Stock Price Prediction Based on Wavelet Analysis and Neural Network

Jing Ren

Nanjing University of Science and Technology, Nanjing, China

Keywords: Wavelet Analysis, BP Neural Network, LSTM Neural Network

Abstract: This paper takes the wavelet analysis into the prediction of BP and LSTM neural networks, and selects the closing price of the Shanghai Composite Index (000001) from July 1, 2016 to June 30, 2021. In order to obtain the yield, this paper takes the first-order difference of the natural logarithm. Through wavelet analysis, the original yield sequence is decomposed and reconstructed. Then, the BP and LSTM neural network models are constructed with the reconstructed sequence, and the prediction effect of the two models is compared and analyzed. The result shows the fitting effect of LSTM is better than that of BP neural network. Further, the BP and LSTM neural network models of the original yield sequence are compared with the wavelet reconstructed model, and it is found that the wavelet decomposition and reconstruction can improve the accuracy of prediction.

1. Introduction

Time series analysis has always been the focus of research in the financial field, many factors affecting the time series data and there may be intricate relationships between different factors. Traditional methods predicting time series includes ARIMA, ARCH, GARCH, these models can fit the linear part of the sequence. In recent years, neural networks have emerged to test and predict time series, study learning can extract the effective information from huge data to test and has unique advantage in dealing with nonlinear problem.

Liu Xiangli and Wang Xupeng divide yield data into high frequency parts and low frequency part and reconstruct the splitting sequence, then they fit the ARIMA model, the results show that the wavelet multi-resolution analysis can filter out days effect, decomposition reconstruction model can improve the prediction accuracy [1]. Yuan Yujia decomposes the raw data into low frequency and high frequency sequences by the wavelet decomposition and reconstruction algorithm, then she take each subsequence into LSTM model training finding WA - LSTM model has certain advantages in fitting precision and prediction accuracy [2].

2. Theoretical framework

2.1. The structure of BP neural network

BP neural network is a multilayer feed-forward neural network trained on an error reverse propagation algorithm. The topology of the BP neural network model can be divided into three levels: input layer, implicit layer and output layer. The forward propagation process begins from the input layer, through the implicit layer processing, and turns to the output layer, if the output layer cannot get the desired output, then turn to backpropagation, the error signal returned along the original connection path, by modifying the weight of each neuron, so that the error signal is minimized.

Assuming that the number of neurons in the input layer, output layer, and hidden layer of a neural network is n, m, and p, respectively, the output of the neuron is: $x_i^j = \sigma \sum_{i=1}^{n} w_i x_i + b_j$, j=1, 2, …,p. The output of the output layer neurons is: $y^k = \sigma \sum_{j=1}^{p} w_{ik} x_j + b_k$, k=1, 2, …, m. Excitation functions usually use S-type functions, such as: $\sigma(x) = \frac{1}{1+e^{-x/Q}}$, where Q is an S-type parameter that adjusts the form of the excitation function.

2.2. The structure of LSTM neural network

Long short-term memory network (LSTM) is a special recurrent neural network (RNN), it trains the input x of the current node and the input h of the previous node through a given function, then obtain the output y of the current node and the output h' passed to the next node, and thus loop down. On the basis of RNN, LSTM divides the input h^t of the previous node into two parts c^t and h^t, respectively, indicating the slow change part and the fast change part, which can well solve the gradient disappearance and gradient explosion problems in the long sequence training process.

3. Empirical analysis

In this paper, the closing price of the Shanghai Composite Index (R1) is selected for research, and the time range is from July 1, 2016 to June 30, 2021, with a total of 1216 observations. Stock-related data comes from Flush.

In this section, the wavelets of the original yield sequence are decomposed and reconstructed, and then BP and LSTM neural networks are built for comparative analysis. Furtherly, BP and LSTM neural networks are built using the original yield sequence to verify whether the wavelet analysis can improve the accuracy of model fitting.

3.1. Wavelet decomposition and reconstruction

(1) The determination of wavelet layers and N

The decomposition and reconstruction of 3-layer, 4-layer and 5-layer wavelet were carried out respectively. After the reconstruction, MAE was used to measure the effect [3].

By comparing db1-db7 with 3-layer, 4-layer and 5-layer, it is found that db3 with 3-layer has the best effect and MAE is the smallest. Therefore, three-layer DB3 is used for wavelet decomposition.

(2) Wavelet decomposition and reconstruction

Shanghai Composite Index R1 was decomposed and reconstructed by three-layer db3 wavelet. During the reconstruction, high-frequency information d2 and d1 were set to zero. The decomposition coefficient and reconstructed sequence are shown in Figure 1 and Figure 2.

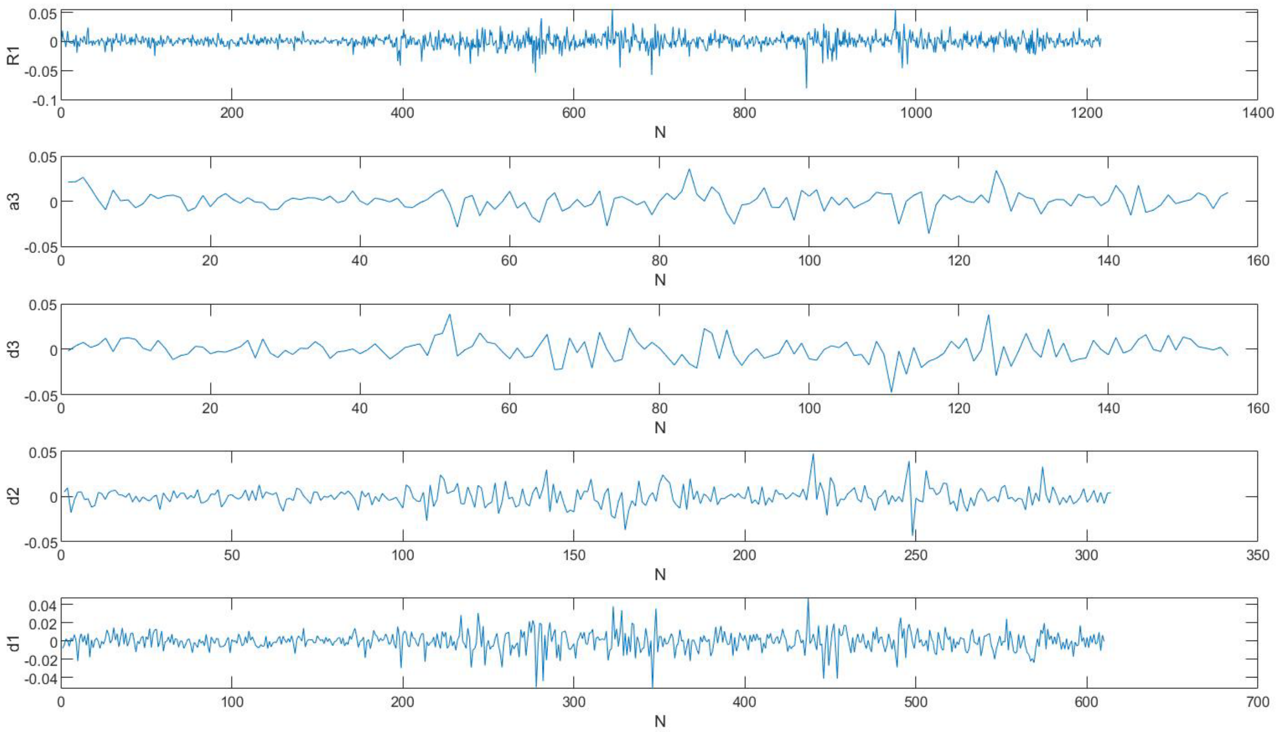

Figure 1. Low frequency and high frequency coefficients after wavelet decomposition

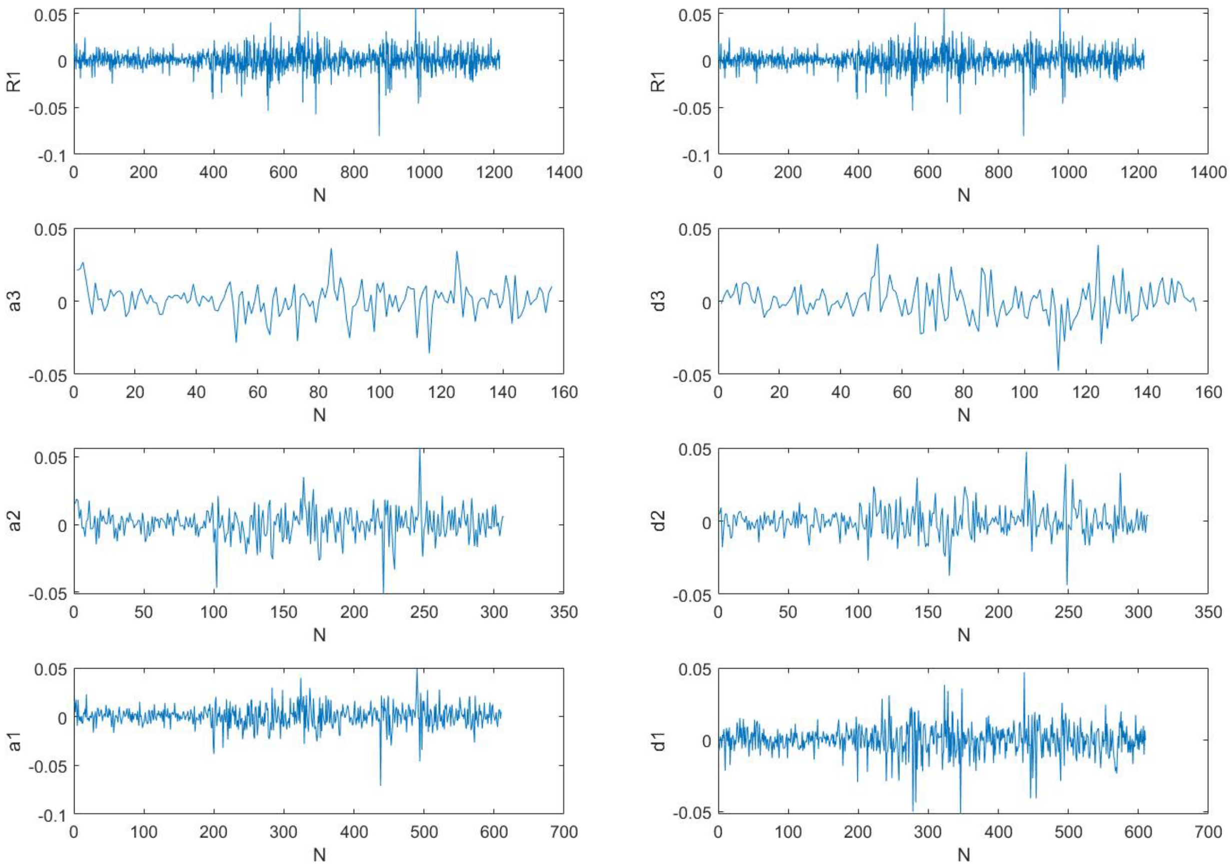

Figure 2. Wavelet reconstruction sequence

R1 is the trend diagram of the original sequence. After 3-layer db3 decomposition, the approximate part a3 and the detail part d1, d2 and d3 of the original sequence are obtained. It is found that the approximate part a3 after 3-layer decomposition retains the trend of the sequence well. Then we reconstructed sequence by zeroing the high frequency information, which not only retains the original features of the sequence, but also removes most of the small fluctuation noise.

3.2. Results

Use Mean Absolute Error (MAE), Mean Square Error (MSE) and Root Mean Square Error (RMSE) to evaluate the fitting effect of WA-BP and WA-LSTM on test sets [4].

Table.1. Comparison of WA-BP and WA-LSTM prediction effect

	MAE	MSE	RMSE
WA-BP	0.001612	5.6653e-06	0.0023802
WA-LSTM	0.00080473	1.4652e-06	0.0012104

According to table 1, the MAE, MSE and RMSE values of WA-LSTM model are all smaller than those of MA-BP model, indicating that the LSTM neural network with wavelet analysis has better prediction effect

To further verify whether wavelet analysis improves the accuracy of neural network model fitting, BP and LSTM neural networks were built respectively with the original rate of return sequence and compared with WA-BP and WA-LSTM. The results are shown in Table 3.

Table.2. Comparison of prediction effect

	MAE	MSE	RMSE
BP	0.0085598	0.00014821	0.012174
LSTM	0.013945	0.00033473	0.018296
WA-BP	0.001612	5.6653e-06	0.0023802
WA-LSTM	0.00080473	1.4652e-06	0.0012104

The fitting effect of BP and LSTM after wavelet reconstruction is better than the original rate of return series, indicating that wavelet reconstruction does improve the accuracy of model fitting.

4. Conclusion

This paper analyses the yield of the Shanghai Composite Index from July 1, 2016 to June 30, 2021 by using wavelet decomposition and reconstruction technology, it is found that the fitting accuracy of BP and LSTM neural networks based on wavelet analysis is improved. Besides, comparing the prediction errors of LSTM and BP, the fitting effect of LSTM is significantly better than that of BP model.

References

[1] Liu Xiangli and Wang Xupeng. Research on high frequency forecast of Stock index futures based on wavelet analysis [J]. Systems Engineering-Theory & Practice,2015,35(06):1425-1432.

[2] Yuan Yujia. Application of short and long time memory model based on wavelet analysis in stock index forecasting [D]. Shandong University,2020.

[3] Fan Wen, Zhang Pulin. Application of wavelet neural Network in stock price prediction [J]. Public Communication of Science & Technology,2011(14):187-188.

[4] OuYang Hongbing, Huang Kang, Yan Hongju. Financial time series prediction based on LSTM neural network [J]. Chinese Journal of Management Science,2020,28(04):27-35.