 of MSE. After training our model achieves a much better performance with 0.238 of MSE. Furthermore, the learning curve also validates the effectiveness of the LSTM. The model is converged after 60 epochs.

Table 2. The MSE score of the proposed LSTM on the initial and final epoch.

Epoch	MSE
0	3.026
100	0.238

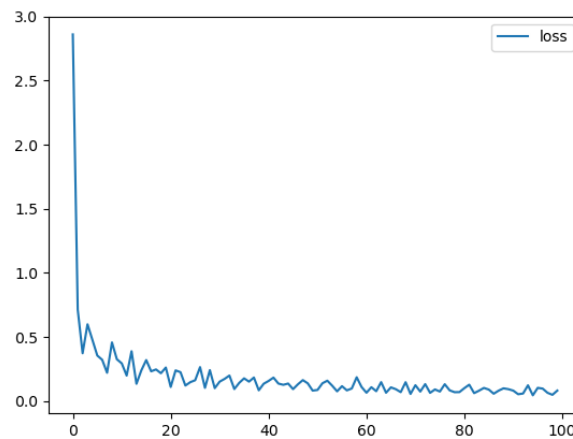


Figure 2. Loss curve of the proposed LSTM.

3.2. Visualization results

We also show the prediction results in Figure 3. When epoch =0, the trend of test results is not fit the ground truth. After training, our LSTM can correctly predict the learning curve without any delay influence, with a satisfying performance.

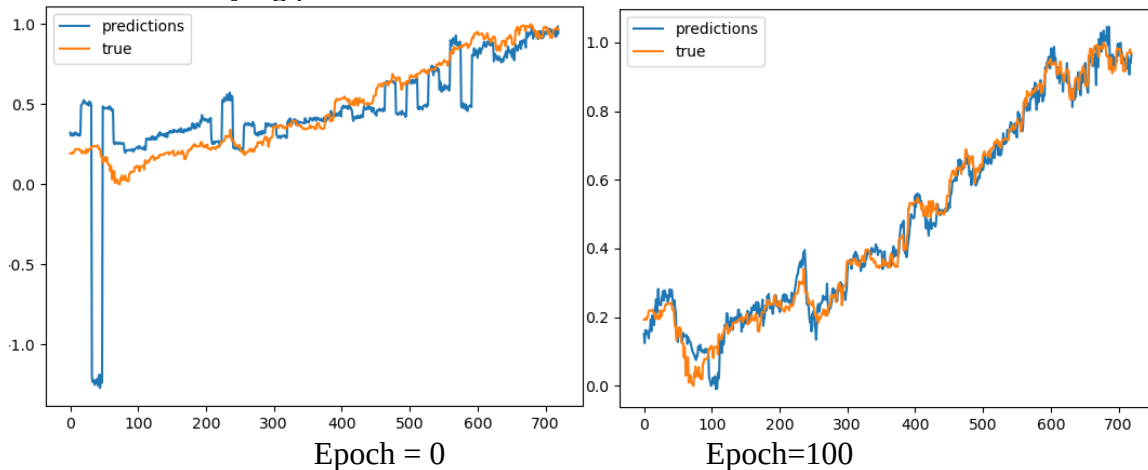


Figure 3. Visualization result between the proposed LSTM in 0th and 100th epoch.

We also evaluate the LSTM with the real stock data, we selected 2020's statistics from Google, Amazon and Apple stock price to validate the model's performance. From Figure 4, we all three data are correctly fit. In specific, stock from google and amazon are with the same trend even the same number, while the stock of Apple has some error points. This is because the prices are changed seriously and our model suffer a wrong prediction with the delay effect. In the future, we will use more robust feature to predict the stock price.

4. Conclusion

In this paper, we aim to exploit the stock price using deep learning model. We utilize the popular deep recurrent network, i.e., long-short term memory(LSTM) model for evaluation. By training on LSTM and Sandp500 data, we get a better stock forecast trend and get a relatively low Mean squared error. At the same time, the visualization results are given to prove the reliability of the results. Such results help investment institutions and individuals with risk aversion and investment advice, allowing them to make better money and avoid large-scale investment failures.

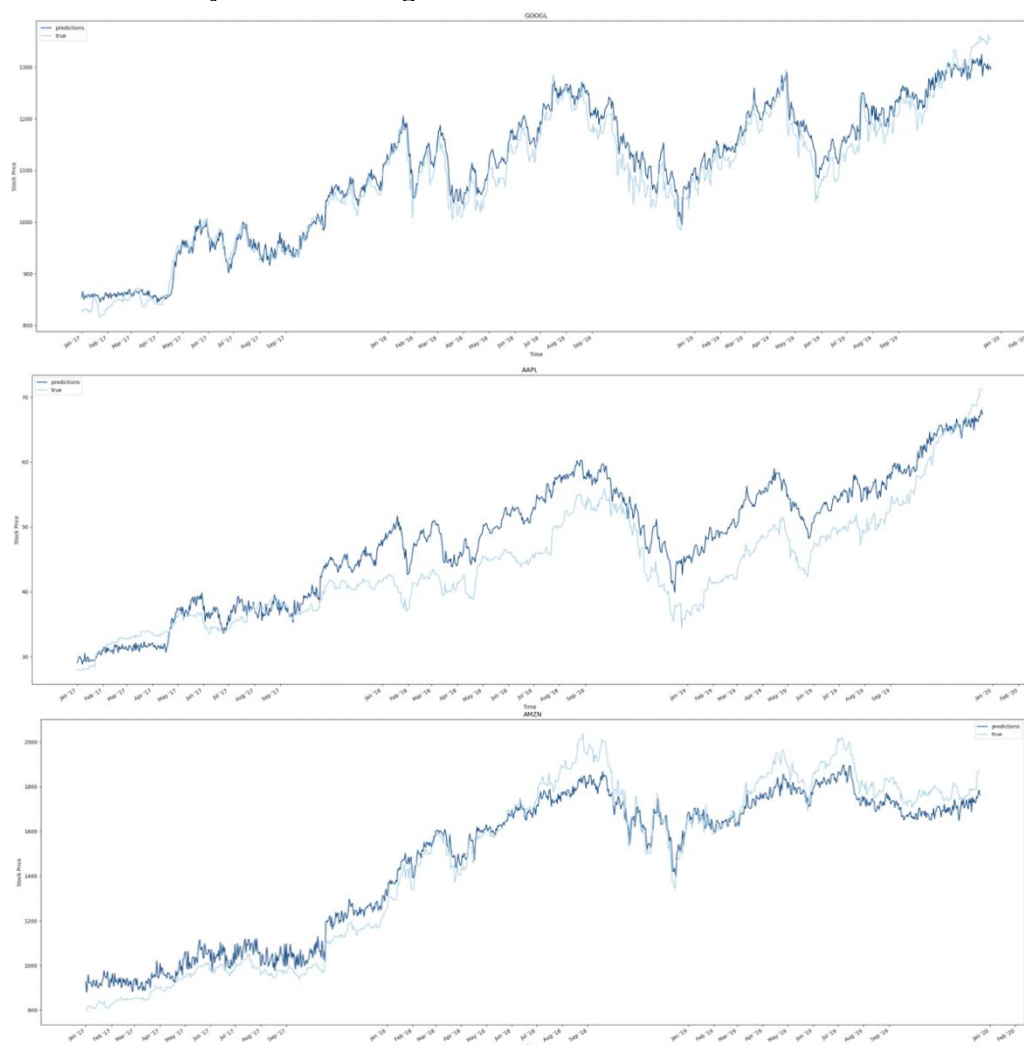


Figure 4. Visualization results on the test set of Google, Apple and Amazon stock

References

- [1] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
- [2] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.

- [3] Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., ... & Iyengar, S. S. (2018). A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 51(5), 1-36.
- [4] Shrestha, A., & Mahmood, A. (2019). Review of deep learning algorithms and architectures. *IEEE Access*, 7, 53040-53065.
- [5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
- [6] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [7] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778)
- [8] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [9] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).
- [10] Sundermeyer, M., Ralf S., and Hermann N.. "LSTM neural networks for language modeling." *Thirteenth annual conference of the international speech communication association*. 2012.
- [11] Shahid, F., Aneela Z., and Muhammad M.. "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM." *Chaos, Solitons & Fractals* 140 (2020): 110212.
- [12] Zhao, Z., Chen, W., Wu, X., Chen, P. C., & Liu, J. (2017). LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2), 68-75.