The Prediction of Stock Price Based on Improved Wavelet Neural Network

Qinglan Ye, Lianxin Wei*

Department of Applied Mathematics, College of Science, University of Shanghai for Science and Technology, Shanghai, China
Email: *weilianxin@usst.edu.cn, 18301913089@163.com

Abstract

To improve the accuracy of forecasting stock prices, a new method is proposed, which based on improved Wavelet Neural Network (WNN). Firstly, the Genetic Algorithm (GA) is used to optimize initial weights, stretching parameters and movement parameters. Then, comparing with traditional WNN, the momentum are added in parameters adjusting and learning of network, what's more, learning rate and the factor of momentum are self-adaptive. The prediction system is tested using Shanghai Index data, simulation result shows that improved WNN performs very well.

Keywords

WNN, Forecasting Stock Prices, Momentum, Learning Rate, Self-Adaptive

1. Introduction

The stock prices are time series data with multiple variables, there are non-linear, time-varying and uncertain relationship, and have always been a challenge to both economists and researchers [1] [2].

A variety of models have been established to predict the stock prices by observing the law of data, such as Autoregressive Integrated Moving Average Model (ARIMA) [3], support vector machine [4] and neural network (NN), wavelet neural network (WNN) [5] model. ARIMA is typical linear time series analysis method, but stock market is a large of nonlinear dynamic system and share price index is out of order. So if ARIMA model predicts complex time series, forecasting result often don't be ideal. SVM is a non-traditional nonlinear forecasting technique based on the structure risk minimization principle. It can be use the theory of minimizing the structure risk to avoid the problems of excessive study, calamity data, local minimal value and so on. However, the regression cannot approach every function. With the rapid development of artificial intelligence technology,

---

*Corresponding author.

NN is employed as the nonlinear predictor for financial time series. Although prediction results usually outperform ARIMA models, traditional NN models have several disadvantages, including: a lot of model parameters should be dependent on, easy to be trapped into local minima, and over-fitting on training data resulting in poor generalization ability. As the wavelet analysis theory developed in the mid-1980, WNN was first proposed by Qinghua Zhang, and has been widely applied [6], WNN has following advantages, one is that as the low correlation of wavelet neuron, WNN has fast convergence. What's more, stretching and movement factors of wavelet function make the approximation capability of network more powerful. The third is wavelet neural network matches the signal with the good partial characteristic and multiresolution learning, which can express function characteristic and has higher forecast accuracy [7]. But there still have two disadvantages: First, the learning rate is constant, so if there is large, the training may over convergence; Second, WNN training method usually includes the stochastic gradient algorithm and declined gradient, those only consider the nth while ignoring the previous direction of n times, which make training trapped into local minima [8].

Aiming at the shortcomings of WNN, the paper proposes the stocks market modeling and forecasting by using improved WNN. In parameters adjusting and learning of network, momentum items are added, in the same time, learning rate and the factor of momentum are self-adaptive and the form of decreasing factor of learning rate is connected with forecasting error. Using improved WNN to predict Shanghai Stock market, it shows that improved WNN is superior to WNN.

2. Models

2.1. Wavelet Neural Network (WNN)

WNN is neural network model composed by wavelet function as its activation function. Structural diagram of three-layer wavelet neural network is shown in **Figure 1**. $X_1, X_2, \cdots, X_k$ are input of network, $Y_1, Y_2, \cdots, Y_m$ are output, W_{ij} represents the weight between input layer and hidden layer, W_{jk} is weight between hidden layer and output layer. In this time, the input and output of hidden layer as follows,

$$net_j = \sum_{i=1}^{k} W_{ij} x_i, \quad j = 1, 2, \cdots, l$$

$$\varphi(j) = \varphi_j \left(\frac{net_j - b_j}{a_j} \right), \quad j = 1, 2, \cdots, l$$

the output of network as follows,

$$y(k) = \sum_{j=1}^{l} W_{jk} \varphi(j), \quad k = 1, 2, \cdots, m$$

φ_j is wavelet function, a_j is movement parament, b_j is stretching parameter, the number of hidden layer
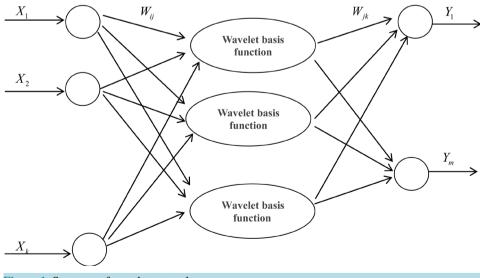

Figure 1. Structure of wavelet network.

nodes is l, k is the number of input layer nodes and m is the output layer nodes'.

The correction of parameters such as weights, movement parameters and stretching parameters as follows,

$$W_{i,j}(n+1) = W_{i,j}(n) - \eta 1 \cdot \frac{\partial e}{\partial W_{i,j}(n)}$$

$$W_{j,k}(n+1) = W_{j,k}(n) - \eta 1 \cdot \frac{\partial e}{\partial W_{j,k}(n)}$$

$$a_k(n+1) = a_k(n) - \eta 2 \cdot \frac{\partial e}{\partial a_k(n)}$$

$$b_k(n+1) = b_k(n) - \eta 2 \cdot \frac{\partial e}{\partial b_k(n)}$$

In the formulas, $e = \sum_{k=1}^{m} yn(k) - y(k)$ and represents the error of output of network, $yn(k)$ represents expected output, $\eta 1, \eta 2$ represent learning rate.

2.2. Improved Wavelet Neural Network

As the learning rate is constant, if there is large, the training may over convergence, in the same time WNN training method usually include the stochastic gradient algorithm and declined gradient, those only consider the nth while ignoring the previous direction of n times, which make training trapped into local minima. In order to overcome the shortcomings in the correcting, we propose the formulas of correction as follows,

$$W_{i,j}(n+1) = W_{i,j}(n) - \eta 1(n) \cdot \frac{\partial e}{\partial W_{i,j}(n)} + k(n) \cdot \left(W_{i,j}(n) - W_{i,j}(n-1) \right)$$

$$W_{j,k}(n+1) = W_{j,k}(n) - \eta 1(n) \cdot \frac{\partial e}{\partial W_{j,k}(n)} + k(n) \cdot \left(W_{j,k}(n) - W_{j,k}(n-1) \right)$$

$$a_k(n+1) = a_k(n) - \eta 2(n) \cdot \frac{\partial e}{\partial a_k(n)} + k(n) \cdot \left(a_k(n) - a_k(n-1) \right)$$

$$b_k(n+1) = b_k(n) - \eta 2(n) \cdot \frac{\partial e}{\partial b_k(n)} + k(n) \cdot \left(b_k(n) - b_k(n-1) \right)$$

where, $\eta 1(n), \eta 2(n)$ represent the nth learning rate and $k(n)$ is the nth momentum factor, and in this paper, the formulas of them are expressed as,

$$\eta 1(n) = \begin{cases} e^{-\Delta E} \cdot \eta 1(n-1) & \Delta E > 0 \\ \eta 1(n) & \Delta E \leq 0 \end{cases}$$

$$\eta 2(n) = \begin{cases} e^{-\Delta E} \cdot \eta 2(n-1) & \Delta E > 0 \\ \eta 2(n) & \Delta E \leq 0 \end{cases}$$

$$k(n) = \begin{cases} 0 & \Delta E > 0 \\ k(n-1) & \Delta E \leq 0 \end{cases}$$

In the formulas, $\Delta E = e(n) - e(n-1)$, if $\Delta E > 0$, the learning error increased, it is better to decrease the learning rate and remove the momentum; if $\Delta E \leq 0$, the learning error decreased and the direction of correction is right, so we keep learning rate and momentum unchanged.

2.3. The Training Process by Improved WNN Model

Step 1. Considering when the initial parameters is not suitably selected, the WNN training will slow or even

diverging [9], so it is vital to use GA to optimize initial weights, stretching parameters and movement parameters [10].

Step 2. Setting initial learning rate $\eta1, \eta2$ and momentum factor k.

Step 3. Dividing sample into training sample and test sample. Training sample is used to train the network and test sample is used to test the prediction accuracy.

Step 4. Putting test sample as input of network and calculating the error between expected output and actual output.

Step 5. Making output of network is close to expected output, using error to adjust weights, stretching parameters and movement parameters.

Step 6. Judging the algorithm is end or not. If it is not, return Step 3.

3. Case Study

3.1. Data Processing

Taking Shanghai stock market for example, the chosen time series is the closing index of daily stock price in Shanghai Stock exchange from March 23th, 2012 to November 28th, 2014. This yields a total of 607 data samples, in which, the former 507 data samples are selected as training samples, the later 100 data samples as test samples. To improve the prediction accuracy, the data are normalized, the formula as follows,

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

where x' represents the data of normalized processing, x represents original data, $x_{\max}$, $x_{\min}$ represent the maximum and the minimum number in data.

3.2. Training of Network

To stock market, taking into account the fact that there are five trading days a week, a WNN forecast model of one five-dimension input data [11] and one output data is established and set that the number of neurons in hidden layer is 6. We choose three-layer network because it can approximate arbitrary function. So the structure of WNN is 5-6-1. If the activation function adopts different wavelet functions, then there will be different performance [12]. In this paper, we choose Morlet wavelet function as activation function. Learning rate $\eta1, \eta2$ are 0.01and 0.001 separately. The expected error is 0.001 and the number of iterations is 100; Improved WNN has the same structure with WNN. The initial learning rate $\eta1, \eta2$ are also 0.01 and 0.001 separately; initial momentum factor k is 0.9; the expected error is 0.001; and the number of iterations is 100. GA is used to optimize initial weights, stretching parameters and movement parameters of WNN and improved WNN. The number of iterations is 50. What's more, in GA, the population size is 40, the number of evolution is 50, crossover probability is 0.2 and mutation probability is 0.1.

3.3. Simulation and Evaluation

The paper use root mean square error (RMSE) and mean absolute percentage error (MAPE) as evaluation standard of network, which are defined as

$$\text{RMSE} = \sqrt{\frac{\sum_i^n (y_i - y_i')^2}{n}}$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - y_i'}{y_i} \right|$$

where n refers to the number of test sample, $(y_1, y_2, \cdots, y_n)$ refers to actual output of network, $(y_1', y_2', \cdots, y_n')$ refers to desired output.

Figure 2 shows the forecasting result of WNN improved WNN whose initial parameters were optimized by GA. Their RMSE and MAPE were shown by **Table 1**.

Figure 2 we can see after optimized by GA, improved WNN has a better forecasting ability than WNN.

Figure 2. WNN forecasting and Improved WNN forecasting.

Table 1. Comparison of forecasting error in different methods.

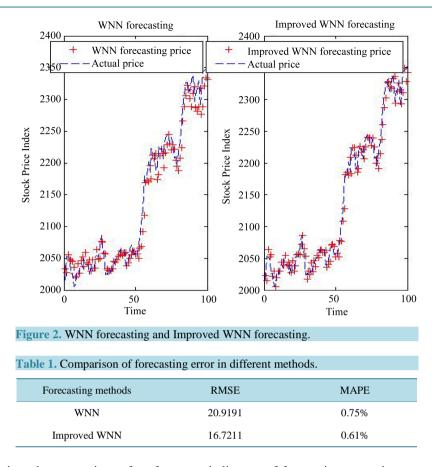| Forecasting methods	RMSE	MAPE
WNN	20.9191	0.75%
Improved WNN	16.7211	0.61%

Table 1 summarizes the comparison of performance indicators of forecasting errors between the two models. Comparing with two methods, it can be see that improved WNN improves forecast accuracy, and it substantially reduces the RMSE by about 20% and MAPE by about 19%.

4. Conclusion

In this paper, the application of improved wavelet neural network to stock market prediction is studied. Considering that the forecasting accuracy is easy to be affected by initial parameters, which can provide GA to optimize it at first. The forecasting simulation results of Shanghai index data show that the improved WNN method is effective and the stock market model is of good prediction performance.

References

[1] Gupta, S. and Wang, L.P. (2010) Stock Forecasting with Feed forward Neural Networks and Gradual Data Sub-Sampling. *Journal of Intelligent Information Systems*, **11**.

[2] Zhu, M. and Wang, L.P. (2010) Intelligent Trading Using Support Vector Regression and Multilayer Perceptrons Optimized with Genetic Algorithms. *The 2010 International Joint Conference on Neural Networks (IJCNN)*, Barcelona, 18-23 July 2010, 1-5.

[3] Cai, Z.X., Wu, C.H. and Zhang, D.X. (2002) Network Simulation FTP Business Prospect of Mathematical Modeling and Simulation Flow. *Computer Simulation*, **19**, 104-106.

[4] Yu, G.Q. and Zhang, C.S. (2004) Switching ARIMA Model Based Forecasting for Traffic Flow. *ICASSP*, **2**, 429-432.

[5] Vapnik, V.N. (1995) The Nature of Statistical Learning Theory. Springer-Verlag, New York.

[6] Donoho, D.L. and Johnstone, I.M. (1994) Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrika*, **3**, 425-455.

[7] Zhao, Y., Zhang, Y. and Qi, C.J. (2008) Prediction Model of Stock Market Returns Based on Wavelet Neural Network. *Computational Intelligence and Industrial Application*, **1**, 31-36. http://dx.doi.org/10.1109/PACIIA.2008.46

[8] Oussar, Y. and Dreyfus, G.H. (2000) Initialization by Selection for Wavelet Network Training. *Neurocomputing*, **34**,

131-143. http://dx.doi.org/10.1016/S0925-2312(00)00295-2

[9]   Donoho, D.L. (1995) Denoising by Soft-Thresholding. *IEEE Transaction on Information*, **3**, 613-627.

[10] Xu, Q. and Shu, L.C. (2010) Application of Optimized Wavelet Neural Network Based on Genetic Algorithm in Groundwater Level Prediction. *Hydrology*, **1**, 27-30.

[11] Zhou, H.R. and Wei, Y.H. (2010) Stocks Market Modeling and Forecasting Based on HGA and Wavelet Neural Networks. *Natural Computation*, **2**, 620-625.

[12] Fang, Y. and Fataliyev, K. (2014) Improving the Genetic-Algorithm-Optimized Wavelet Neural Network for Stock Market Prediction. 2014 *International Joint Conference on Neural Networks* (*IJCNN*), Beijing, 6-11 July 2014, 3038-3042.