Stock Transaction Prediction Modeling and Analysis Based on LSTM

Siyuan Liu[1,2] , Guangzhong Liao[1,2] ,Yifan Ding[1,2]
[1]Hubei Key Laboratory of Intelligent Information Processing and Real-time Industrial System
[2]School of Computer Science and Technology,
Wuhan University of Science and Technology
Wuhan, China, 430065
*Corresponding email: 15671628256@163.com

***Abstract*—Stock price volatility is a highly complex nonlinear dynamic system. The stock's trading volume affects the stock's self correlation, self correlation and inertial effect, and the adjustment of the stock is not to advance with a homogeneous time process, which has its own independent time to promote the process. LSTM (Term Memory Long-Short) is a kind of time recurrent neural network, which is suitable for processing and predicting the important events of interval and long delay in time series. Based on temporal characteristics of stock and LSTM neural network algorithm, this paper uses the LSTM recurrent neural networks to filter, extract feature value and analyze the stock data, and set up the the prediction model of the corresponding stock transaction.**

Keywords—machine learning; neural network; stock transaction prediction; LSTM

I. Introduction

The human cognition of the logic of the stock market volatility is a very challenging and difficult problem in the world. So far, there is no theory and method that is convincing and can stand the test of time.

In October 14, 2013, the Royal Swedish Academy of Sciences awarded the American economist Eugene Fama, Lars Peter Hansen and Robert Schiller the Nobel prize in economics pointed out: There is little way to accurately predict the direction of the stock market in the next few days or weeks, but may be able to predict the price of more than three years. From the characteristics and perspective of the research paradigm, the analysis method of stock investment mainly has three kinds: basic analysis, technical analysis, evolutionary analysis. The technical analysis methods commonly used are based on statistics, such as trend analysis, morphological analysis, line analysis and technical index analysis, cycle analysis etc [1].These methods are simple and easy to grasp, but the stock market is a very complex nonlinear dynamic system, which is influenced easily by many factors, such as the political form, the financial policy, the situation of the company and the important news [2].

Some researchers have studied the development trend of stock price from the traditional analysis method. Sun Jihong [3] used the long time clustering method to predict the future price of stocks, which filled the shortage of domestic stock prediction research.Dose and Cincotti [4] studied the stochastic optimization method of cluster analysis to select stocks for investment. However, the analysis and prediction accuracy of stock prices is not high. So, traditional technical analysis methods are difficult to obtain satisfactory results.

In recent years, the artificial neural network is a hot research field in the computer field [2]. And because of its nonlinear mapping ability, good self-learning and adaptive performance, analysis and prediction of the neural network has been widely used in the field of stock price, financial income, and the exchange rate risk [5]. Fenu G [6] use artificial intelligence neural network to study the best time of stock investment. Artificial intelligence neural network(ANN) has strong nonlinear approximation ability to nonlinear relation.

As the stock price, volume and other data contains a large number of information affecting the stock price changes, ANN can learn the historical data of the stock, so as to find the law of stock prices. But the financial data is affected by many factors in reality, and the time series formed by it is more random and random, and it usually has multi level and multi scale characteristics. Therefore, The prediction model of single neural network has limitations and has a certain impact on the prediction accuracy of stock prices.

LSTM neural network is a special kind of recurrent network [7]. LSTM can keep the error, for the reverse pass along the time and layer. LSTM keep the error at a more constant level, so that recursive network can be a lot of time to learn, so as to open the establishment of a long distance causal link. In this paper, we use characteristics of LSTM neural network algorithm on time series to predict short-term changes of the corresponding stock transaction.

II. LSTM Neural Network

In a traditional recurrent neural network, during the gradient back-propagation phase, the gradient signal can end up being multiplied a large number of times (as many as the number of timesteps) by the weight matrix associated with the connections between the neurons of the recurrent hidden layer. This means that, the magnitude of weights in the transition matrix can have a strong impact on the learning process.

If the weights in this matrix are small (or, more formally, if the leading eigenvalue of the weight matrix is smaller than

978-1-5386-3758-6/18/$31.00 ©2018 IEEE

1.0), it can lead to a situation called vanishing gradients where gradient signal gets so small that learning either becomes very slow or stops working altogether. It can also make more difficult the task of learning long-term dependencies in the data. Conversely, if the weights in this matrix are large (or, again, more formally, if the leading eigenvalue of the weight matrix is larger than 1.0), it can lead to a situation where the gradient signal is so large that it can cause learning to diverge. This is often referred to as exploding gradients.

These issues are the main motivation behind the LSTM model which introduces a new structure called a memory cell (see "Fig.I" below). A memory cell is composed of four main elements: an input gate, a neuron with a self-recurrent connection (a connection to itself), a forget gate and an output gate [8]. The self-recurrent connection has a weight of 1.0 and ensures that, barring any outside interference, the state of a memory cell can remain constant from one timestep to another. The gates serve to modulate the interactions between the memory cell itself and its environment. The input gate can allow incoming signal to alter the state of the memory cell or block it. On the other hand, the output gate can allow the state of the memory cell to have an effect on other neurons or prevent it. Finally, the forget gate can modulate the memory cell's self-recurrent connection, allowing the cell to remember or forget its previous state, as needed.

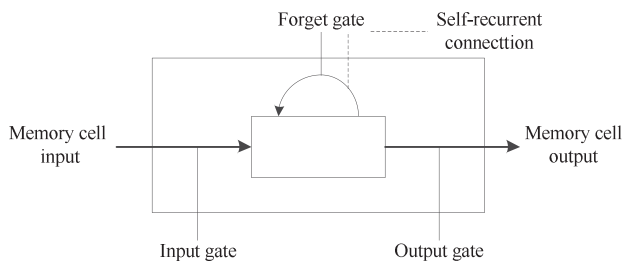

Fig. 1. Illustration of an LSTM memory cell.

A standard neural network unit only consists of the input activation and the output activation which are related, when a activation function is used by [9]

$$b_i = \tanh(a_i) \tag{1}$$

The equations below describe how a layer of memory cells is updated at every time step t. In these equations :

- x_t is the input to the memory cell layer at time t.
- W_i, W_f, W_c, W_o, U_i, U_f, U_c, U_o and V_o are weight matrices.
- b_i, b_f, b_c and b_o are bias vectors.

First, we compute the values for i_t, the input gate, and $\tilde{C}_t$ the candidate value for the states of the memory cells at time t :

$$\begin{aligned} i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \\ \tilde{C}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \end{aligned} \tag{2}$$

the

Second, we compute the value for f_t, the activation of the memory cells' forget gates at time t :

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{3}$$

Given the value of the input gate activation i_t, the forget gate activation f_t and the candidate state value $\tilde{C}_t$,we can compute C_t the memory cells' new state at time t :

$$C_t = i_t * \tilde{C}_t + f_t * C_{t-1} \tag{4}$$

With the new state of the memory cells, we can compute the value of their output gates and, subsequently, their outputs [10]:

$$\begin{aligned} o_t &= \sigma(W_o x_t + U_o h_{t-1} + V_o C_t + b_o) \\ h_t &= o_t * \tanh(C_t) \end{aligned} \tag{5}$$

III. Modeling

A. Data Source

The experiment data comes from the stock data interface of JoinQuant [1] platform. In this platform , we can get stock historical transaction data and obtain the k line data by setting related parameters. The basic information of these data is as shown in "TABLE I".

TABLE I. THE DESCRIPTION OF THESE STOCK DATA'S INDEX

Index	Description
open	price at the beginning of time
close	price at the end of time
low	minimum price
high	highest price
volume	number of stock transaction
money	price of stock transaction
limit_up	limit up
limit_down	limit down

In this experiment, we select data is the closing data of CSI 603899 Index from 2014-05-18 to 2017-01-29.

For example, some CSI 603899 Index data is as shown in "TABLE II" .

B. Data Pre-Processing

Firstly, calculate the stock's MA , EMA index by the closing price data.

MA : Moving Average. C_i is the closing data on a certain day.

[1] https://www.joinquant.com/

TABLE II. THE STOCK DATA SAMPLE

Date	Open	Close	High	Low	Volume	Limit_up	Limit_down
2016-08-22	17.56	18.15	18.19	17.49	1378206.0	19.36	16.18
2016-08-23	18.02	17.92	18.11	17.80	1656454.0	19.85	16.31
2016-08-24	18.13	18.11	18.25	17.87	1415056.0	19.82	16.16
2016-08-25	17.91	17.85	18.08	17.73	1033421.0	19.79	16.25
2016-08-26	18.01	18.20	18.42	17.89	1734215.0	19.80	16.17

$$MA = \sum_{i=1}^{n} X_i / n \tag{6}$$

EMA : Exponential Moving Average. X is a variable, N is a certain day, Y' is EMA of last cycle.

$$EMA(X,N) = Y = (2*X + (N-1)*Y')/(N+1) \tag{7}$$

Secondly, preprocess the correlation index of stock data according to the following methods.

$$\begin{aligned} oc &= (close - open)/open \\ oh &= (high - open)/open \\ ol &= (low - open)/open \\ ch &= (high - close)/close \\ cl &= (low - close)/close \\ lh &= (high - low)/low \end{aligned} \tag{8}$$

C. LSTM Model

- The feature of the training samples are MA, EMA, oc, oh, ol, ch, cl and lh of the closing data of CSI 300 Index from 2014-05-18 to 2016-12-25.

$$\begin{aligned} &faclist = ['MA','EMA','oc','oh','ol','ch','cl','lh'] \\ &del\_list = ['volume','open',close','high','low'] \end{aligned} \tag{9}$$

- The classification of the training samples are the ups and downs of each day from 2014-05-18 to 2016-12-25. Ups is true, downs is false.
- The test samples are MA, EMA, oc, oh, ol, ch, cl and lh of the closing data of CSI 300 Index from 2016-12-26 to 2017-01-29.

Then, use the above conditions and keras2 package to structure LSTM Networks mode.

[2] https://keras.io/

IV. EXPERIMENT RESULTS AND DISCUSSIONS

A. Experiment Results

The stacked LSTM obtains the instance training algorithm by extracting the hidden state information of the lower LSTM by stacking multi-layer LSTM, which has achieved good results. But it's limited by tanh, the LSTM model needs to consume more computing resources after more than five layers. So, this paper uses a not more than three layer of the stacked LSTM model.

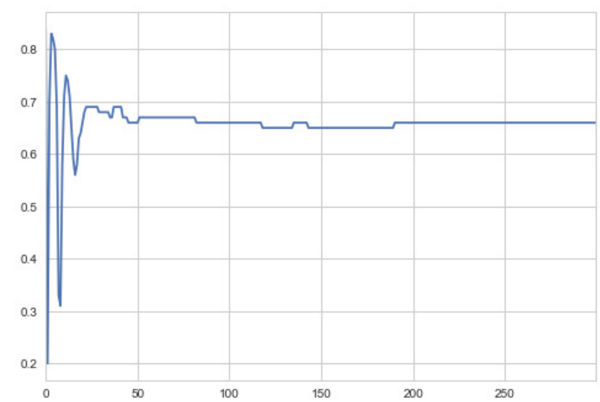

Fig. 2. The predicted results of the single layer LSTM Model

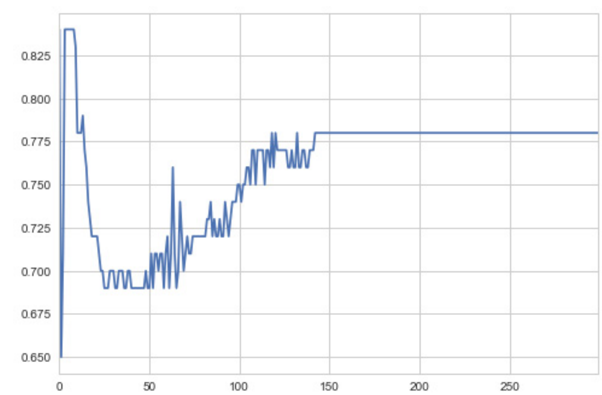

Fig. 3. The predicted results of the three layer LSTM Model

B. Experiment Discussions

The experimental result shows that the sample accuracy rate of the single layer LSTM Model is 0.66. However, the sample accuracy rate of the three layer LSTM Model is up to 0.78. Thus, the more stack layers of LSTM model, the higher the accuracy of prediction results. Of course, if the stack layer is higher, the computational resources consumed by the model will increase.

Therefore, in order to obtain more accurate prediction results, it is necessary to combine computing resources within the acceptable time range to choose the LSTM model with the proper number of stacked layers. At the same time, It is necessary for LSTM network to be combined with existing clustering techniques to gain large speed ups in training and testing times at a small loss in performance.

V. CONCLUSION

In this paper, we uses the LSTM recurrent neural networks to extract feature value and analyze the stock data. The experimental results show our model can play a better forecasting effect, even though the accuracy is not very high, only about 72% for the short period of data. But we believe the model still has a lot of improvements to improve its accuracy.

In the future, we will extract more feature values to training our model and improve the model. In the same time, we will also compare the pros and cons of different neural networks model in stock forecasting to find better model to improve the prediction accuracy.

REFERENCES

[1] Wang Bo, Zhang Fengling. "Comparison of Artificial Neural Network and Time Series Model for Forecasting Stock Prices." Journal of Wuhan University of Technology (Information and Management Engineering) 27.6(2005):69-73.

[2] Lin Jie, Guo Yaohuang. "Short Term Prediction of Stock Prices Based on Neural Networks." Journal of Southwest Jiaotong University 33.3(1998):299-304.

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

[4] Dose, Christian, and S. Cincotti. "Clustering of financial time series with application to index and enhanced index tracking portfolio ☆." Physica A Statistical Mechanics & Its Applications 355.1(2005):145-151.

[5] Yin Hongcai, Zhao Chunyan. "Research on Stock Forecasting Based on Neural Network." Natural science journal of Harbin Normal University 23.3(2007):47-49.

[6] Fenu, Gianni, and S. Surcis. "A Cloud Computing Based Real Time Financial System. " ACM Symposium on Applied Computing ACM, 2009:1219-1220.

[7] Colah. Understanding LSTM Networks. http://colah.github.io/posts/ 2015- 08-Understanding- LSTMs/. 27.8(2015).

[8] Gers F., A., J. Schmidhuber, and F. Cummins. "Learning to Forget: Continual Prediction with LSTM." Neural Computation 2.10(1999): 2451-71.

[9] Sundermeyer, Martin, R. Schlüter, and H. Ney. "LSTM Neural Networks for Language Modeling." Interspeech 2012:601-608.

[10] Graves, Alex. "Supervised Sequence Labelling with Recurrent Neural Networks." Studies in Computational Intelligence 385(2012).