Prediction of Stock Price Based on LSTM Neural Network

Dou Wei
College of Management and Ecomomics
Tianjin University
douweidw1113@163.com

***Abstract*—This study, based on the demand for stock price prediction and the practical problems it faces, compared and analyzed a variety of neural network prediction methods, and finally chose LSTM (Long Short-Term Memory, LSTM) neural network. Then, through in-depth study on how to predict the stock price by the LSTM neural network optimized by MBGD algorithm, the feasibility of the method and the applicability of the model are analyzed, and finally the conclusion is drawn. It is found that historical information is very important to investors as the basis of investment decisions. Past studies have used opening and closing prices as key new predicators of financial markets, but extreme maxima and minima may provide additional information about future price behavior. Therefore, the index of three representative stocks in China's stock market are selected as the research objects, and the key data collected from them include the opening price, closing price, lowest price, highest price, date and daily trading volume. The results show that although LSTM neural network model has some limitations, such as the time lag of prediction, but with attention layer, it can predict stock prices. Its main principle is to discover the role of time series through analyzing the historical information of the stock market, and to deeply explore its internal rules through the selective memory advanced deep learning function of LSTM neural network model, so as to achieve the prediction of stock price trend.**

Keywords—prediction of stock price, LSTM, neural network

I. INTRODUCTION

The stock market is the place where the stocks are transferred, traded and circulated. On the one hand, the issuance of stock provides a legal and reasonable channel for capital flow, which enables a large amount of idle capital to be gathered in the stock market. Such effective accumulation of capital can improve the organic composition of enterprise capital and greatly promote the development of economy. On the other hand, the circulation of stock enables the capital to be collected effectively and the accumulation of capital is effectively promoted. Based on this, the stock market is generally regarded by scholars from all walks of life as an intuitive reflection of the economic development of a country or region in a certain period. One of the main reasons lies in the stock market trading prices can objectively reflect the stock market supply and demand relations. Moreover, the stock market is often regarded as an indicator of stock prices and quantities. However, due to the complexity, variability and uncertainty of the stock market, the stock price formation mechanism presents the characteristics of complexity and unpredictability. Stock prices not only from the political, economic, market, technology, and investor behavior aspects such as individual factors, influenced by various factors in the interaction between the role at the same time, these will lead to changes in stock prices, the existence of the various uncertain factors lead to the complexity of the stock price changes. It is the constant changes in stock prices that provide the breeding space for speculation in the capital market and increase the risk of the stock market. Investors and analysts individuals tend to be "irrational", relying on personal experience and intuition to make decisions and determine there is a certain limitation, namely with experience and intuition to forecast the stock price is not accurate, this is not accuracy under the guidance of related behavior caused greater risks, could lead to economic losses to investors. Therefore, how to accurately analyze, judge and predict stock prices for investors to make decisions is very critical [1-10].

Deep learning based on neural network has attracted extensive attention from scholars in the field of deep learning. Neural network is a complex nonlinear dynamic system, therefore, what other method can't deal with low efficiency of complex and nonlinear system can be through neural network method, this method is characterized by handling mechanism can highly parallel, topological structure performance is very flexible, operation ability is extremely powerful, nonlinear operation faster and stronger ability of self learning organization. This method has been paid attention to in the research field.

The following section are as follows: section II gives a brief description of proposed framework. In section III experimental result is presented. A conclusion is also included in section IV.

II. PROPOSED FRAMEWORK

This study proposes an attention-based long short-term memory model to predict stock price trend. The model consists of four parts: input layer, hidden layer, attention layer, output layer . The input layer cleans the input data to meet the input requirements of the model. The hidden layer is connected to the line model network through LSTM unit. The attention layer weighted the feature vector. The output layer gets the calculated results. The model training is solved by gradient descent algorithm. The proposed framework is illustrated in Fig. 1.

A. Input Layer:

(1) take the date, closing price, opening price, maximum price and minimum price of the stock as input data to form a time series; (2) split the input data into training set and test set according to the ratio of 7:3; (3) convert each component of input data into the interval [0,1] after standardization.

978-1-7281-4691-1/19/$31.00 ©2019 IEEE
DOI 10.1109/AIAM48774.2019.00113

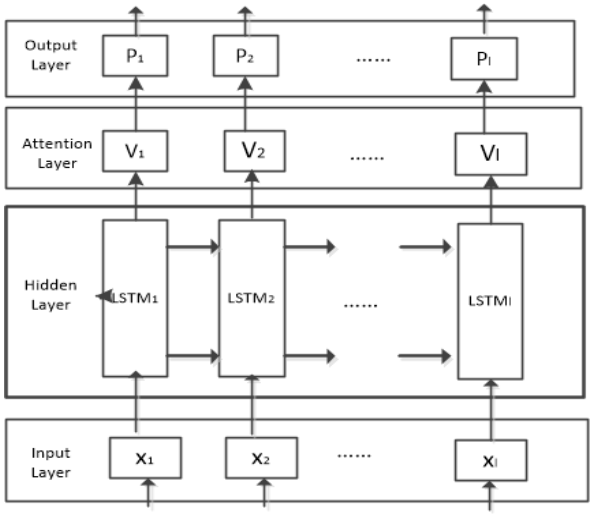

Fig. 1. Proposed framework

B. Hidden Layer

The hidden layer is formed by the LSTM unit, which is affected by the input data of the current moment and the previous moment.

C. Attention Layer

By calculating the weights of the input data in the attention layer, the model can select and learn the input data. The higher the weights obtained by the model training, the closer the input data is to the target value.

D. Output Layer

After the model training is completed, stock time series is input for prediction, that is, stock data is input for N days to predict the stock trend on the N+1 day. The trained model uses the trading data of the first four trading days to predict the closing price of the fifth trading day.

III. Experiment

A. Raw Data

The six characteristics of trading date, opening price, closing price, lowest price, highest price and daily volume of the stock are selected as the input of the model. Data is crawled from the public information of Oriental fortune website by using crawler, and stored in the stock database through data cleaning operations such as removing null value and invalid value.

B. Data Nomalization

Because the feature vector of the model in addition to time has five different dimensions, and its property is different, have different dimension and orders of magnitude, usually stock trading volume of the order of magnitude of orders of magnitude greater than the stock price, if not through data standardization directly using the original indexes were analyzed, and the order of magnitude larger trading volume index in the model will be prominent, and the order of magnitude smaller price index will be weakened. Therefore, in order to ensure the reliability of the results, five eigenvectors other than time need to be standardized. In this paper, the maximum and minimum method is selected for the standardization of input data, and the formula is as follows:

Standardized input data = (original data - minimum)/(maximum - minimum)

In order to ensure the consistency of business meaning of input and output data, the inverse operation of standardization processing is needed. These two parts of the operation are solidified in the program, automatic completion.

C. Traning Detail

We trained model by mini-batch gradient descent, with learning rate of 0.001. We used minibatch size of 64, and normalization was applied for each vector of a sequence by using the mean and standard deviation of each stock computed from the training set. We used a PC server as training platform (CPU:E5-2620, 64G memory, GPU Titan),we chose CentOS 7,Tensorflow as deep learning platform. The duration of one epoch is about 2000 second.

D. Results

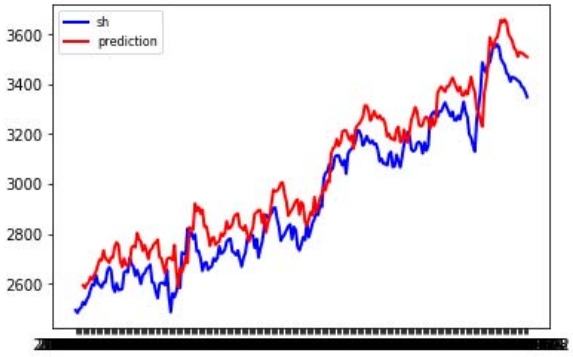

Fig. 2. Shanghai stock index over 1 year

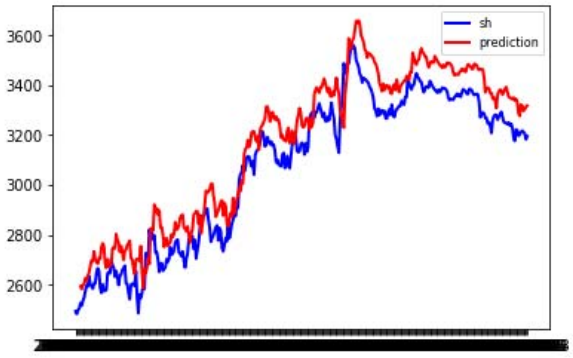

Fig. 3. Shanghai stock index over 18 months

This article proposed a combination of attention mechanism model for the four groups when the length of the time span of different test data of the output and the real value close, but also for different stock index, there was no significant difference between model prediction results, basic can complete forecast movements in stock market, in view of the Shanghai composite index, shenzhen index, the Shanghai and shenzhen 300 index of representative significance for China's stock market as a whole, this article proposed combining attention mechanism when the length of the

memory model in different time Windows has universality to the stock market trends.

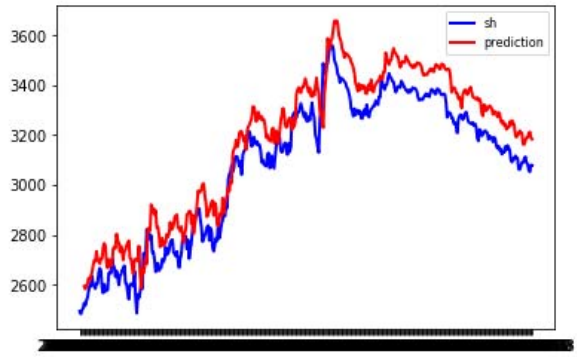

Fig. 4. Shanghai stock index over 2 years

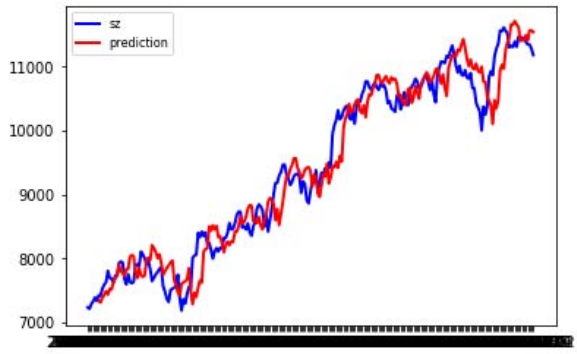

Fig. 5. Shenzhen stock index over 1 year

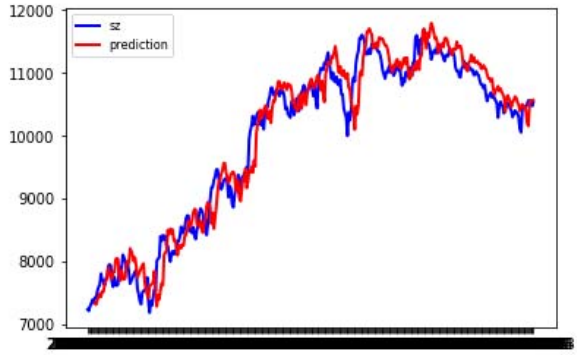

Fig. 6. Shenzhen stock index over 18 months

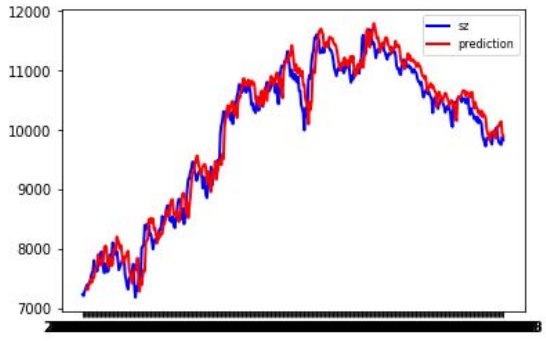

Fig. 7. Shenzhen stock index over 2 years

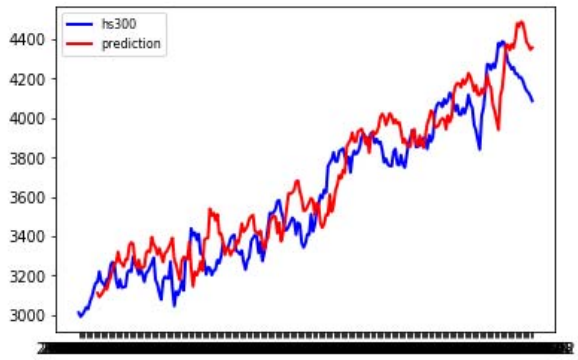

Fig. 8. HS300 index over 1 year

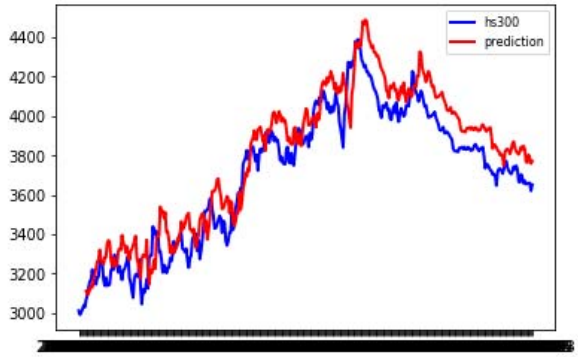

Fig. 9. HS300 index over 18 months

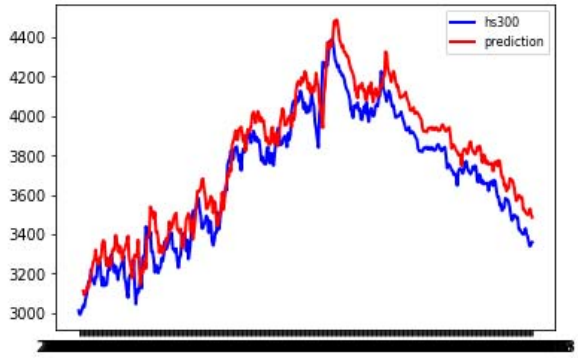

Fig. 10. HS300 index over 2 years

The output of the model has hysteresis in the scene with a small time window, and the peak and trough of the predicted value have a certain time delay compared with the real value. As can be seen from the figure, for the two groups of data with a small time span, the predicted value of blue has a certain time lag compared with the real value of red. The predicted time delay will have a certain impact on the validity of the model, but with the increase of the time span of the input data, the time delay will gradually decrease. In the last group of data with the largest time span, the time synchronization between the predicted value and the real value can be basically guaranteed.

It can also be seen from Fig. 2.-Fig. 10. that in addition to the time window of input data, the size of the data input will also have an impact on the model effect. This is also in line

with the advantages of deep neural network over shallow neural network. The number of hidden layers of deep neural network is large, the number of model parameters is large, the amount of input data used for training is larger, the number of iterations of the model is larger, and the model is more conducive to finding global minimum value.

IV. Conclusion

In this paper, for stock trend prediction method based on attention mechanism LSTM deep learning model, and the representative of China's stock market index (the Shanghai index, shenzhen index and the csi 300 index) was carried out on the feasibility analysis, and verify the effectiveness of the proposed algorithm, the innovation of this article point can be summarized as the following two aspects:

(1) Attention mechanism combined with the depth of the neural network, this paper puts forward a kind of stock trend prediction model of Attention - LSTM, the length of the memory model and the opening price, highest price, closing price of the stock as the input variables of the model, the hidden layer and the Attention after calculation, forecasts the next trading day's closing price of stock deep learning model.

(2) In Attention-LSTM algorithm training, small batch gradient descent algorithm was used to optimize the model in the way of small step iteration adopted by small batch gradient descent algorithm, so that the model could be approximated at a faster speed with a smaller error.

References

[1] B. Wang, F. L. Zhang, "Comparison of artificial neural network and time series model for forecasting stock prices", *Journal of Wuhan University of Technology (Information and Management Engineering)*, vol. 27, no. 6, pp. 69-73, 2005.

[2] J. Lin, Y. H. Guo, "Short term prediction of stock prices based on neural networks", *Journal of Southwest Jiaotong University*, vol. 33, no. 3, pp. 299-304, 1998.

[3] J. H. Sun, "Long time series clustering method and its application in stock price", *Diss. WuHan University*, 2011.

[4] D. Christian, S. Cincotti, "Clustering of financial time series with application to index and enhanced index tracking portfolio", *Physica A Statistical Mechanics & Its Applications*, vol. 355, no. 1, pp. 145-151, 2005.

[5] H. C. Yin, C. Y. Zhao, "Research on stock forecasting based on neural network", *Natural Science Journal of Harbin Normal University*, vol. 23, no. 3, pp. 47-49, 2007.

[6] F. Gianni, S. Surcis, "A cloud computing based real time financial system", *ACM Symposium on Applied Computing ACM*, pp. 1219-1220, 2009.

[7] Colah, Understanding LSTM Networks, 2015, [online] Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[8] F. A. Gers, J. Schmidhuber, F. Cummins, "Learning to forget: continual prediction with LSTM", *Neural Computation*, vol. 2, no. 10, pp. 2451-71, 1999.

[9] S. Martin, R. Schlüter, H. Ney, "LSTM neural networks for language modeling", *Interspeech*, pp. 601-608, 2012.

[10] G. Alex, "Supervised sequence labelling with recurrent neural networks", *Studies in Computational Intelligence*, vol. 385, 2012.