Stock Market Prediction based on Deep Long Short Term Memory Neural Network

Xiongwen Pang[1], Yanqiang Zhou[1], Pan Wang[2], Weiwei Lin[3] and Victor Chang[4]
[1]*School of Computer, South China Normal University, Guangzhou, China*
[2]*China Merchants Bank Branch in Wuhan, Wuhan, China*
[3]*School of Computer Science and Engineering, South China University of Technology, Guangzhou, China*
[4]*International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Suzhou, China*

Keywords: Embedded Layer, Long-Short Term Memory Neural Network (Lstm Neural Network), Stock Market Prediction, Stock Vector.

Abstract: To study the influence of market characteristics on stock prices, traditional neural network algorithm may also fail to predict the stock market precisely, since the initial weight of the random selection problem can be easily prone to incorrect predictions. Based on the idea of word vector in deep learning, we demonstrate the concept of stock vector. The input is no longer a single index or single stock index, but multi-stock high-dimensional historical data. We propose the deep long-short term memory neural network (LSMN) with embedded layer to predict the stock market. In this model, we use the embedded layer to vectorize the data, in a bid to forecast the stock via long-short term memory neural network. The experimental results show that the deep long short term memory neural network with embedded layer is state-of-the-art in developing countries. Specifically, the accuracy of this model is 57.2% for the Shanghai A-shares composite index. Furthermore, this is 52.4% for individual stocks.

1 INTRODUCTION

Since the establishment of the socialist market economy, the financial market plays an irreplaceable role in China. In recent years, with the development of national economy and improvement of financial awareness, the financial market has attracted the attention of domestic and foreign scholars and investors. They constantly put forward various theories which have been applied to the practice, hoping to predict market trends. However, as the market is influenced by the national policy, economic situation, as well as psychological, human and other factors, the financial market forecasts fail to achieve the desired results on a frequent basis (Hsu et al., 2016).

Based on various researches, neural network has been used in many areas such as pattern recognition, financial securities and signal processing. Additionally, it has been widely regarded for its advantage on regression and classification in the stock market forecast. However, traditional neural network algorithms may also fail to predict the stock market precisely, since the initial weight of the random selection problem is easy to fall into the local optimal resulting in incorrect predictions.

Most of the current researches are based on foreign stock markets, such as the Standard & Poor (S&P) index, the Nasdaq index and so on. The stock markets in developing countries such as the Shanghai stock market make fewer forecasts, and thus the effect of the prediction is not significant. Based on studies of the deep learning (Mikolov et al., 2013)(Chen et al., 2014), this paper introduces the concept of stock vector by referring to the "word vector of natural language processing", and performs the simulation experiment on the Shanghai A-shares market through the improved long-short term neural network techniques. The neural network can make an effective forecast for the financial market in China. It can read and understand the market trend and eventually provide useful analysis for investors.

Pang, X., Zhou, Y., Wang, P., Lin, W. and Chang, V.
Stock Market Prediction based on Deep Long Short Term Memory Neural Network.
DOI: 10.5220/0006749901020108
In *Proceedings of the 3rd International Conference on Complexity, Future Information Systems and Risk (COMPLEXIS 2018)*, pages 102-108
ISBN: 978-989-758-297-4

2 RELATED WORKS

The methods of shallow machine learning commonly used in the current stock forecasting model are neural networks, support vector machines, or models that combine them with other algorithms. Hsu demonstrated that the method of machine learning could be used to predict financial market more accurately than economic methods (Hsu et al., 2016). He also proved that the forecasting effect of the financial market was affected by the maturity of the market, the input variable, the base forecasting time and the forecasting method.

Niaki used the neural network with 27 types of economic variables to predict the S&P index. The result showed that the neural network model, compared with trading strategy, had increased the profit of transactions (Niaki et al., 2013), which relied on the method of initialize the weight. Zhong used 60 kinds of economic variables and three kinds of dimensionality reduction, including PCA, Kernel PCA (KPCA), Fast Robust PCA (FRPCA), combined with neural network to predict the S&P index. The result showed that PCA was the best, achieving 57% of accuracy (Zhong et al., 2017).

In these traditional neural networks, there are some issues such as falling into local optimal value and a single hidden layer. These problems eventually lead to the accuracy of forecast is not high in stock market. Hinton published the "Reducing the Dimensionality of NN" (Hinton et al., 2006) that set off an upsurge of deep learning. Subsequently, a variety of deep learning models have been developed, and these models have been widely used after.

Zhu used 14 indicators as input such as the opening, higher, lower, closing price, related with technical analysis indicators ROC and RSI, so as to learn the historical data of the S&P index through the DBN. Since then, he forecasted the stock price. The profit was higher than the ordinary trading strategy (Zhu et al., 2014). Chang proposed business analytics of stock market performance, by which investors can trace or review the market performance of their chosen stocks. He selected the Heston Model and its associated API to compute predicted stock index movement and offer a good extent of accuracy (Chang, 2014). Batres Estrada selected as the DBN-MLP model to predict the income of the S & P index. The model had a 53% of accuracy, exceeding the regularized logistic regression and some MLP baselines (Batres Estrada, 2015). Shen applied DBN model to predict the weekly exchange rate and compared with FFNN. It also indicated that the predictive effect of DBN could reach 63% better than 41% of FFNN (Shen et al., 2015).

Yu applied the deep neural network and LSTM to forecast the trading data of the Amazon stock. It was found that the effect of the deep neural network was better than LSTM, and the prediction accuracy was 54% (Yu, 2016). Gao used LSTM to forecast six kinds of different industries in the US stock market, using 359 stock features as input, with an average accuracy rate of 54.83% (Gao, 2016). Sean used the recurrent neural network-LSTM, which was optimized by Bayesian, to predict the price of the coin, reaching 52% of accuracy (Sean et al., 2016).

The inputs of these models are the basic index based on the composite index or the economic variable of the technical index, so the dimension of input data is not too high. In the literatures above, the comparison baseline is based on the traditional machine learning algorithm, in order to prove the effectiveness of those models. Most of the selected objects are more mature stock markets such as the S&P index and the Nasdaq index, etc. Instead, the minority of them are immature stock markets such as the Shanhai stock market, which prediction thereby is not distinct. In the deep learning, the LSTM is used in natural language processing and other serial data because of its unique memory function, but there are not many predictions for stock time series. This paper uses the deep LSTM to obtain useful information from the stock time series and try to predict the immature stock market.

3 STOCK FORECASTING MODEL BASED ON LONG SHORT TERM MEMORY NEURAL NETWORK

3.1 Stock Vector

In natural language processing, there is a large number of unlabeled texts. When computers process text, it is necessary to convert the text into a format that the computer can understand. Therefore, researchers proposed the "word vector", which uses a series of numbers to express the word. The simplest method is the one-hot vector.

In order to fully extract information from the corpus, researchers introduced new machine learning methods, which mainly include RBM, neural network and the correlation between word and context (Chen et al., 2014).

However, RNN based on language model in learning process is prone to exploding and vanishing gradients. For this reason, the extended model of RNN was proposed (Mikolov et al., 2013), which added a feature layer to the original structure. The structure is shown in Figure 1. For a specific language model, given a word sequence $(w_1, w_2, \cdots w_T)$, hidden layer sequence $(h_1, h_2, \cdots h_T)$ and the output layer sequence $(y_1, y_2, \cdots y_T)$ in RNN are described in Eq. (1).

$$\begin{aligned} \mathrm{h}_t &= \tanh(Ww_t + Uh_{t-1} + Ff(t) + b_h) \\ y_t &= Ah_t + b_y \end{aligned} \tag{1}$$

In Figure 1, w_t is the one-hot vector of the t word, and $\theta = \{b_h, b_y, W, U, A\}$ is the set of learning parameters. The embedded vector of the word is the matrix W. When the feature layer *f(t)* is added, the loss information is not attenuated to 0, thus the vanishing gradient problem is avoided.

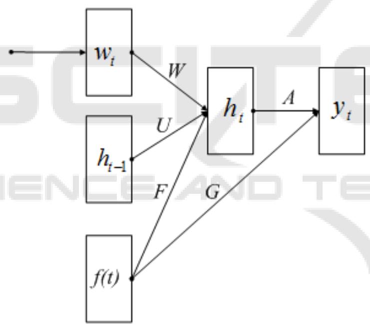

Figure 1: RNN structure of feature layer.

In the past stock market forecasting, the input is often a single stock or a single index. Thus the input dimension will not be too large. However, when the input is historical data of multiple stocks, the input dimension will increase to thousands or even millions. At this time, if we use these original information to predict the stock market directly, it may result in greater errors due to the impact of information redundancy and irrelevant information.

Therefore, we introduce the concept of vectorization for stock market, named stock vector. The stock vector refers to the idea of word vector. First of all, the dimension of the stock vector is reduced, and then it is expressed in a low dimension space. Finally, the stock market is forecasted by stock vector. Based on the method of word vector, a kind of model of reducing dimension and predicting stock price is proposed: the long short term memory neural network with embedded layer (ELSTM).

3.2 The Deep Long Short Term Memory Neural Network with Embedded Layer based on Stock Vector (ELSTM)

Since the RNN is prone to exploding and vanishing gradients during the training process, the gradient can't be passed down in long sequence. As a result, RNN cannot capture the effect of long distance. Therefore, LSTM, an improved RNN, is used to predict the stock sequence. The LSTM has four layers, which are shown in Figure 2.

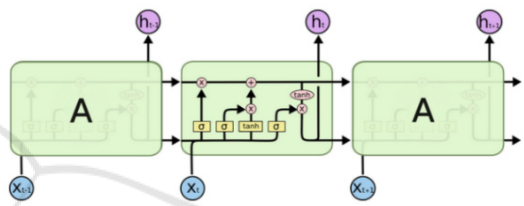

Figure 2: LSTM module diagram.

3.2.1 Construction of ELSTM

The opening time of each stock in the A-shares market is not the same, which makes the stock gap too big. In order to align the data, the missing data is padded with 0, so there will be some sparse sub-matrices. At the same time, in order to reduce the dimension of data and filter the irrelevant information, we add the embedded layer in front of the LSTM layer to achieve precisely stock prediction. The structure of the entire network is shown in Figure 3.

It can be seen from the Figure 3 that the main part of the framework is embedded layer and LSTM layer. The working principle of each layer is as follows:

(1) Embedded layer. The embedded layer is initialized to random matrix. We convert high-dimensional data into low-dimensional data by matrix transformation. In Figure 3, the A-shared dataset is converted to a stock vector that is the input of neural network. In the whole model training process, the matrix in the embedded layer is trained as the parameters of the network. The training matrix is the prerequisite and basis of further feature extraction. Throughout the Error Back Propagation (EBP) process, the training goal is minimum error.

(2) LSTM layer. After the reduced dimension processing in the embedded layer, we got the stock vector. The LSTM circularly read the stock vector, and further extracts the feature information. Then it predicts the stock value and compares it with the standard value. Finally, it updates the parameters via EBP in the LSTM model.

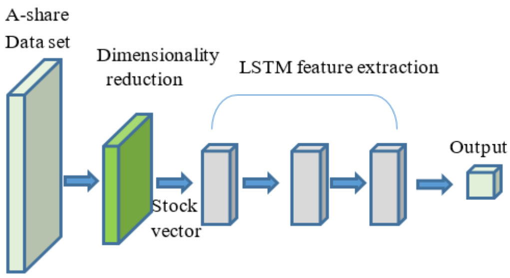

Figure 3: Network structure diagram.

(3) Output layer. For the regression prediction, the output layer has a single neuron. While the direction of ups and downs forecast, the output layer has two neurons.

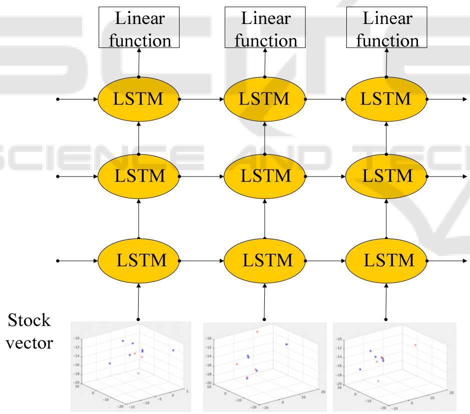

Figure 4: Timing diagram of LSTM prediction.

Figure 4 shows the timing structure of the LSTM module. The stock vector is transferred to the LSTM of the three layers, and the output sequence is calculated iteratively. Finally, the output is processed by the activation function to get the result. The input of each LSTM is not only affected by a lower layer, but also receives the effect of the same layer output, taking full account of the characteristic of time series. When regression prediction has been performed, the last layer uses a linear function to predict the output. While the ups and downs are predicted, the classification function is *logistic*.

4 EXPERIMENTAL RESULTS AND ANALYSIS OF THE MODEL

4.1 Experimental Preparation

The hardware environment of this paper is Centos 64bit operating system, 62G RAM, Intel(R) Xeon(R) CPU, and the software environment is Theano framework based on Python language.

4.1.1 Model Input

We use crawler technology to grab information about the stock. This article includes two experiments:

(1) We filter the historical data of single stock on the Shanghai A-share market, including the opening price, the highest price, the lowest price, the closing price and the volume. With them, it is possible to predict the price and trend of the Shanghai A-share Composite Index.

(2) Based on the original five indicators, we add some additional indicators, such as daily amplitude, 5-day amplitude, 10-day amplitude and amplitude of fluctuation of the volumes, amounting to a total of 9, so as to predict the price and trend of a single stock.

The data of 2006/1/1-2016/10/19 for 10 years were selected as the object of study, 70% of which were training set, 10% were validation set and 20% were test set. First, the dimensions of the input dataset are aligned, and then the data is normalized on [0, 1]. The specific description is as follows:

$$x^* = \frac{x - \min}{\max - \min} \tag{2}$$

The max is the maximum value of the sample, and the min is the minimum of the sample, so that the x is mapped to [0, 1].

4.1.2 Model Output

In order to analyze the effect of the model more comprehensively, this paper evaluates the performance of the model for short-term stock price forecast from two aspects, namely, Mean Square Error (MSE) and Data Accuracy (DA).

4.2 Comparison and Analysis of Models

The models of ELSTM, DBN, DBN-MLP and Multilayer Perceptrons (MLP) are respectively experimented on two datasets. The results are shown in Figure 5-6.

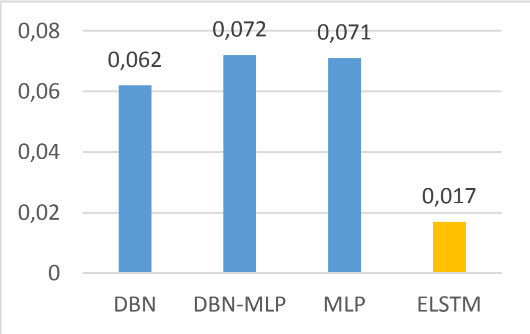

Figure 5: Comparative MSE of Shanghai A-share Composite Index.

According to Figure 5 and 6, it can be seen that the MSE and accuracy of the LSTM is the best. Specifically, the accuracy is about 10% higher than other models and approximately 7% higher than the stochastic prediction. In additions, the MSE reaches the minimum. The MSE of the four contrast models is as the same overall.

The accuracy of the Shanghai A-share Composite Index by LSTM is 57.2%, the same with literature (Zhong et al., 2017) on the S&P 500. From the literature (Hsu et al., 2016), we can see that the market forecast accuracy is affected by market maturity. The more mature the market is, the higher the accuracy rate is. In fact, the Chinese stock market is less mature than the US. Fortunately, we took the Chinese market and index into account, and the effectiveness of the method used in this paper is still demonstrated better than other competing methods to some extent.

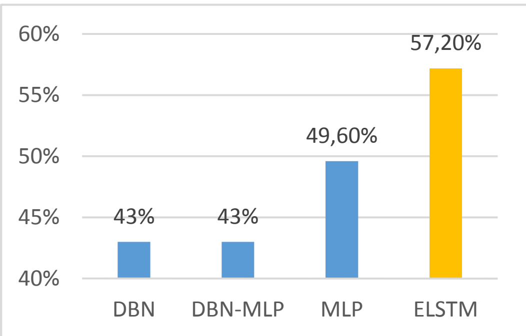

Figure 6: Comparative accuracy of Shanghai A-share Composite Index.

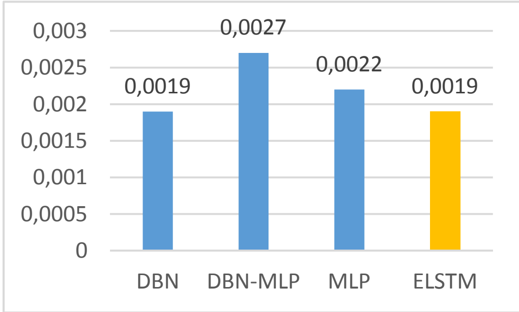

Figure 7: Comparative MSE of Sinopec.

For the prediction of Sinopec, the results of several forecasting methods are similar. From Figure 7, we can see that the best is ELSTM and DBN, then MLP, DBN-MLP in MSE. It is easy to draw a conclusion from Figure 8 that such models have minor differences in accuracy. As for all models, the accuracy is always higher than 50%, which indicate that the input information of stock in this paper does improve the accuracy of Sinopec.

Taking into account the results of predicted Shanghai A-share Composite Index and Sinopec, the results are more stable with the ELSTM method, since it will not fluctuate greatly due to the randomization of the initial weights. Therefore, we choose ELSTM model to conduct empirical analysis of the stock. The ELSTM on the Shanghai A-share and the Sinopec forecast trend as shown in Figure 9 and 10 respectively.

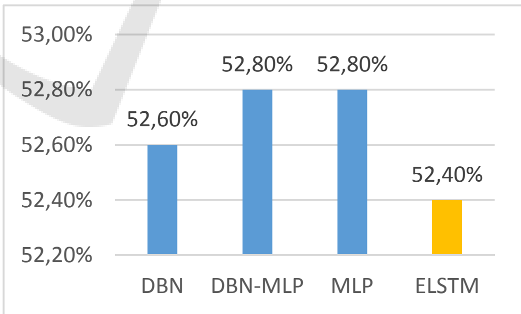

Figure 8: Comparative accuracy of Sinopec.

It can be seen from Figure 9 that the waveform of the predictive curve and the actual curve for Shanghai A-share Composite Index are mostly consistent. Between 100 and 300 days, the peaks of the two curves are the same, while the small fluctuation in other states is kept in a similar state. From Figure 10, it is depicted that the actual waveform of Sinopec and the predicted are also similar to a degree. Nevertheless, the price gap between 200 and 500 is comparatively larger. Thus, the effect is not as good as the Shanghai A-share Composite Index.

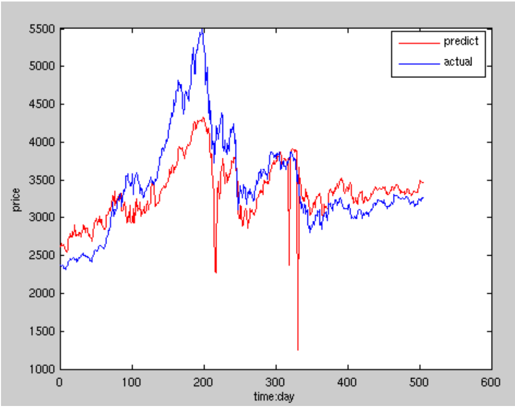

Figure 9: Comparison of predictive value and true value of the Shanghai A-share Composite Index.

In summary, the individual stock is sensitive to external factors such as the suspension of the stock market and some error of the training data, which affected the test results. The Shanghai A-share Composite Index is determined by A-share data rather than a certain stock. Hence, it is predictive effect is relatively satisfying.

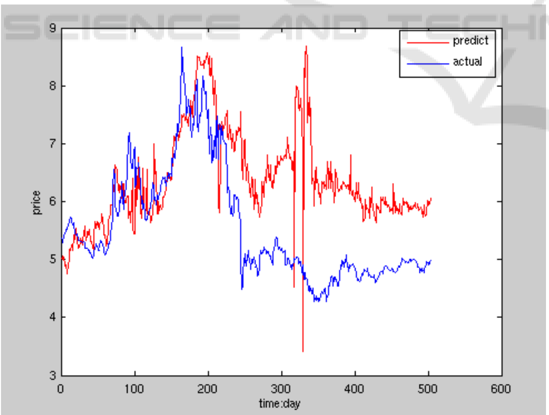

Figure 10: Comparison of predictive value and true value of the Sinopec.

5 CONCLUSION

In this paper, the LSTM neural network with embedded layer is proposed on the basis of LSTM neural network. Firstly, we verify the performance of the models in Shanghai A-share Composite Index and Sinopec. Then we use a better ELSTM model for other selected stocks. The average accuracy of the three stocks is 53.2% and the A-share Composite Index is 57%, which are higher than the stochastic forecast. In summary, the methods mentioned in this paper have better predictive performance for the Shanghai A-share Composite Index.

Although the models can improve the predicted effect of Shanghai A-share Composite Index to a certain extent, there are still some deficiencies in the input of historical data. The text information in the stock market such as news is not fully utilized. So in the next step, we will consider adding text information factor to the model to further improve the performance.

ACKNOWLEDGEMENTS

This work is partially supported by the National Natural Science Foundation of China (Grant Nos. 61772205 and 61402183), Guangdong Provincial Scientific and Technological Projects (Grant Nos. 2017A010101008, 2017B010126002, 2017A010101014, 2017B090901061, 2016A010101007 and 2016B090918021), Guangzhou Science and Technology Projects (Grant Nos. 201607010048, 201604016013 and 201604010040), and Special funds for the development of industry and information of Guangdong Province (big data demonstrated applications) in 2017.

REFERENCES

Hsu M.W., Lessmann S., Sung M.C., Ma T., Johnson J. E.V, 2016. Bridging the divide in financial market forecasting: machine learners vs. financial economists. *Expert Systems with Applications, 61:215-234.*

Niaki S.T.A., Hoseinzade S., 2013. Forecasting S&P 500 index using artificial neural networks and design of experiments. *Journal of Industrial Engineering International, 9(1):1-9.*

Zhong X., Enke D., 2017. Forecasting daily stock market return using dimensionality reduction. *Expert Systems with Applications, 67: 126-139.*

Hinton G.E., Salakhutdinov R.R., 2006. Reducing the dimensionality of data with neural networks. *Science, 313(5786):504.*

Chengzhang ZHU, Jianping YIN, Qian LI, 2014. A Stock Decision Support System based on DBNs. *Journal of Computational Information Systems, 10: 2 883–893.*

Victor Chang, 2014. The Business Intelligence as a Service in the Cloud. *Future Generation Computer Systems, 37(2014)512 -534.*
Batres-Estrada B., 2015. Deep learning for multivariate financial time series. KTH Royal Institute of Technology.
Shen F., Chao J., Zhao J., 2015. Forecasting exchange rate using deep belief networks and conjugate gradient method. *Neurocomputing, 167(C):243-253.*
Yu Fu-fu, 2016. Based on the neural network of stock forecast [D]. Harbin, Harbin Institute of Technology.
Gao Q., 2016. Stock market forecasting using recurrent neural network. University of Missouri--Columbia.
da Silva I.N., Hernane Spatti D., Andrade Flauzino R., Liboni L.H.B., dos Reis Alves S.F., 2017. Forecast of Stock Market Trends Using Recurrent Networks. In: Artificial Neural Networks. Springer, Cham.
Chen E., Qiu S., Xu C., Tian F., Liu T., 2014. Word embedding: Continuous space representation for natural language. *Data Acquisition and Processin, 29 (1): 19-29.*
Mikolov T., Zweig G., 2013. Context dependent recurrent neural network language model. *Spoken Language Technology Workshop. IEEE, 234-239.*