Cloud Publications
International Journal of Advanced Computer Science and Information Technology
2020, Volume 6, Issue 1, pp. 101-114
ISSN 2320–0235, Crossref: 10.23953/cloud.ijacsit.461

Research Article

Comparison of Back Propagation, Long Short-Term Memory (LSTM), Attention-Based LSTM Neural Networks Application in Futures Market of China using R Programming

Wang Shuangao[1], Liu Yi[1], Rajchandar Padmanaban[2,3], Mohamed Shamsudeen[4] and Subalakshmi R[5]

[1]Business School, China University of Political Science and Law, Xueyuan Road Campus: 25 Xitucheng Lu, Haidian District, Beijing, China
[2]NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Campus de Campolide, Lisbon, Portugal
[3]Forest Research Centre, School of Agriculture, University of Lisbon, Tapada da Ajuda,, 1349-017 Lisbon, Portugal
[4,5] University V.O.C. College of Engineering, Anna University Thoothukudi Campus, 7th Street West, Bryant Nagar Main Road, Thoothukudi, Tamil Nadu, India.

Correspondence should be addressed to Wang Shuangao, *wangshuangao@outlook.com*

Publication Date: 14 May 2020

DOI: https://doi.org/10.23953/cloud.ijacsit.461

Abstract Artificial neural network is widely used in the financial time series, but Long short-term memory (LSTM) neural network is rarely used in the futures market in China. In this paper, the LSTM neural network is studied by using futures data. The daily trading data of four groups of futures such as silver, copper, lithium and coking coal from December 2014 to December 2018 are used as the training object to make short-term prediction of the closing price. By comparing the Back Propagation (BP) neural network, general multi-layer LSTM neural network, and using the attention mechanism optimization LSTM contrast test, the result of the experiment shows that the futures price trend forecast time sequence, attention mechanism to promote significant effect of time sequence, and LSTM combined effect, by adjusting the parameters setting, using the improved LSTM neural network for time series prediction accuracy is higher, better generalization ability.
Keywords *LSTM Neural Network; Futures Forecasting; Attention Mechanism; Financial Engineering*

1. Introduction

Forecasting stock and futures prices is an interesting and difficult task for many analysts and academics, and because of the inherent complexity and dynamics of such price movements, improving their accuracy is particularly difficult (White, 2002). In the past few decades, economists have tried to predict stock and futures markets using linear measurement tools (AR, MA, ARIMA) and nonlinear algorithms (ARCH, GARCH, neural network) (Padmanaban and Karuppasamy, 2018).

In recent years, a small number of scholars began to study the application of data mining methods to the stock market. Compared with the traditional measurement method, neural network has better prediction performance and learning performance in financial time series, and has more comparative advantages (Gencay, 1996). For the domestic stock market, a large number of scholars use various optimized BP neural networks to predict the stock market price and trend. With the further development of computer technology, some domestic scholars began to use Long Short Term Memory (LSTM) neural network for stock market prediction (Rajchandar, 2012). However, there are relatively few researches on the price prediction of futures market using neural network, and few researches on domestic futures market using LSTM neural network. The scale of futures market is huge, and there is still a large space for neural network research (Hinton, 2012).

In 1988, H. White used the neural network to study the daily stock return of IBM, which was the world's first prediction study on time series using machine learning. However, he failed to achieve the expected effect, because the model got into the problem of local minimum value in the training process, that is, the gradient explosion. Grudnitski used the basic neural network to predict the gold futures price in 1993, mainly to verify the applicability of neural network, and found that the prediction ability of neural network was better than the traditional data model (White, 2002). In 1996, German scholar Gencay used forward neural networks and perceptron to make an empirical analysis of Dow Jones industrial average index, which was mainly based on average data analysis, and the prediction effect was good (Gencay, 1996). In 2004, G. Peter Zhang used the method of combining ANN neural network with ARIMA traditional time series algorithm to accurately prove that the neural network was more accurate than the results obtained by ARIMA in nonlinear data analysis (Grudnitski and Osburn, 1993). In 2004, Shaikh and Iqbal used the neural network method to reveal the difference between implied volatility and temporal volatility to study the price changes of standard & poor's 500 index futures (Hamid and Iqbal, 2010); Hamid and Zakaria used four BP neural networks to predict the price of Iranian crude oil. On the whole, the optimized BP model was more effective than the ARIMA model. The larger the amount of training data, the more accurate the effect. Jan-chung Wang using a general equilibrium model to forecast the Taiwan stock market, in the model to join the market after the stochastic interest rate and market volatility, found that the prediction effect of this method than the general equilibrium pricing model prediction effect is strong, Jan-chung Wang also found that using the EWMA model with GARCH (1, 1) model method of combining the forecasting error of the futures market volatility than other models (Visalatchi and Padmanaban, 2012).

This paper summarizes the typical literatures on the prediction of non LSTM neural network model of futures as follows: by combing the related literatures on the prediction of futures price, it is found that the main prediction methods for futures can be roughly divided into two categories: one is to predict through GARCH and its modified model; the other is to comprehensively use other models, such as principal component analysis method, on the basis of neural network, to overcome the problem. The limitation of neural network improves the accuracy of prediction. For the price prediction using neural network, Yu Wen uses CNN LSTM neural network to analyze the secondary financial market data (Hochreiter and Jurgen Schmidhuber, 1997). Compared with the traditional simple statistical methods and some other neural network methods, such as logistic regression, convolutional neural network (CNN) is a more effective analysis of financial secondary market (Monishiya and Padmanaban, 2012). The market price is predicted in a relatively short period of time, and the prediction for a longer period of time is significantly improved, which is 10% higher than the simple statistical method and 5% higher than other neural networks (Dashti and Hamid, 2011). Using neural network to forecast futures, generally using LSTM combined with other methods. Yishun Liu (2019) found this model combines the VMD (variation pattern decomposition) method and LSTM (long-term short-term memory) network, constructs the prediction model, and forecasts the prediction trend and the inverse neural network model (BPNN) (Qiu and Akagi, 2016).

Leiji forecasts the price of carbon futures based on ARIMA CNN LSTM model. Chenhao Wang (2018) established a model to predict the high and low price of soybean futures through LSTM neural network (Rajchandar and Bhowmik, 2017). Although the volume of China's futures market is huge, but generally speaking, there are few comparative studies on the futures market using neural network in China (Baek and Kim, 2018). Moreover, there are many problems in data selection range, model applicability and so on. In this paper, the LSTM based on the attention mechanism is compared with LSTM and BP neural network, which has some innovations.

2. Introduction to the Model Principle

2.1. Model Principle

LSTM is a special RNN (Recurrent neural network) Recurrent neural network. In 1997, Sepp Hochreiter and Jurgen Schmidhuber proposed LSTM algorithm (Hochreiter and Jurgen Schmidhuber, 1997). The recurrent neural network (RNN) is a time-depth neural network capable of processing sequence data. Since the significant disadvantage of RNN is "insufficient memory length", that is, it is unable to remember too far back or too far forward, the selection of hidden layer weight has a great impact on the learning and training process (Venkatesan and Padmanaban, 2012). If the weight is small, the gradient will disappear. Since the error gradient can be accumulated in the update, if the weight is large, it will become a very large gradient, which will result in a large update of the network weight, which will make the network unstable and eventually lead to "gradient explosion", which will result in unconvergence (Rumelhart and David, 1986). In order to deal with the gradient problem existing in RNN, the LSTM model of "peep hole connection" will be used for reference in this study. The advantage of the LSTM model is that it is better for processing and predicting time series events with long intervals and delays (Fakhruddin and Mahalingam, 2018). The reason for this is that the basic unit of its short and long memory network consists of one or more memory blocks and three adaptive multipliers, namely the input gate, the output gate and the forgetting gate (see Figure 1). Through the three gates to achieve the preservation and control of information. The process of studying and predicting the futures time series is shown in the following Figure 1:

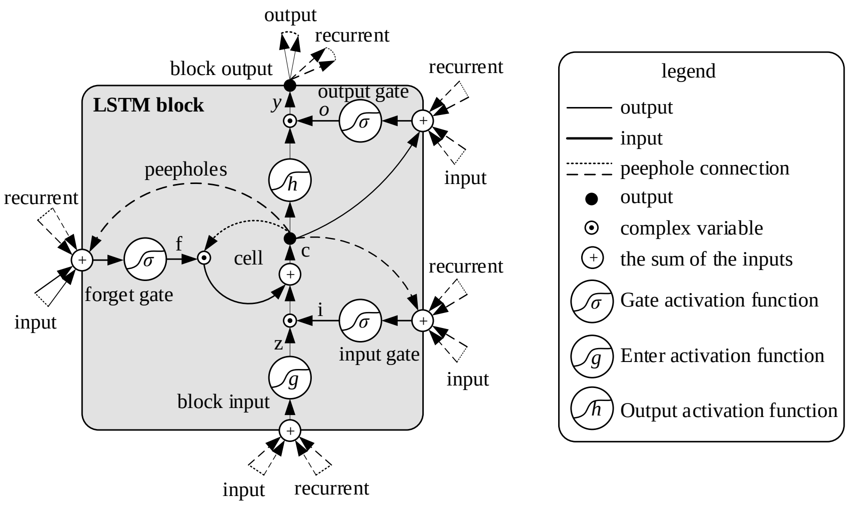


Figure 1: *Schematic diagram of LSTM model, made by the author*

2.2. LSTM Related Formulas

The relevant mathematical formulas in this paper are obtained according to Ren Jun's research and appropriately revised on this basis (Wen and Yuan, 2018). The input data is set as X, and Zc is the current state value of the neuron Cell. Each Cell is updated at time t and LSTM. The condition of the input gate is as follows:

$$Z_l^t = \sum_{i=1}^{l} w_{il} x_i^t + \sum_{hh=1}^{H} w_{hl} y_h^{t-1} + \sum_{c=1}^{C} w_{cl} Z_c^{t-1} + b_i \quad (1)$$

In formula (1), the variable with subscript l is related to the input gate. The first term is the input of the external unit, and the third term is the input from the dashed part of the Cell. The second term with the subscript h is a generic term, because both the elements and the hidden nodes in the LSTM model can be interconnected, so a portion of the external input can also be represented by it, where it represents the bias vector of the input gate b_i. For forgotten doors gate is as follows:

$$Z_\phi^t = \sum_{ii=1}^{l} w_{i\phi} x_i^t + \sum_{h=1}^{H} w_{h\phi} y_h^{t-1} + \sum_{c=1}^{C} w_{c\phi} S_c^{t-1} + b_\phi \quad (2)$$

2.3. BP Neural Network Model

BP neural network is Rumelhart & McClelland two people first proposed in 1986, it is a special kind of multilayer forward neural network, BP algorithm is to solve the multi-layer forward neural network weights optimization and proposed, is characterized by forward signal transmission, error back propagation, is currently the most widely used in neural network learning algorithm of a class of the diagram above is a typical BP neural network. Where, is the input of the input layer, and is the weight between the JTH node of the hidden layer and the node of the input layer. $x_1 \mathrm{L}\ x_i \mathrm{L}\ x_n$ $w_{1j} \mathrm{L}\ w_{ij} \mathrm{L}\ w_{nj}$. (Figure 2)

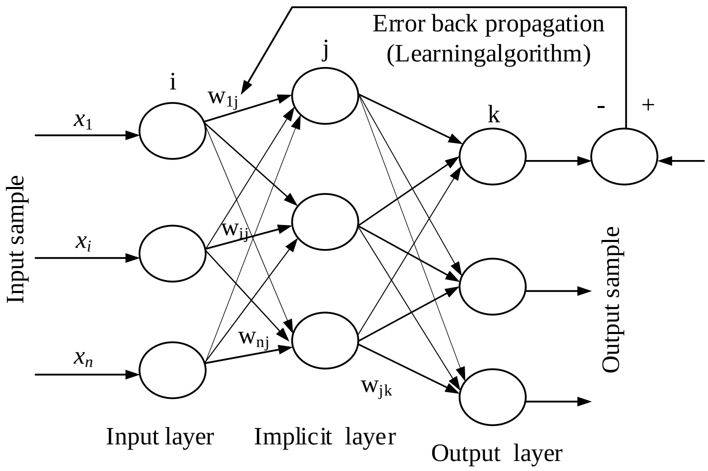


Figure 2: *BP model diagram*

The BP neural network algorithm in this paper refers to the research of Hongwei Ding, and its steps are as follows:

1) Initialization weight and threshold value: the random initialization interval is [-1,1], and a threshold value is set for each cell.
2) propagation forward from the input layer

Input calculation formula of hidden layer:

$$I_j = \sum_{i=1}^{n} W_{ij} O_i + \theta_j \quad (3)$$

Where, is the threshold value of the JTH neuron in the hidden layer.O_j The actual output calculation formula of the hidden layer:

$$O_j = f(I_j) = \frac{1}{1+e^{-I_j}} \quad (4)$$

Where, f is the excitation function. In the same way, the output value of the output layer can be obtained. The derivation steps are omitted.

2.4. LSTM based on Attention

Attention mechanism means to focus on something and get more Attention. The attention mechanism in deep learning is supposed to pay more attention to certain factors when processing data. In a broad sense, attention is an integral part of the network architecture, responsible for managing and quantifying interdependencies:

- Between the input and output elements (general note)
- In the input element (self care)

In this paper, the most important part of LSTM optimization strategy is the addition of the attention mechanism, whose motivation is to consider the different input time series and their different influences on the predicted output. In order to realize this kind of network, softmax function is added to the time sequence dimension of the input end to calculate the weight of attention, as shown in equation (5):

$$weights = \mathrm{soft\,max}(Input) \quad (5)$$

After the weights are obtained, they need to be applied to the input to assign different weights to the timing of the input, so as to make the prediction better.

$$output = weights * (Input) \quad (6)$$

3. Empirical Method

This article uses R programming and Python to process the data and generate training diagrams.

3.1. Data Preprocessing

The raw data of silver, copper and lithium futures in this paper are from Shanghai futures exchange, and the raw data of coking coal futures are from Dalian Commodity Exchange. Since the original futures data are discontinuous or missing due to the stop of trading and so on, this paper standardized the data with excel. The data set is the daily trading data combination of December 16, 2014, December 17, 2018 for a total of 4 years. The sample form after data preprocessing is shown in the following Table 1:

Table 1: *original data styles (silver AG for example)*

Date	Opening price	Highest price	Lowest price	Closing price	Before the settlement price	Settlement price
20141216	3650	3667	3503	3525	3682	3615
20141217	3602	3602	3434	3465	3615	3477
20141218	3434	3520	3418	3496	3477	3472

...	...	...	...	...	...	...
20181217	3550	3550	3500	3508	3524	3511

3.2. Parameter Setting

Learning rate is one of the most important over parameters for deep neural network adjustment. The learning rate is too low, the training is more reliable, and the optimization is too time-consuming, because each step toward the minimum loss function is small. However, if the learning rate is too high, the training may not converge. The amount of weight change can be so large that the optimization passes the minimum and the loss function becomes worse. The number of iterations is the number of times the neural network learns on the training set, and is also the number of times the weight item is updated. In general, the selection of the number of iterations is mainly to make the training loss value of the neural network close to or reach the minimum. When more training times are given, the result of the neural network is no longer greatly improved, and the value nearby is the number of iterations. In general, the three most important parameters, training set, learning rate and iteration times, constitute the input parameters of the prediction function in the neural network.

The structural parameters of BP and LSTM models are designed as follows: the number of independent variables is consistent with the number of neurons; each layer of input and output is available; the division of time points is consistent with the number of model layers; the input is input one by one in accordance with the time sequence. In order to prove the effect of the model, this study tested the closing price, the highest price and the lowest price respectively. Namely, the opening price, the lowest price, the highest price, the trading volume, and the closing price and other five independent variables (parameters).

In this study, BP and LSTM models were used to predict the closing price on the 8th day with 7 transaction days. The partition proportion of the data set is: 85% of the data blocks are used for training, and the remaining 15% are tested in chronological order. The network parameters of BP model are designed as follows. (1) Correction rate: 0.1; (2) number of hidden layers: 1100; (3) learning rate: 0.05; (4) iteration: 1000;

The network parameters of LSTM model are designed as follows: (1) the number of the two hidden layers is: 50,100; (2) learning rate: 0.005; (3) time steps: 30; (4) number of iterations:200. LSTM parameters after adding CNN convolution and Attention: (1) the number of hidden layers is: 128,64; (2) time steps 6;(3) number of iterations: 200.

4. Analysis of Experimental Results

4.1. Model Training Effect

The goal of training (learning) is to set the learning rate as 0.005, and the output error will become smaller and smaller after the iterative training through the neural network until the error is stable in a small interval. The learning rate of the following four futures is set at 0.005, the total number of iterations is 200, and the proportion of data set training and testing is 85%:15%. The image reflects

the convergence of four different training errors in the data set training. The blue line represents the training situation, the yellow line represents the test situation, and the index of the vertical coordinate is the sum of squares of errors. The abscissa represents the number of iterations of the neural network. The loss function of model training is MSE.

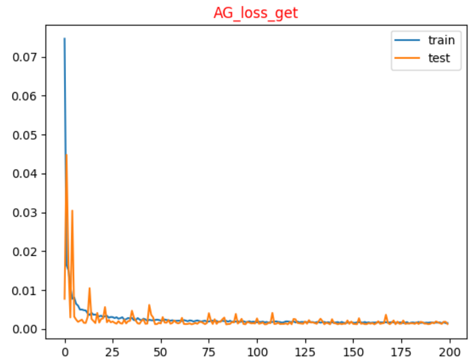


Figure 3: *Silver training renderings*

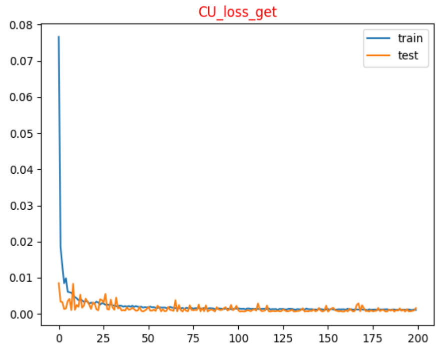


Figure 4: *Copper training renderings*

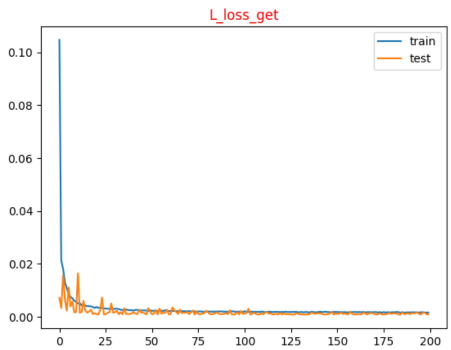


Figure 5: *Lithium training effect diagram*

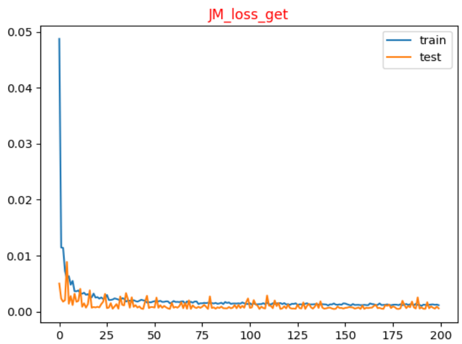


Figure 6: *Coke training effect diagram*

The general LSTM training effect is shown in the Figure 3, 4, 5, 6. It can be seen that silver stays stable after 50 iterations, copper stays stable after 25 iterations, lithium stays stable after 25 iterations, and coke stays stable after 50 iterations.

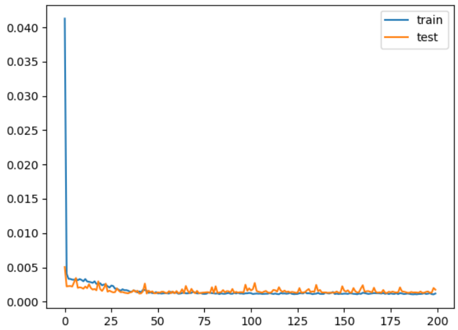


Figure 7: *Training effect of silver closing price*

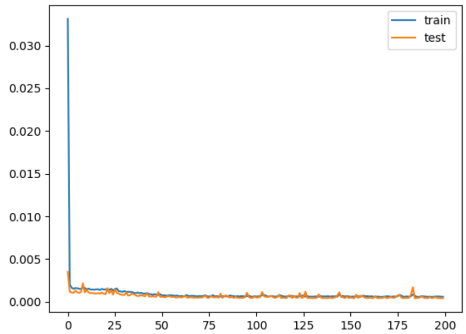


Figure 8: *Training effect of copper closing price*

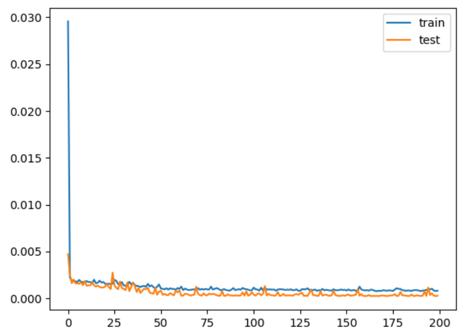


Figure 9: *Training effect of coking coal closing*

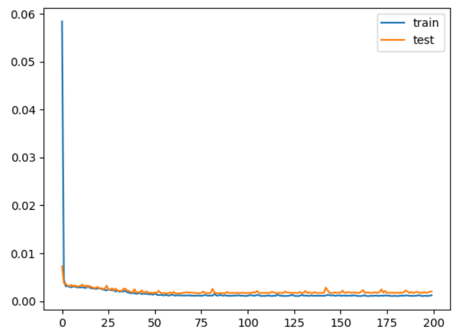


Figure 10: *Training effect of lithium closing price*

The training effect of LSTM model based on the adaptive learning rate based on attention is shown in the Figure 7, 8, 9 and 10. According to the LSTM learning and training based on the above four futures closing prices, the image converges rapidly when the average number of iterations is within 25. The training of the model in the data set can make the error of the neural network converge rapidly to the range of (0,0.1).

Learning from the general training, the number of iterations are within 100 images to achieve rapid convergence, model in the training data set, can make the error of the neural network is fast convergence in (0,0.1) interval, in the number of iterations is lower than 50 times of model prediction error was rapidly decreases, that model has fast convergence characteristics, all futures prediction and stability in the lower interval training error, show that model is stable. In general, the model after further optimization has good applicability and stable prediction performance.

4.2. Comparative Analysis of the Prediction Effect between BP and LSTM

The graph comparison and analysis of the predicted results are shown in the following legend. The x-coordinate in the Figure 11, 12, 13, 14, 15 and 16 represents the number of iterations, and the y-coordinate is the error value and the futures price, respectively. The yellow curve represents the actual value and the blue curve represents the predicted value.

The prediction results after BP and LSTM neural network training (taking silver and copper as an example for comparison): the left side is BP's prediction results, and the right side is LSTM's prediction results. It can be seen that the LSTM's silver prediction results are very good, and the copper prediction results are also good, slightly separated from the actual value.

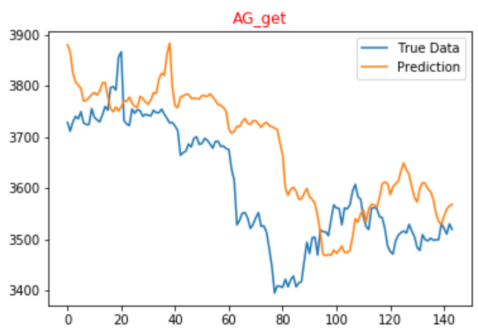


Figure 11: *Silver BP prediction*

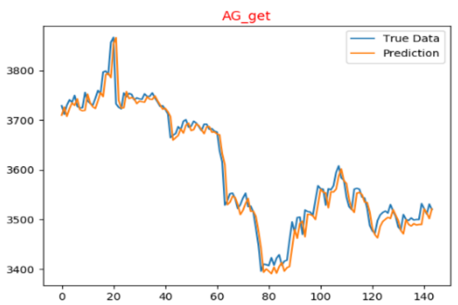


Figure 12: *Silver LSTM prediction renderings*

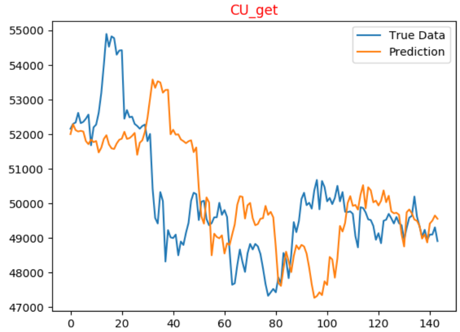


Figure 13: *Copper BP prediction effect*

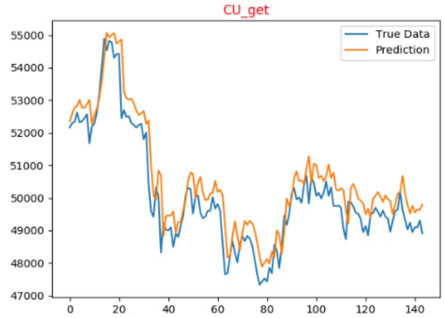


Figure 14: *Copper LSTM prediction*

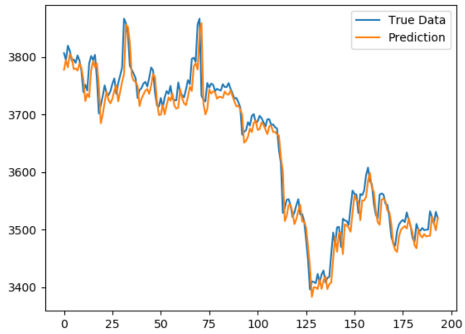


Figure 15: *Optimization of LSTM effect for sliver*

Figure 16: *Optimization of LSTM effect for copper*

By comparing BP neural network and data set training, the fitting result is not ideal, indicating that BP neural network is not a special time series of data for prediction. In addition, the study also predicted the lowest and highest prices of all futures. In the case of copper and silver, it can be seen from the forecast and the actual trend chart that the prediction effect is very good for the peak price of silver (AG_low).The lowest price of copper (CU_low) and the highest price of silver (AG_high) result in a better prediction effect, and the prediction of the highest price has deviation. By observing the curves of the above four futures prediction results and comparing the prediction results with the actual trend, it can be clearly found that the predicted output of LSTM model on the four futures prices is basically consistent with the real output on the overall trend.

4.3. Model Prediction Accuracy Analysis

In the evaluation research of accuracy, the evaluation index of regression algorithm, namely root mean square error (RMSE), is used to evaluate the accuracy of prediction. The root mean square error is also known as the standard error, which is the square root of the loss function for linear regression. You take the true value, the predicted value, and then you square it and average it, and then you take the square root. Intuitively, MSE is equivalent to putting the loss function on the test set to see the loss value. Due to the limitation of the number of observations, the best value is usually used to replace the real value. The advantage of the square root error is that it is very sensitive to the very small or very large errors in the measurement, and the precision of the measurement can be well reflected. In the RMSE calculation formula, the value on the Y test set is the actual value, and Yi is the output value of the neural network.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - Y)^2}{n}} \quad (7)$$

When the size of RMSE value reflects the size of the predicted result deviating from the actual result when multiple measurements of a variable are made, that is, the higher the measurement accuracy is, the smaller the RMSE will be. The Table 2 shows the RMSE results when four futures varieties are predicted.

Table 2: *Comparison of root mean square error of LSTM neural network in futures price prediction*

varieties	silver	copper	Coking coal	lithium
RMSE	0.0125	0.0208	0.0219	0.0233

MAE=average absolute error (MAE) is another loss function used for regression models. MAE is the sum of the absolute value of the difference between the target value and the predicted value

$$MAE = \frac{1}{m}\sum_{i=1}^{m} |\, h(x^i - y^i)\, | \qquad (8)$$

Table 3: *Comparison of the average absolute error of LSTM neural network in the prediction of futures prices*

Varieties	Silver	Copper	Coking coal	Lithium
RMSE	0.0111	0.0159	0.0181	0.0119

As can be seen from the above table, RMSE of copper, coking coal and lithium futures is larger than that of silver, and the RMSE of silver is 0.017. The predicted value is the most consistent with the graph of the real value, with almost no separation, and the prediction accuracy is the highest.

In order to further explore the reason why RMSE and MAE are not the same futures, we further analyzed the original data of silver, copper and lithium. The following four charts show the raw data distribution of futures. It can be clearly seen that the distribution of the original data of silver is relatively regular and concentrated, and the overall figure is closer to the normal distribution, while that of coking coal is poor in terms of distribution regularity, and the data distribution is more dispersed. The lowest and highest values of lithium

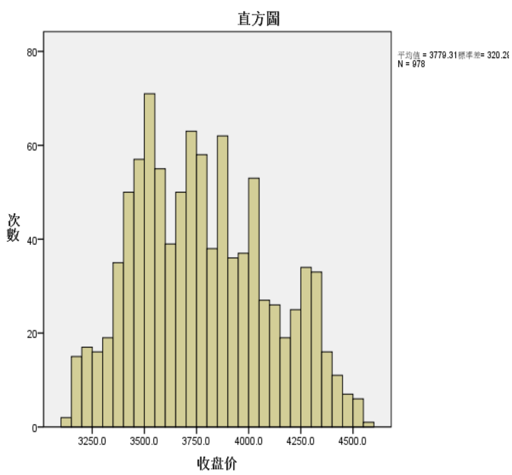


Figure 17: *Histogram of silver original data*

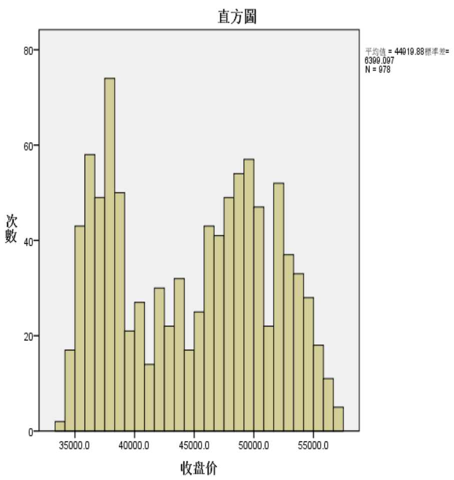
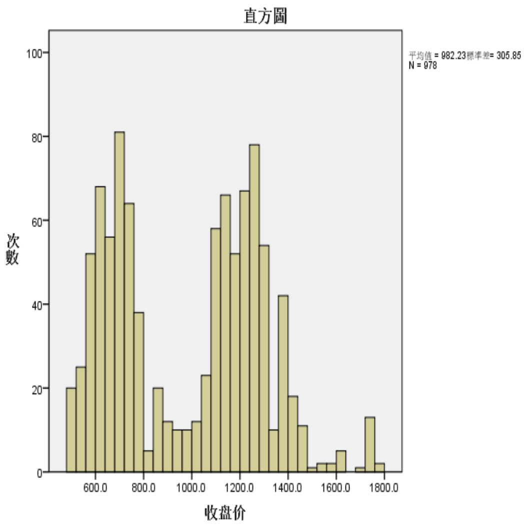


Figure 18: *Box diagram of copper original data*



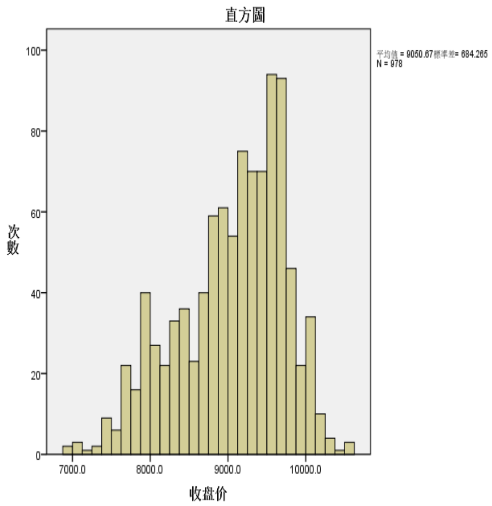


Figure 19: *Box diagram of raw data of coking coal*

Figure 20: *Distribution histogram of raw data of lithium*

The RMSE value of lithium is the largest among the four futures, and its predicted value is the most discrete with the actual result, and the prediction effect is the worst. The difference between the maximum value and the minimum value of lithium is prominent. RMSE is more sensitive to outliers, and the above results are consistent with the principle of RMSE.MAE of coking coal is the largest, that is, the absolute average value is the largest, which means the dispersion degree of coking coal is large.

Through the distribution of the original data, it also shows that the LSTM neural model has different prediction effects on different futures, which is related to the model and the dispersion degree of the distribution of individual futures data. The LSTM neural network model in this study has a root-mean-square error value in the range of (0,0.1) for the prediction of the future price trend, which fully shows that the prediction accuracy is very high, and the model is applicable to the future price with good performance and significant prediction effect.

5. Conclusions and Prospects

All BP neural networks applied in stock market prediction belong to the theoretical category of static neural networks. Such neural networks have no memory function and are excessively dependent on the current input. Therefore, static neural networks cannot well reflect the dynamic characteristics of stock and futures markets. LSTM neural network increases the horizontal connection between the hidden layer units, has the characteristics of memory and the influence of internal time sequence. As an RNN with memory and feedback type, LSTM neural network is a type of dynamic neural network. Compared with the original model, a hidden layer is added to improve the performance of LSTM model in predicting financial time series. The LSTM dynamic neural network is suitable for the high noise, nonlinear and unstable random price time series in the market.

For gradient problems of RNN and neural network fitting problem, the empirical use data normalization and dropout method to modify the model, dropout by adjusting the LSTM model parameters (not modifying the objective function) under the same conditions, through the contrast experiment, found that the modified LSTM model has better generalization ability, don't rely too much on some local characteristics, such as silver futures has better prediction effect.

The study by optimizing the LSTM model for time series prediction calculation of the biggest problems are often a fitting, and this article very good correction of this problem, due to network with multiple hidden layer structure, better able to learn the characteristics of futures data in the past, and be able to find out the relationship between the time series, also can use selective memory function, and do dig deep to the internal laws of the futures price. The experimental results show that no matter the closing price, the lowest price or the highest price, the results of this study are satisfactory.

In the time series prediction, uses convolution to further optimize. The size of the convolution kernel in the convolution layer is 3*3, the number of the convolution kernel is 4, and Max pooling is added to the output of the convolution kernel. However, the empirical results show that LSTM attention has a better effect, indicating that convolutional neural network is not necessarily the optimal choice for time series, while parameter adjustment is more effective. Of course, it is possible that this is the reason why the author chose only one type of convolution. In the next article, the author will further verify the convolution optimization.

The distribution of the original data directly determines the prediction accuracy of the model for the futures data. The model optimization in this study improves the applicability of the model to the new sample data by adjusting the parameters rather than the objective function. Therefore, in order to make good use of this model and obtain good returns in financial practice, it is suggested that futures varieties with the characteristics of relatively concentrated and regular distribution of the sorted original data should be selected. For futures with poor data distribution regularity, it may take too long to train the model, and the prediction effect may not be as good as expected.

Influence the cause of the futures prices is diverse, the change of the futures market is complex, both macroeconomic factors, also has the market supply and demand factors, and both parties speculative factors, so the market traders can't rely too much on technology, and to see the problem from the intrinsic nature of the economy. LSTM neural network also has its own disadvantages. In the current situation, machine learning cannot predict the long-term trend and cannot include all the influencing factors.

The selection of the overall data quantity is limited, and the results of the study, after the training model of effective for individual products, also hard to adapt to more varieties of futures market, this also is the common problem facing all the neural network, this study only discuss the forecasting accuracy and stability, not to the model in the case of applicability of futures to do deep research, so still have considerable space to further research.

Futures market, there are many factors that affect the direction of the trading price, so it is difficult to predict. The advantages of LSTM neural network model, such as the rapidity, determine the ability to quickly analyze a huge amount of data to achieve real-time prediction results, while the stability reflects the reliability of the model. These application characteristics are very important for price prediction and will have a great impact on the future returns of investors. It can be predicted that artificial neural network has wide application space in futures and stock market prediction.

References

Baek, Y., and Kim, H.Y. 2018. ModAugNet: A new forecasting framework for stock market index value with an over fitting prevention LSTM module and a prediction LSTM module. *Expert Systems with Applications*, 113, pp.457-480.

Dashti, R.A.S.E., Hamid, M., and Zakaria, F. 2011. Forecasting Iranian crude oil price using artificial neural network and arima models. *Journal of Social Policy*, 7(3), pp.382-383.

Fakhruddin, B., Mahalingam, R., and Padmanaban, R. 2018. Sustainable development goals for reducing the impact of sea level rise on mangrove forests. *Indian J. Geo-Marine Sci.*, 47(10), pp.1947-1958.

Gencay Ramazan. 1996. Non-linear prediction of security returns with moving average rules. *Journal of Forecasting*, 15(3), pp.165-174.

Grudnitski, G., and Osburn, l. 1993. Forecasting S&P and gold futures prices: An application of neural networks. *Journal of futures Markets*, 13(6), pp.631-643.

Hamid, S.A., and Iqbal, Z. 2004. Using neural networks for forecasting volatility of s & p 500 index future sprices. *Journal of Business Research*, 57(10), pp.1116-1125.

Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, Ruslan R. Salakhutdinov. 2012. Improving neural networks by preventing co-adaptation of feature detectors. *https://arxiv.org/abs/1207.0580*

Hochreiter, Sepp, and Jurgen Schmidhuber. 1997. Long short term memory. *Neural Computation,* 9(8), pp.1735-1780.

Monishiya, G.B., and Padmanaban, R.C. 2012. Mapping and change detection analysis of marine resources in the tuticorin and vembar group of island using remote sensing. *Int. J. Adv. Forest Sci. and Management*, 2(2), p.132.

Padmanaban, R., Bhowmik, A.K., and Cabral, P. 2017. A remote sensing approach to environmental monitoring in a reclaimed mine area. *ISPRS Int. J. Geo-Information*, 6(12), doi:10.3390/ijgi6120401.

Padmanaban, R, Bhowmik, A.K., and Cabral, P. 2019. Satellite image fusion to detect changing surface permeability and emerging urban heat islands in a fast-growing city. *PLoS One*, 14(1):e0208949.

Padmanaban, R. 2019. A remote sensing approach to the quantification of local to global scale social-ecological impacts of anthropogenic landscape changes. *https://www.semanticscholar.org/paper/A-remote-sensing-approach-to-the-quantification-of-Padmanaban/b9bb88620d546bdd5837c834445c8deff7c7020d*

Padmanaban, R.C. 2012. Integrating of Urban Growth Modelling and Utility Management System using Spatio Temporal Data Mining. *Int. J. Adv. Earth Sci. Eng,* 1, pp.13-15.

Padmanaban, R., Karuppasamy S., and Narayanan R. 2018. Assessment of pollutant level and forecasting water pollution of Chennai coastal, TamilNadu using R. *Indian J. Geo-Marine Sci.* 47(7), pp.1420-1429.

Padmanban, R, and Painho, M. 2017. Urban Agent Based Model of Urban SlumDharavi, Mumbai, *International Journal of Earth Sciences and Engineering*, 10(6), pp.1110-1117.

Qiu, M., Song, Y., and Akagi, F. 2016. Application of artificial neural network for the prediction of stock market returns: the case of the Japanese stock market. *Chaos, Solitons & Fractals*, 85, pp.1-7.

Rajchandar, P. 2012. Modelling the Transformation of Land use and Monitoring and Mapping of Environmental Impact with the help of Remote Sensing and GIS. *Int. J. Adv. Altern. Energy Environ. Ecol.* 1, pp.36–38.

Rajchandar, P., Bhowmik, A.K., Cabral, P., Zamyatin, A., Almegdadi, O., and Wang, S. 2017. Modelling Urban Sprawl Using Remotely Sensed Data: A Case Study of Chennai City, Tamilnadu. *Entropy*, 19(4), p.163.

Rumelhart, David E., Geoffrey E. Hinton, and James l. McClelland. 1986. A general framework for parallel distributed processing. Parallel distributed processing: Explorations in the microstructure of cognition, 1, pp.45-76.

Venkatesan, G., and Padmanaban, R. 2012. Possibility Studies and Parameter Finding for Interlinking of Thamirabarani and Vaigai Rivers in Tamil Nadu. *India. Int. J. Adv. Earth Sci. Eng.* 1(1), pp.16–26.

Visalatchi, A., and Padmanaban, R. 2012. Land use and land cover mapping and shore line changes studies in Tuticorin coastal area using remote sensing. *Int. J. Adv. Earth Sci. Eng*. 1, pp.1-12.

Wang, J. 2010. Stock Market Volatility and the Forecasting Performance of Stock Index Futures Journal of Forecasting, 28(4), pp.277-292.

Wen, Y., and Yuan, B. Use CNN-LSTM network to analyze secondary market data. Proceedings of the 2nd International Conference on Innovation in Artificial Intelligence. 2018, pp.54-58.

White. Economic prediction using neural networks: the case of IBM daily stock returns. IEEE 1988 International Conference on neural networks.

Zhang, G. Peter. 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, pp.159-175.