LSTM Neural Network in Stock Price Prediction

Kejia Zhang(✉)

Department of Computer Science, University of Bristol, Bristol BS8 1TH, UK
zhangkejia2019@163.com

Abstract. Being an important part of the national economy, the stock market has always performed an important role in economic development. Following the growth of the times and variations in the investment philosophy of people, an increasing number of people have become involved in the stock market and predicting stock prices has become a popular topic. Stock prices are a kind of time series data, it has a large amount of data, high variability, high noise, and volatility, thus the traditional statistical analysis methods cannot properly capture the characteristics of these non-linear data and resulting in poor prediction accuracy. LSTM is a type of recurrent neural network, and it has become an effective learning model to deal with time series data. This paper will focus on LSTM neural networks and investigate the effect of the memory length for past data on the accuracy of the model prediction. In the experiments, the model prediction results will be compared using data from the past day, past week, past month, and the past year as inputs. The results show that the best predictions are made using short-term past data as input, particularly when using data from the past day as input.

Keywords: machine learning · stock price prediction · long short-term memory · neural network

1 Introduction

The stock market has a significant impact on the interests of individual investors and the national economy. Therefore, stock price trend prediction has become a popular topic in the stock market. It can help people to judge the current situation, which can reduce the risk of loss and increase returns. However, forecasting stock prices is a complicated task as it does not follow any pattern. Stock prices change purely based on supply and demand over a period.

Many methods have been shown to be applicable to stock price forecasting. Fundamental analysis is a type of investment analysis used by investors when making investment decisions. They attempt to estimate the intrinsic value of a stock by examining a range of economic factors such as a company's sales volume, earnings, and profits. This method is more useful for long-term investors [1]. Additionally, technical analysis is another type of investment analysis. This method attempts to predict the future price of a stock based on the study of its past price action, which focuses on the price of the stock

J. Yen et al. (Eds.): ICBIS 2023, AHCS 14, pp. 848–856, 2023.
https://doi.org/10.2991/978-94-6463-198-2_87

and ignores all fundamental information about the company [1, 2]. This analysis method was usually suitable for short-term investment [2]. Another type of investment analysis is time series forecasting, where time series data is analysed to predict the future based on trends in past data. For example, regression moving averages (ARMA) and multiple regression models usually find trends in past data to predict future data. However, if a stock has little historical data or is in a period of rapid change, accurate analysis and forecasting can be difficult.

Therefore, neural networks are a method to solve the above situation. Neural networks can adapt to noisy data and establish input-output relationships for non-linear data, thus making stock price prediction possible. There is a significant amount of research work has been done trying to predict stock prices through neural networks. For example, LSTM has been shown to be applied to stock prediction and the predictions are accurate [3].

Long Short-Term Memory (LSTM) is one of the most famous series of predictive networks, which is a type of recurrent neural network (RNN) that can remember information for a long time, which makes them well suited to processing time-series data. This type of network can process the data incrementally and maintain an internal state, which caches the information that it has seen to date in a summarised manner.

This paper will discuss the use of LSTM neural networks for stock price prediction. In this experiment, there is an LSTM model will be constructed. The prediction performance of the LSTM model will be compared across different company stock price datasets, with different training and prediction durations set for comparison. There is also some discussion on the effect of memory length from past data as input to the LSTM model for predicting stock prices.

2 Literature Review

2.1 Stock Price Prediction with Deep Learning

In recent research, many deep learning approaches were proven can be used in prediction, such as Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN).

In 2003, Zhang combined the autoregressive integrated moving average (ARIMA) with ANN and used this hybrid model to forecast three different types of time series, with the results showing that the ANN was more advantageous in dealing with non-linear data [4]. After that, in 2017, Rout et al. applied a low-complexity artificial neural network for stock price prediction on the Bombay Stock Exchange dataset and showed that the ANN performed better in predicting stock prices compared to other methods [5]. In the same year, Selvin et al. compared the CNN-sliding window model with other neural networks and showed that the CNN-sliding window model gave better results because it only used the current window to predict future prices [6].

In 2018, Achkar et al. compared the prediction accuracy of LSTM with that of BPA for stock prices and found that LSTM was more accurate [7]. In addition, in 2020, Mehtab and Sen compared the prediction accuracy and prediction speed of CNN and LSTM. Although CNN predictions were faster, LSTM was expected to be more accurate [3]. In summary, LSTM provides a stronger performance in stock prediction compared to other prediction methods. Therefore, this paper will choose the LSTM neural network as the experimental model.

2.2 Long Short-Term Memory (LSTM)

LSTM was proposed by Hochreiter and Schmidhuber in 1997 and has been shown to handle time series problems better than Elman nets, real-time recursive learning, neural sequence chunking, recursive cascade correlation, and backpropagation through time [8]. In 2015, Chen et al. used LSTM to forecast Chinese stock market returns and the results show that LSTM has significant potential for stock forecasting [9]. After that, in 2020, Guo combined LSTM with Natural Language Processing (NLP) for analysing the news headline and text and train the LSTM model by using these mixed features. The empirical results indicate that combining the usage of characteristics from the news can effectively increase the accuracy of stock prediction by LSTM models [10]. In the same year, Yu and Yan used a DNN neural network that was designed based on LSTM neural networks and the time series phase-space reconstruction (PSR) method, which showed high prediction accuracy on the stock market index dataset [11].

According to the above, the previous study focused on how to improve the prediction accuracy of the LSTM, but they did not notice much about the effect of the memory length of the LSTM on the prediction accuracy. This paper will focus on comparing the prediction accuracy by using the stock close price data with different memory lengths.

3 Method

3.1 Model Selection

Based on the above analysis of past research, the LSTM is one of the neural networks that are suitable for the task of predicting stocks. Therefore, this paper uses the LSTM model as the experimental model and object. LSTM is a special type of RNN neural network.

As shown in Fig. 1, a conventional RNN is unfolded by several identical units connected consecutively. However, the actual structure of the RNN is the one on the left of Fig. 1, which can be found to be a self-constantly cyclic structure. In other words, as the input data keeps increasing, the above self-looping structure passes the last state to the current input, together with the new input data for the current round of training and learning, until the input or training is finished, and the final output is the final prediction

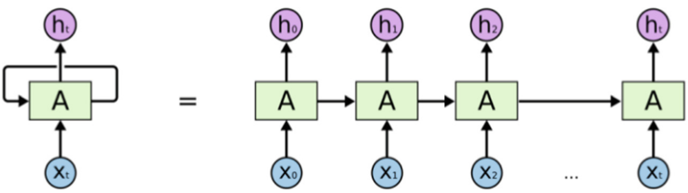

Fig. 1. RNN Structure [12]

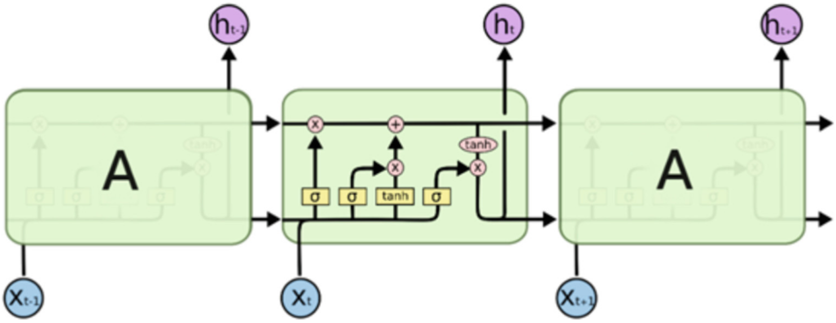

Fig. 2. The structure of the LSTM neural network [12]

result. Compared to the traditional RNN neural network, the LSTM has a memory cell to determine whether the information is useful or not. The memory cell is composed of an input gate, an output gate and a forget gate. Compared to a conventional RNN where there is only one state inside a single loop structure, the LSTM has three gates, which are the input gate, forgetting gate and output gate, inside a single loop structure and a persistent cell state is passed on between the loop structures to determine which information to forget or to pass on. The input gate measures the quantity of inputs to the network at the current moment that are stored in the unit state. The forget gate is responsible for the decision of whether to keep the cell state of network from the previous moment to the current moment. The output gate relates to the amount of cell state that is exported to the current outcome of the LSTM model. The structure of LSTM is shown in Fig. 2.

3.2 Data Selection

The stock price data in this experiment was downloaded from Yahoo Finance website [13]. For reducing bias, the stock data used in this experiment are from a similar type of technology companies, which are Apple and Microsoft. All stock price of each company is the daily stock close price from July 2010 to August 2021.

3.3 Data Processing

The stock price dataset is normalized by using the MinMaxScaler function from the sklearn library. After the predicted values are derived, they are processed by using MinMaxScaler again to reverse the normalization. The reason for using normalization is that it puts all the data in the same range to reduce the gap between data, which will improve the speed and accuracy of the training of the model.

The LSTM network requires the data to be imported as a 3D array. For the purpose of converting two-dimensional arrays into three-dimensional arrays, the normalised data will be looped through, and smaller partitions will be created based on the time step, which is one-day, one-week, one month and one year. Then, they will be regrouped into the final sample dataset. As only the closing price of the stock is considered in this experiment, each partition has only one feature value. The final sample data will be separated into two sets, one set to train the model and another set to test the model. The training set includes the stock price data from 29/06/2010 to 26/03/2019, and the test set includes the stock price data from 27/03/2019 to 19/08/2019. The period of training dataset and test dataset are the same for each stock price dataset of company.

3.4 Construct Model

In this experiment, the entire model was constructed with the Tensorflow library. Figure 3 demonstrates the architecture of the model that was used in this experiment. The experimental model has three LSTM layers, three Dropout layers and a Dense layer. Each LSTM layer will be followed by a Dropout layer, and the Dense layer will be used as the final layer to output the predictions. The first LSTM layer will accept the shape of the training data set and it will output 70 units of neurons as the output. The Dropout layer will then randomly drop 30% of neurons and take the remaining input into the second LSTM layer. After the second layer accepts the results of the previous layer, the 70 unit neurons will be recalculated as the output. Then, the Dropout layer will randomly drop by 30% and continues to feed the remainder into the third LSTM layer. When the third layer has been calculated, the third Dropout layer will accept the output of the third layer and randomly drop it by 30%. The remaining neurons will then be fed into the final layer, called the Dense layer, which will output one neuron, which will be the final predicted result.

4 Experimental Result and Analysis

This experiment will use data from Apple (AAPL) and Microsoft (MSFT) and use the model that was presented in the previous section as the object of the experiment. This experiment will be conducted to examine the relationship between the LSTM model and the amount of input past data, i.e. the memory length. Therefore, this experiment was designed by changing the time step of the input data as the experimental variable and outputting the predicted next daily stock price, with all other parameters remaining the same. Because the stock data is time series data and for the reliability of the experiment, the data set will not be disrupted, and all data will be trained for 500 rounds using the same model with the same parameters.

The following images show the results of this experiment, with the red line being the real stock price and the blue line being the predicted stock price. For clarity of the graphs, only 150 days from the test dataset have been taken as displayed in the graph. For comparisons on the same data set for the same company, the corresponding times, and positions in the data set after data processing are different because the time steps were chosen differently. Even with the 150 days chosen for their position in the same training set, their corresponding times are different. This is the reason that the real stock price lines are not the same on the graph with the same data set.

According to the experimental results, it can be found that the model remembers more data from the past and does not predict the next daily data properly when a longer time step is chosen. For example, the model has almost no predictive power when the time

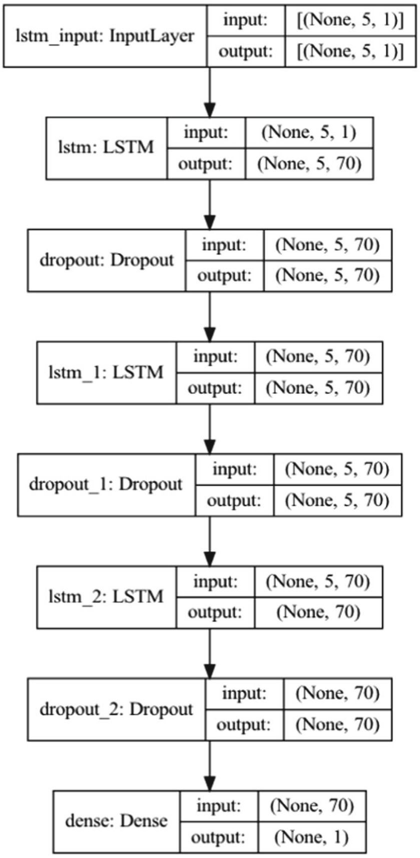

Fig. 3. Model Structure

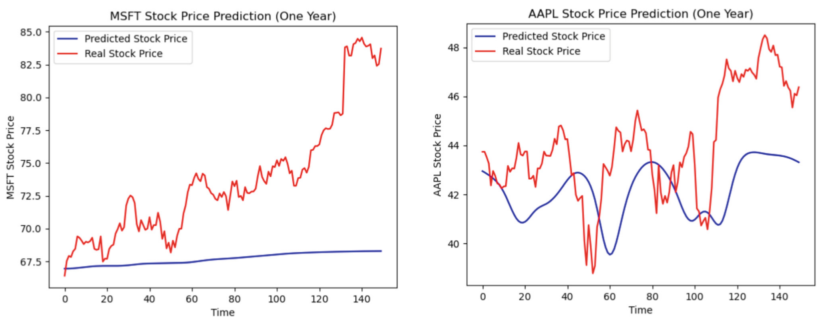

Fig. 4. One-year time step

step was one-year and one-month, which was displayed in Fig. 4 and Fig. 5. However, as the length of the time step decreases, the predicted stock price line gradually approaches the real stock price line. Figure 6 is an example, when the time step is one-month, the

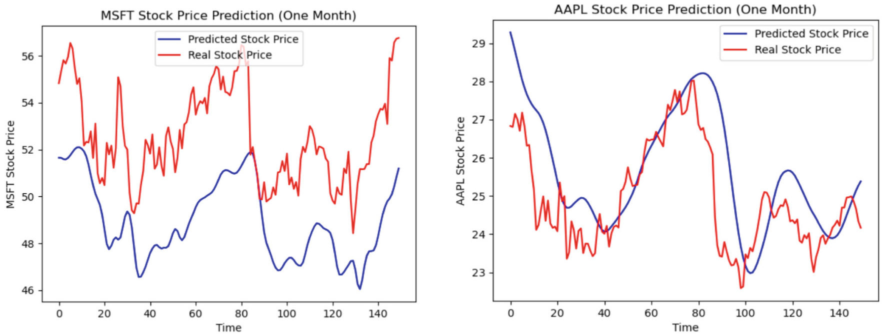

Fig. 5. One-month time step

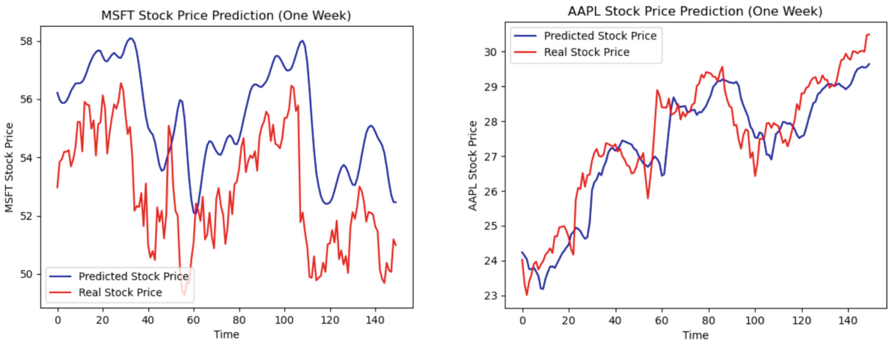

Fig. 6. One-week time step

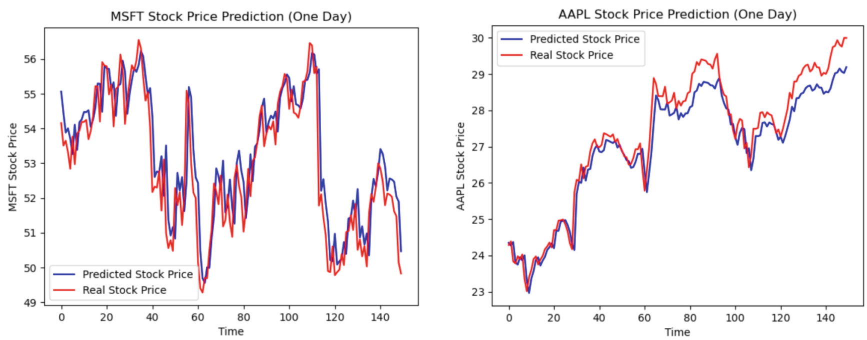

Fig. 7. One-day time step

overall trend of the predicted price line is essentially the same as the overall trend of the real stock price line. The best result is when the time step is one-day, which is the predicted stock price line and real stock price line almost coincide (Fig. 7). In addition, the delay is also an important factor in predicting the stock price. The results of this experiment show that the delay in prediction is also related to the time step. The images of the results show that as the time step became shorter, the number of inflection points in the forecast line increased and each inflection point became closer to the inflection point of the real price data, meaning that the model was able to better capture the timing of the price change. However, although the inflection points of the predicted price line are very close to the real price inflection points for the model using a one-day time step, there is still a delay, such as the MSFT result in Fig. 7.

5 Conclusion

In summary, this paper first summarizes past approaches to predicting stocks and highlights neural networks as an effective method. Then, based on the other research, the LSTM neural network was proven it was suitable for predicting time sequence data. Thus, the experiment in this paper constructed a three layers LSTM model and apply it to Apple and Microsoft stock price datasets with different time steps. The results of the experiments were plotted as several line graphs for comparison. It was found that as the time step, i.e. the amount of input past data, decreased, the predictions from the LSTM model were closer to the real stock price of the stock. Furthermore, when shorter time steps are used, the delay of the prediction from the model also was improved significantly. Although the results obtained using the precious daily data as input are close to the real stock price, there is still a delay.

In future research, it is possible to try to improve the problem of delayed stock price prediction by combining the use of temporal data, such as dates, in training LSTM models, or by replacing the dataset with a near real-time or smaller time unit dataset, such as minute stock price data, for stock price prediction.

References

1. Devadoss, A.V. and Ligori, T.A.A., 2013. Forecasting of stock prices using multi-layer perceptron. International journal of computing algorithm, 2(1), pp. 440-449
2. Vanstone, B.J. and Finnie, G.R., 2006, August. Combining technical analysis and neural networks in the australian stock market. In Artificial Intelligence and Soft Computing (pp. 126-131)
3. Mehtab, S. and Sen, J., 2020, November. Stock price prediction using CNN and LSTM-based deep learning models. In 2020 International Conference on Decision Aid Sciences and Application (DASA) (pp. 447-453). IEEE
4. Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, pp. 159-175
5. Rout, A.K., Dash, P.K., Dash, R. and Bisoi, R., 2017. Forecasting financial time series using a low complexity recurrent neural network and evolutionary learning approach. Journal of King Saud University-Computer and Information Sciences, 29(4), pp. 536-552
6. Selvin, S., Vinayakumar, R., Gopalakrishnan, E.A., Menon, V.K. and Soman, K.P., 2017, September. Stock price prediction using LSTM, RNN and CNN-sliding window model. In 2017 international conference on advances in computing, communications, and informatics (icacci) (pp. 1643-1647). IEEE
7. Achkar, R., Elias-Sleiman, F., Ezzidine, H. and Haidar, N., 2018, August. Comparison of BPA-MLP and LSTM-RNN for stocks prediction. In 2018 6th International Symposium on Computational and Business Intelligence (ISCBI) (pp. 48-51). IEEE

8. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in Neural Computation, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997. https://doi.org/10.1162/neco.1997.9.8.1735
9. Chen, K., Zhou, Y. and Dai, F., 2015, October. A LSTM-based method for stock returns prediction: A case study of China stock market. In 2015 IEEE international conference on big data (big data) (pp. 2823-2824). IEEE
10. Guo, Y., 2020, November. Stock price prediction based on LSTM neural network: the effectiveness of news sentiment analysis. In 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME) (pp. 1018-1024). IEEE
11. 12. Yu, P. and Yan, X., 2020. Stock price prediction based on deep neural networks. Neural Computing and Applications, 32(6), pp. 1609-1628
12. Yan, S., 2015. Understanding LSTM networks. (Online). (Accessed on August 11, 2022)
13. Yahoo Finance - Stock Market Live, quotes, Business & Finance News! Available at: https://finance.yahoo.com/ (Accessed: November 22, 2022)