Prediction on the Highest Price of the Stock Based on PSO-LSTM Neural Network

YUSHAN ZHANG
School of Statistics and Mathematics
Guangdong University Of Finance & Economics
Guangzhou, China
676044395@qq.com

SITONG YANG
School of Statistics and Mathematics
Guangdong University Of Finance & Economics
Guangzhou, China

***Abstract*—With the development of artificial intelligence technology, the stock market has entered a new research stage. Machine learning algorithm is widely used in the prediction of financial time series. The highest price of a stock is an important factor to measure the price of a stock. For this reason, this paper uses three representative historical data of American stocks as the research object to analyze the short-term price trend. In addition to the original short and long term memory networks, in order to improve the accuracy of prediction, we put forward a prediction model combining particle swarm optimization algorithm and long - and short-term memory neural network. Through empirical research, the PSO optimization algorithm proposed for LSTM model can quickly find the optimal network weight of the neural network, reduce the loss function, achieve the effect of rapid fitting, and has a more accurately predicted results.**

Keywords: LSTM Neural Network; PSO Algorithm; The Highest Price of the Stock; Prediction

I. INTRODUCTION

In the recent years, machine learning has been being applied to various fields including finance, transportation and medicine. As an important part of financial products, the stock not only reflects peoples' living standard but also indicates the market economy so the prediction on the stock trend is becoming more and more popular with researchers [1]. Though the rich historical stock information has provided the mass investors and analyzers with data resources, it is apparently unreliable and inefficient to analyze and judge merely depending on the analyzers' personal experience and intuition. Therefore, people are in need of an intelligent, scientific and efficient research method to help with and guide the stock trade [2]. In the past, the scholars mainly applied various traditional statistical methods to time series model, such as GARCH (Auto-Regressive Conditional Heteroscedasticity Model), ARMA (Auto-Regressive Moving Average Model) [3] and Multiple Linear Regression Model [4]. However, such models usually can fit the linear relationship between stock price and influencing factors only. And consequently, faced with prediction on the stock price with high complexity, it is hard for the prediction accuracy to be convincingly persuasive. In the past few years, the machine learning algorithms, including decision tree [5], Support Vector Machine [6], improved clustering algorithm [7] and improved neural network [8], have been applied to stock prediction [9] and have achieved relatively good prediction results. Though neural network has achieved relatively good results in stock prediction, the deeper neural network may be more suitable for such nonlinear complex system [10]. RNN (Recurrent Neural Network) is frequently applied to analyze the prediction series data [11], but RNN has the problem of gradient disappearance in handling with long time series problems. The time memory unit owned by LSTM neural network can learn the long and short term information of time series so it is applicable to predict the time series problems with interval and delay events. Hence, LSTM neural network is applied to the prediction on the highest price of the stock and LSTM is optimized by PSO in this paper. The previous scholars usually conducted the parameter optimization of network model through optimizing iteration number and learning rate by PSO algorithm while, in this paper, PSO is introduced to optimize the updating of internal weights of neural network, w and b, by searching for the global optimum instead of optimizing the traditional external parameters of neural network so that the neural network model can rapidly fit the target data to construct the optimal network structure. Furthermore, LSTM neural network optimized by PSO and traditional LSTM neural network are compared from two aspects, prediction result and loss function change.

II. PRINCIPLES OF LONG AND SHORT TERM MEMORY (LSTM) NEURAL NETWORK

In LSTM model, the hidden-layer neurons in RNN are replaced by LSTM neurons to offer the ability of long-term memory. The structure of LSTM neurons is shown as Fig.1 whose forward computing can be expressed as:

$$i_t=\sigma(W_{xi}x_t+W_{hi}h_{t-1}+W_{ci}c_{t-1}+b_i)$$

$$f_t=\sigma(W_{xf}x_t+W_{hf}h_{t-1}+W_{cf}c_{t-1}+b_f)$$

$$c_t=f_tc_{t-1}+i_t tanh(W_{xc}x_t+W_{hc}h_{t-1}+b_c)$$

$$o_t=\sigma(W_{xo}x_t+W_{ho}h_{t-1}+W_{co}c_t+b_o)$$

$$h_t=o_t tanh(c_t)$$

Where i, f, c, o refer to input gate, forget gate, internal status, output gate, respectively. x_t refers to the current input. h_{t-1} and h_t refer to the previous output and the current output, respectively. c_{t-1} and c_t refer to the last cell status and the current cell status. W and b refer to the corresponding weight and bias term, respectively. σ and tanh refer to sigmoid and hyperbolic tangent activation function. The structure of gate enable LSTM to allow the information to selectively affect the status of each moment in RNN. The value is between 0-1 and it

978-1-7281-3584-7/19/$31.00 ©2019 IEEE

is feasible to describe how much information is being input currently through this structure. The forget gate determines which part in the previous status can be forgotten while the input gate determines which part can enter the current status. The output gate determines the current output based on the newest status, previous output and current input.

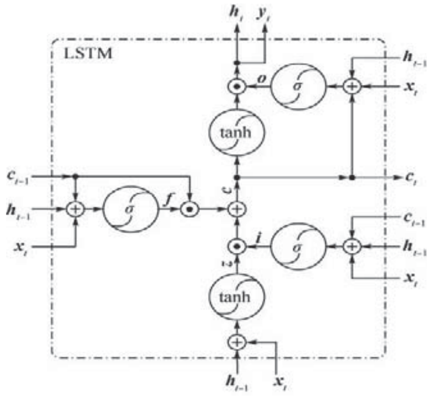


Fig.1 Basic Structure of LSTM

III. PARTICLE SWARM OPTIMIZATION (PSO)

PSO is a swarm intelligence algorithm designed by simulating the foraging behavior of birds. If there exists the optimal solution to the problem in the area (the mapping refers to the food that birds forage), a type of particles is used to simulate the bird individuals mentioned above with two attributes of speed and position. The current position of the particle can be regarded as a candidate solution to the optimal solution. Each particle independently searches for the optimal solution in the space and records it as the current individual extremum, P_{best}. Through continuously updating the speed and position and iterating, the global optimal solution, G_{best}, that satisfies the conditions is finally acquired.

The particle update its speed and position according to the following formula:

$$V_{i,t+1}=\omega C_1 random(0,1)(P_{besti}-X_{it})+C_2 random(0,1)(G_{besti}-X_{it})$$

$$X_{i,t+1}=X_{it}+V_{i,t+1}$$

Where, ω is named as inertia factor whose value is negative. The local and global optimizing ability can be adjusted through adjusting the value of ω. C_1 and C_2 refer to the individual learning factor and social learning factor. $random(0,1)$ refers to a random number in the interval [0,1]. The flow chart of PSO is shown as Fig.2.

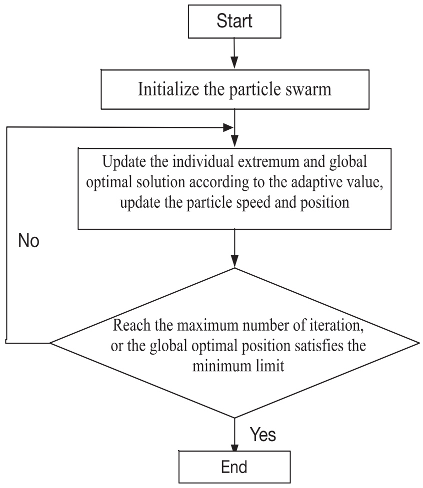


Fig.2 Flow Chart of PSO

IV. PSO OPTIMIZATION OF LSTM NEURAL NETWORK ALGORITHMIC FLOW

As for LSTM-PSO model, PSO is introduced based on traditional LSTM to optimize the internal parameters of neural network. The traditional PSO usually conduct the parameter optimization of network model through optimizing the iteration number and learning rate. However, PSO introduced in this paper aims at enabling the neural network model to rapidly fit the target data by rapidly searching for the global optimum to optimize the weight updating of internal weights of neural network, w and b, instead of traditionally optimizing the external parameter of neural network. The specific steps are listed as below:

Step 1: initialize the particle swarm: we confirm the population scale, learning factors, iteration number, constrained intervals of position and speed.

Step 2: A population particle, $X_{i0}(h_1, w, b)$ randomly generates. h_1 refers to LSTM neurons. W refers to the weight matrix of LSTM, $w_{m\times n}$. b refers to the bias of LSTM, $b_{n\times 1}$.

Step 3: Start to train to search for the optimum: we input all the particles into the model for training and evaluate the quality of particles through calculating the adaptive degree function of particle swarm, based on which the position and speed of each particle are adjusted. The smaller the numerical value of the adaptive degree function is, the higher the adaptive degree is. The optimization function is:

$$Fit=\frac{1}{m}\sum_{p=1}^{m}\sqrt{(P_{px}-T_{px})^2}$$

Where, m refers to the number of samples for training. P_{px} and v refer to the predicted input result and actual input result of samples in the particle swarm.

Step 4: calculate the adaptive degree value corresponding to the position of each particle, X_i; determine the individual extremum and group extremum according to the initial particle adaptive degree value; select the best position of each particle as its best position in history. The optimum is calculated according to the following formula:

$$V_{i,t+1}=\omega C_1 random(0,1)(P_{besti}-X_{it})+C_2 random(0,1)(G_{bestt}-X_{it})$$

Where, P_{besti} refers to the optimal position of No. I particle in the population of generation t in d-dimensional space, that is, the individual optimal solution. G_{bestt} refers to the optimal position in the population of generation t, that is, the global optimal solution. w refers to the weight. C1 and C2 refer to the learning factors.

Step 5: update the speed and position of particles themselves through individual extremum and global extremum; calculate the adaptive degree value of new particles and update the individual extremum and group extremum of particles according to the adaptive degree value of new population particles. After satisfying the maximum iteration number of PSO, the parameters of LSTM will also update to the optimum with inputting X. input the prediction data into the LSTM model trained by optimal particles to output the prediction results of the stock.

V. Basic Experiment Contents

In this paper, Python is used for programming and TensorFlow is used as the deep learning framework for training. The programming and debugging of prediction on the highest price of the stock based on PSO-LSTM neural network are realized while image fitting and error evaluation are carried out. In view of rigor, rationality and accuracy of the experiment, the listed individual stocks with relatively huge market values on the US Stock Market, AMZN, AAPL and MSFT, are finally chosen in this paper to reduce the relevant influence of banker control or policy.

A. Data Source and Pre-processing

With AMZN, AAPL and MSFT as the research targets, the daily K data during October 28, 2010 and October 26, 2018 is downloaded by Yahoo Finance and there are csv files of seven attributes, Date, Open, Close, High, Low, Ajd Close and Volume, in the data.

Due to the differences in dimensions of parameters in the data, there exist larger differences in order of magnitude among the attributes. Therefore, normalization is conducted to control the data dimension within (0,1) and the normalization function is:

$$X_{new}=\frac{X_{old}-min(X)}{max(X)-min(X)}$$

Thus, the results of model training should be treated by anti-normalization and the function is:

$$y_{new}=y_{old}(max(x_i)-min(x_i))+min(x_i)$$

Where y_{old} refers to the output of neural network.

B. Evaluation Indexes

In order to evaluate the performance of the prediction model, we select three indexes to measure the prediction accuracy of the model in this paper, including mean absolute error (MAE) that can reflect the actual situations of errors in prediction value, mean average percentage error (MAPE) that takes both the error between prediction value and actual value and the proportion between error and actual value into account, and root-mean-square error (RMSE) that is used for measuring the deviation between observation value and actual value. The calculating formulas are shown as below:

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i|$$

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right|\times 100\%$$

$$RMSE=\frac{1}{n}\sqrt{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$$

C. Construction of LSTM & PSO-LSTM Comparative Experiment

LSTM and PSO-LSTM models are established, respectively, on the data of the three stocks to verify the proposed methods. The prediction method is the highest price of the stock at 20 consecutive time points in the first [I, I+19] and the output is the prediction value of the next moment I+20. The first 1980 data is mainly selected for training. The latter 56 items are used as the test set and the latter 34 items are used as the prediction data set. When the model is used for analysis, the selection of model parameters is also an important aspect. In LSTM model, the activation function of LSTM module is tanh and the activation function of completely-connected artificial neural network that confirms to receive the LSTM output is linear. The epoch of model training is 50 and batch_size is 200. Some codes in the model construction are shown as Fig.3.

Because the data sets have been standardized before, the acquired prediction values need to be treated by anti-standardization. A cycle is set for the random seeds of each stock and the mean value of 20 cycles is output as the result. The specific flow is shown as Fig.4.

```
model = Sequential()
#model.add(LSTM(100, input_shape=(
#   train_X.shape[1], train_X.shape[2])))
model.add(LSTM(20, activation='sigmoid',
               input_shape=(train_X.shape[1],
                            train_X.shape[2]),
                            return_sequences=True))
model.add(LSTM(20))
model.add(Dense(1,activation='linear'))
model.compile(loss=rmse_koss, optimizer='adam')

# fit network
history =model.fit(train_X, train_y, epochs=50,
                   batch_size=200,verbose=1,
                   validation_data=(test_X, test_y))
```

Fig.3 Partial Network Framework Code Built

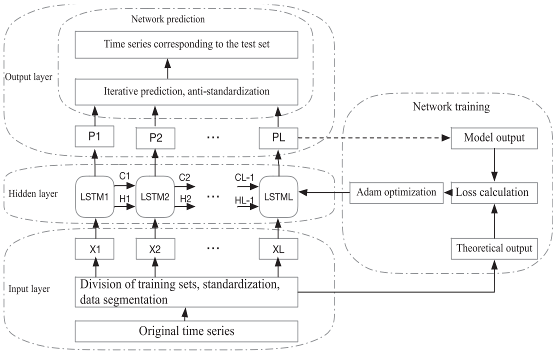


Fig.4 Prediction Framework of the Highest Price of the Stock Based on LSTM

VI. EXPERIMENT RESULTS

The series of the highest prices of the three stocks in the original data are shown as Fig.5. It can be seen from the figure that, in the recent years, the stock trends of Apple and Microsoft are relatively steady while the highest stock price of Amazon fluctuates considerably. There are 2014 data points in total. In order to sufficiently reflect the fluctuation range of stock price in the test set, the first 1980 items are used as the training data set while the latter 34 items are used as the prediction data set.

Fig.5 Trend of the Highest Stock Price in History

A. AMZN Stock

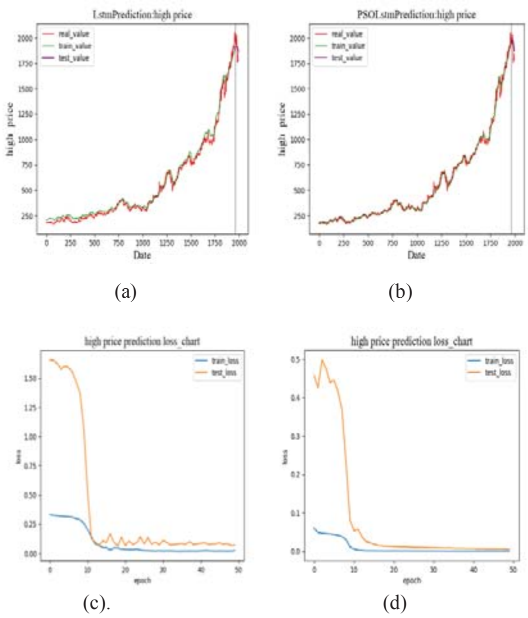


Fig.6 Prediction Results and Loss Function of AMZN Stock

Notes: (a) stock prediction results based on LSTM. (b) stock prediction results based on PSO-LSTM. (c) changes in LSTM loss function. (d) changes in PSO-LSTM loss function.

As shown by Fig.6, as for AMZN stock, the three curves in Fig. (a) and (b) refer to the actual value, training set value and test set value of the stock while the two curves in Fig. (c) and (d) refer to the loss functions of training set and test set. If the proposed PSO is applied to LSTM model, the optimal network weight of neural network can be rapidly found and the loss function can also be rapidly lowered. After optimizing the neural network by PSO-LSTM, the number of iteration reaches about 15, the model has already been able to fit the data set of the stock. Therefore, the results are predicted and the results are shown as Tab.1. In view of each evaluation index, the prediction results of PSO-LSTM are more accurate than those of LSTM where RMSE on the test set is 7.56.

Tab.1 Results of Prediction Indexes of AMZN Stock

Index		Training	Test
LSTM	RMSE	8.94	35.76
	MAPE	0.02	3.27
	MAE	0.02	0.06
LSTM-PSO	RMSE	4.78	7.56
	MAPE	0.01	1.61
	MAE	0.01	0.03

B. AAPL Stock

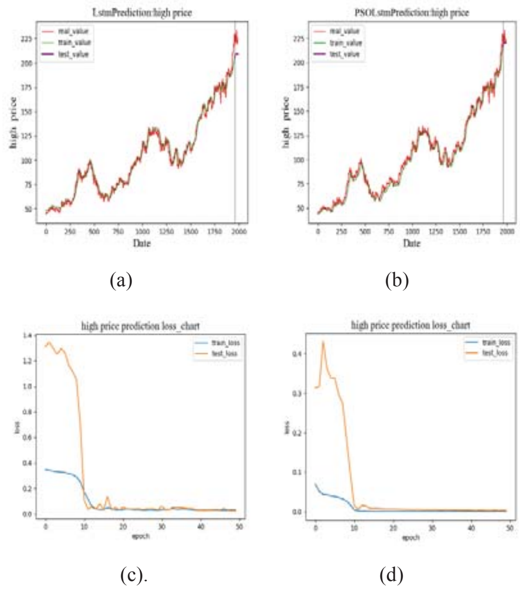


Fig.7 Prediction Results and Loss Functions of AAPL Stock

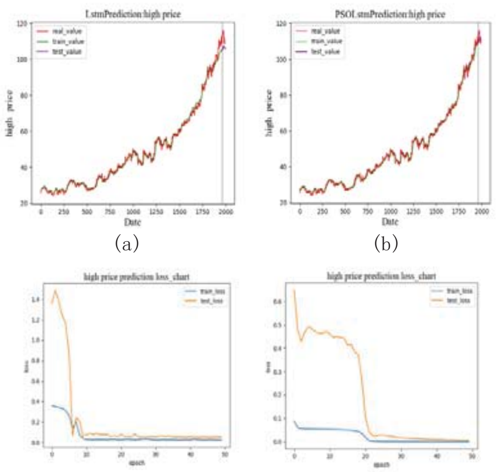
Notes: (a) stock prediction results based on LSTM. (b) stock prediction results based on PSO-LSTM. (c) changes in LSTM loss function. (d) changes in PSO-LSTM loss function.

The prediction of the highest price of AAPL stock is shown as Fig.7. If the proposed PSO is applied to LSTM model, the optimal network weight of neural network can be rapidly found and the loss function can also be rapidly lowered. It can be inferred from the iteration number of loss function that the gradient descent of PSO-LSTM loss function is relatively steady and there is no major fluctuation in the loss function of test set. When the number of iteration reaches about 17, the model is able to fit the data set of the stock. Therefore, the result prediction is conducted and the prediction results are shown as Tab.2. PSO-LSTM is always more optimal than LSTM in each index where RMSE on the test set is 3.71

Tab.2 Results of Prediction Indexes of AAPL Stock

Index		**Training**	**Test**
LSTM	**RMSE**	4.55	10.28
	MAPE	0.06	0.27
	MAE	0.02	0.05
LSTM-PSO	**RMSE**	3.43	3.71
	MAPE	0.01	0.09
	MAE	0.01	0.02

C. MSFT Stock



(a) (b)

(c) (d)

Fig.8 Prediction Results and Loss Functions of MSFT Stock

Notes: (a) stock prediction results based on LSTM. (b) stock prediction results based on PSO-LSTM. (c) changes in LSTM loss function. (d) changes in PSO-LSTM loss function.

The prediction of the highest price of AAPL stock is shown as Fig.8. The prediction results of the highest price of LSTM neural network and PSO-LSTM neural network are shown as Fig.(a) and (b) while the changes in loss function are shown as Fig.(c) and (d). It can be inferred from the iteration number of loss function that the gradient descent of PSO-LSTM loss function is relatively steady and there is no major fluctuation in the loss function of test set. When the number of iteration reaches about 20, the model is able to fit the data set of the stock, indicating that PSO-LSTM model can rapidly fit the stock price data more more accurate prediction. The results of each evaluation index are shown as Tab.3 where RMSE on the test set is 2.54.

Tab.3 Results of Prediction Indexes of MSFT Stock

Index		**Training**	**Test**
LSTM	**RMSE**	2.05	7.04
	MAPE	0.06	0.19
	MAE	0.02	0.07
LSTM-PSO	**RMSE**	1.74	2.54
	MAPE	0.01	0.07
	MAE	0.02	0.02

VII. CONCLUSION

The highest price of the stock is an important factor to measure the price of a stock, so predicting the highest price of the stock is the standard for purchasing the stock. AMZN, APLL and MSFT are chosen as the research objects and LSTM neural network and LSTM neural network improved by PSO are used to predict the highest price of the stock, respectively, in this paper. As shown by the experiment, machine learning can better find and use the inter-data interaction. Meanwhile, LSTM neural network model optimized by PSO improves the accuracy of prediction and is advantageous in rapid search of optimal network weight in the neural network.

ACKNOLEDGEMENT

This work is supported by National Natural Science Foundation of China (61876207) ()()(61876207)Identify applicable sponsor/s here. (sponsors)

REFERENCES

[1] Z. Su, M. Lu, and D. X. Li, "Financial empirical application of deep learning:Dynamics, contributions and prospects," J. Financ. Res. Beijing, pp. 111-126, March 2017.

[2] F. G. Deng and H. L. Wang, "Application of LSTM neural network in prediction on stock price trends, based on research on individual stock data in US and HK stock market," Financ. Econ. Beijing, pp. 96-98, July 2018.

[3] X. Y. Li, X. L. Chen, and K. H. Zeng, "Research on prediction of stock return rate of four banks based on the three models," Math. Econ. Changsha, vol. 35, pp. 21-27, October 2018.

[4] C. X. Wang, "Research on the impact of overseas stock ownership on mainland stocks," Master's thesis, Nanjing University, Nanjing, June 2018.

[5] R. W. Lin and Y. Yang, "Trend prediction on Chinese stock market, based on multidimensional interactive verification method," J. Shanghai Univ. Int. Bus. Econ. Shanghai, vol. 25, pp. 40-49, November 2018.

[6] K. J. Kim, "Financial time series forecasting using support vector machines," Neurocomputing Netherlands, vol. 55, pp. 307-319, September 2003.

[7] J. Zakaria, A. Mueen, and E. Keogh, "Clustering time series using unsupervised-shapelets," in IEEE Int. Conf. Data Min. Brussels, Belgium, December 2012, pp. 785-794.

[8] J. V. Hansen and R. D. Nelson, "Neural networks and traditional time series methods: a synergistic combination in state economic forecasts," IEEE Trans. Neural Netw. New Jersey, vol. 8, pp. 863-873, January 1997.

[9] B. Krollner, B. J. Vanstone, and G. R. Finnie, "Financial time series forecasting with machine learning techniques: A survey," in ESANN, 2010, 18th Eur. Symp. Artif. Neural Netw., Bruges, Belgium, April 2010, pp. 25-30.

[10] K. Ji, "Prediction on Stock Price Trends Based on BP Neural Network," Master's thesis, Zhengzhou University, Zhengzhou, June 2018.

[11] J. Li and Y. F. Lin, "Time series data prediction based on multi-time scale RNN," Comput. Appl. Software Shanghai, vol. 35, pp. 33-37, 62, October 2018.