Stock Price Prediction using LSTM Model

Yifei Zhang*

Faculty of arts and science, University of Toronto, Toronto, Ontario, M5S 1A1, Canada

*Corresponding author's e-mail: 1294383039@qq.com

Abstract. Stock prices represent the development prospects of a company and play an important role in predicting future economic trends. With the development of deep learning, in this paper, our goal is to predict stock prices through deep learning algorithms. The author uses the Long-Short Term Memory (LSTM) network to make predictions by predicting the real stock prices of Google, Amazon, and Apple, and gives the visualization results of the training and testing stages. The results validate the effectiveness of the LSTM model.

Keywords: algorithms, predictions, forecasts, LSTM model.

1. Introduction

In the stock market, investors' subjective trading often has some defects. Investors can be swayed by their own emotions as well as by outside interference. In short, it is difficult for investors to maintain a consistent strategy in actual trading. Investors want machines to make trades that analyze past data to build models and predict future stock movements. This will allow investors to firmly implement their trading strategies and reduce the interference of other factors. 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. Although these approaches optimize machine learning from different perspectives and contribute to the stock forecast. For example, the AlexNet successfully solved the gradient dispersion problem of sigmoid when the network is deep and used GPU to accelerate the operation [5], and the Structure of the VGGNet is straightforward. Almost none of them focus on the effect of the time series length on the results. As the time horizon for predicting stock movements grows longer, the accuracy of the forecasts might decline. To sum up, this study will use LSTM suitable for multivariate time series prediction problems to conduct research, hoping to improve stock prediction accuracy through LSTM.

The stock represents the ownership of a listed company, and its price reflects investors' expectations for the company's future and constantly changes due to supply and demand. For example, investors judge the company's value from different angles and choose to buy or sell the stock according to their judgment to earn profits from the price difference. [10] The stock market is a market where stocks are bought and sold. It comprises companies, investors, and a securities exchange center, which essentially acts as a guarantee for transactions. Predicting the rise and fall of stock in advance and investing all your money in it is the fastest way to make a profit, but investors cannot expect the future price of a stock.

This essay will show how to use the machine learning algorithm LSTM (Long Short-Term Memory) to analyze stock data and predict stock prices, reflecting the contribution of the machine learning algorithm to the economic market. Since investors will integrate their emotions into the investment judgment during the investment process, investors can obtain more objective results after investment through LSTM analysis:

(1) This research will introduce the stock data involved in the research.

(2) The machine learning algorithm LSTM will be introduced.

(3) The article will shed light on the implementation details and demonstrate the predicted values' accuracy by calculating the RMSE.

2. Method

2.1. Dataset Description

This research uses daily stock data of Apple, Google, and Amazon from January 1981 to May 22 as a reference. The seven variables included opening prices, closing prices, highest and lowest prices, turnover, and stock dividends. The image below shows a sample of data.

Table 1. Data sample of the utilized public data.

Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
1981/1/2	0.120214455	0.121085523	0.120214455	0.120214455	21660800	0	0
1981/1/5	0.118036746	0.118036746	0.117601223	0.117601223	35728000	0	0
1981/1/6	0.112809631	0.112809631	0.112374097	0.112374097	45158400	0	0
1981/1/7	0.108018811	0.108018811	0.107583277	0.107583277	55686400	0	0
1981/1/8	0.105841092	0.105841092	0.105405569	0.105405569	39827200	0	0

The data starts at the opening price, set by investors making bids before the market opens and closing prices before the market closes. In the second place, the stock's turnover is decided according to investors' trading activity, the stock with frequent turnover is significant, and the stock with slow turnover is negligible. In order to increase the liquidity of the company's stock and to allow investors to buy and sell the stock more actively, the company conducts stock splits. To be specific, a stock split is the division of shares issued by a company into more shares so that each share can obtain a lower price and higher liquidity. A stock dividend is an interest paid to investors in the stock situation, a non-cash return. In general, these data are reflected more intuitively of the market's attention to the stock. For example, when the opening price of a stock is equal to the closing price, it indicates that the buying and selling of this stock are in balance. Therefore, we combine these data for comprehensive analysis and judgment.

2.2. Introduction of LSTM

LSTM is one of the representations of RNN [11]. Although RNN can obtain information at every moment, RNN is more influenced by the information at a recent time than the information at a longer time. Thus, Sepp Hochreiter and Jurgen Schmidhuber created the LSTM in 1997, which consists of individual memory cells [12]. This kind of temporal recurrent neural network avoids the influence of time on the results. In other words, information that is older but more important to the results will still have a more significant impact [14]. The core of LSTM is to filter information through the gate structure composed of the input gate, forget gate, and output gate, to maintain and update the state of the storage unit. Each storage unit has three Sigmoid layers and one TANH layer. The following figure shows the structure of an LSTM storage cell.

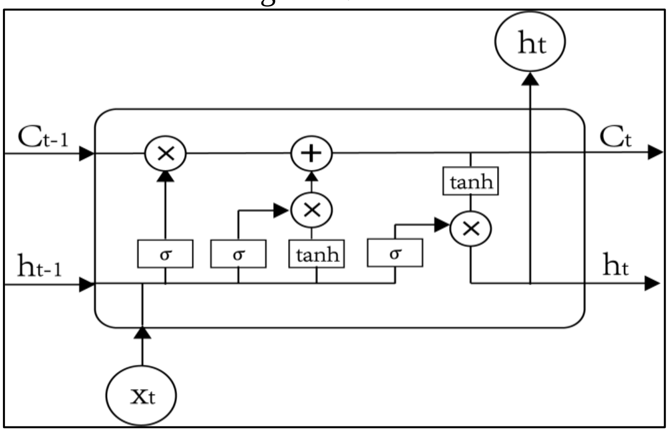

Figure 1. Framework of LSTM network.

Firstly, the function of forgetting gate in gate structure is to decide what information to forget. The forgetting gate determines how much to let CT-1 forget by output HT-1, input XT to this stage, and sigmoid. For example, sigmoid = 1 means that a large proportion of CT-1 is retained, while sigmoid= 0 means that the previous CT-1 is wholly forgotten.

As shown in the figure, [HT-1, XT] is the vector of the output at the last moment and the input at this moment in the memory cell. At the same time, WF and BF are the weight matrix and deviation of the forgetting gate respectively, and sigmoid is the sigmoid function.

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

Secondly, similar to the forgetting gate, the input gate combines HT-1 and XT to decide which information to keep through the sigmoid function. We call the retained information IT. It then creates a spare-CT and determines the contents of the spare-CT using the tanh function.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$-c_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

The magnitude of the impact on CT is determined by multiplying the two parts. Plus, the impact of the previous forgetting gate, it can be written as follows:

$$c_t = f_t * c_{t-1} + i_t * -c_t \tag{4}$$

Finally, we processed the input with sigmoid function, and then brought it into CT with tanh function to determine the output body, and multiplied it with the processed input to get the final output.

$$h_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) * tanh(c_t) \tag{5}$$

2.3. Implementation details

Memory footprint refers to the amount of memory currently consumed by the program. CPU (central processing unit) usage is the amount of CPU the programs use. More carefully, the CPU usage of one program impacts the speed of other programs. A program memory that is too large will affect the machine's overall performance. As shown in the figure, when running the program, the CPU usage of the computer fluctuated between 18% and 23%, and the memory usage reached approximately 1500MB. On the one hand, we can optimize CPU and memory usage by improving the performance of our test program. On the other hand, we can improve CPU and memory usage by optimizing the code itself.

3. Results and Discussion

We divided the dataset into two groups, took the model obtained from the experimental group into the control group, and summarized the reasonable prediction image. We first used the experimental group data for training, a total of 100 rounds of training. As seen from the figure, the loss value of training decreases continuously with increasing the number of training rounds.

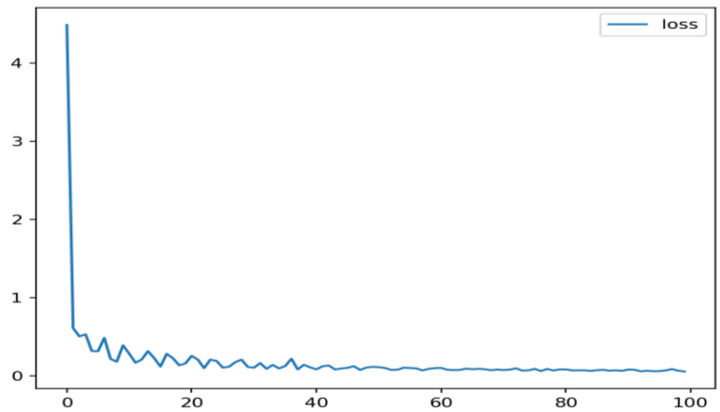

Figure 2. Training loss visualization of LSTM.

In Figure 3, the yellow line is the actual value, and the blue line represents the predicted value of the LSTM. As seen from the figure, the stock trend prediction on the model data set we proposed is consistent with the actual value of the trend, and the predicted value fluctuates around the actual value.

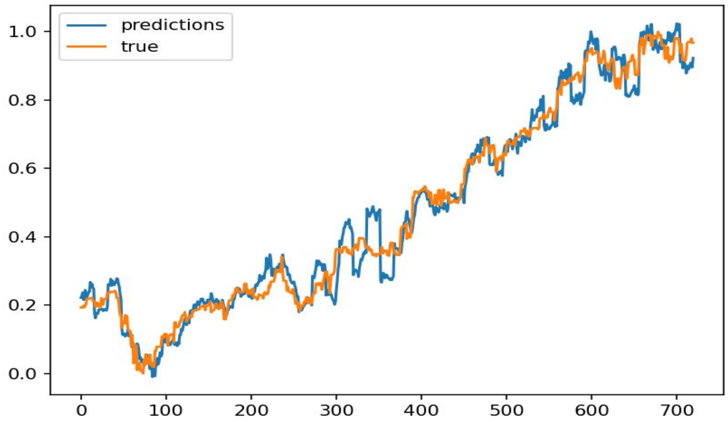

Figure 3. Visualization results of LSTM in training parse.

We process three stock datasets in the LSTM model. We compare the results and find that on the Google data set, the predicted value of our model is significantly better than the predicted value of the other two stocks.

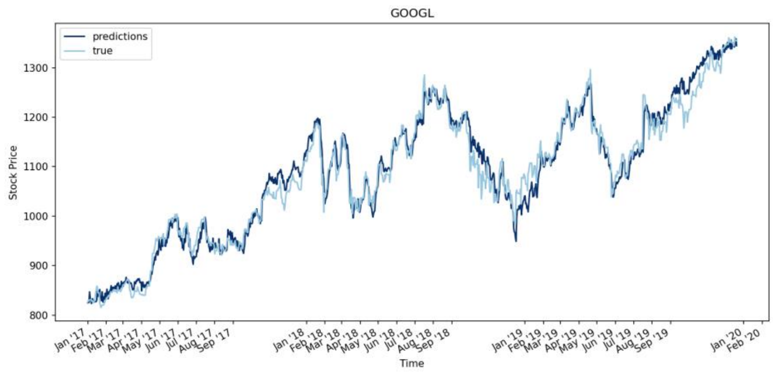

Figure 4. Visualization results of LSTM in testing parse (Google Stock).

There are significant differences in the prediction accuracy of different data sets, which may be due to the decrease of the prediction trend accuracy caused by the different accuracy of data, or the frequent price fluctuations may lead to the decrease of the prediction trend accuracy. Although the predicted value has some errors, the predicted trend is very close to the real value, which proves that our model is adequate.

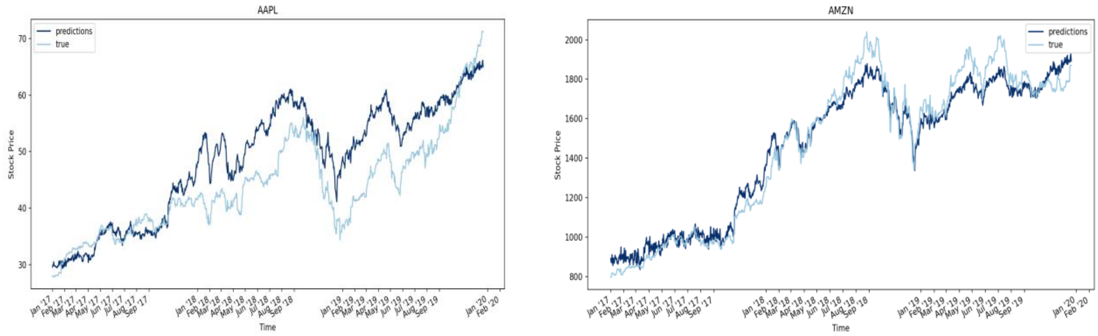

Figure 5. Visualization results of LSTM in testing parse (Apple and Amazon stock).

4. Conclusion

This paper develops a forecasting framework to simulate stock movements. LSTM neural network algorithm is used to predict the stock price changes, and good results are obtained. The experimental results show that the proposed LSTM model has a better fitting degree and more accurate prediction results. In quantitative trading, because it is a set of scientific methods, there is rigorous analysis and calculation, and data and models make decisions. Even simple, standard P/E investment methods can yield huge profits if they are strictly executed. LSTM neural network algorithm can almost perfectly fit the stock trend. Therefore, the model has a broad application prospect. Our research finds that the LSTM algorithm has a strong prediction effect when using relatively objective data. However, it is not enough to only consider historical data to predict stocks easily affected by various aspects[15]. For example, stock prices may fluctuate on a given day due to relevant policies and news, which cannot be observed using historical data. In future research, we hope to add subjective factors such as stock-related news and policies, which allows the prediction model to adapt to changes and make accurate judgments.

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] Kimberly A. (2021) How the stock market works. https://www.thebalance.com/how-does-the-stock-market-work-3306244.Mettam, G.R., Adams, L.B. (2009) How to prepare an electronic version of your article. In: Jones, B.S., Smith, R.Z. (Eds.), Introduction to the Electronic Age. E-Publishing Inc., New York. pp. 281-304.

[11] Abedinia O, Amjady N, Zareipour H.A new feature selection technique for load and price forecast of electrical power systems[J].IEEE Transactions on Power Systems, 2017,32 (1):62–74.

[12] Lecun Y, Bengio Y, Hinton G. Deep learning.[J]. Nature, 2015, 521(7553):436. pmid:26017442

[13] Qiu, J., Wang, B., & Zhou, C. (2020). Forecasting stock prices with long-short term memory neural network based on attention mechanism. PloS One, 15(1), e0227222–e0227222. https://doi.org/10.1371/journal.pone.0227222.

[14] Xu T, Zhang J, Ma Z, et al. Deep LSTM for Large Vocabulary Continuous Speech Recognition[J]. 2017.