

A Novel Data Mining Algorithm based on BP Neural Network and Its Applications on Stock Price Prediction

Shu Zhan^{1, a}, Weihao Li², Xuan Zhuang³

¹University of Iowa, Tippie College of Business, Iowa, USA;

²School of Electronic Engineering, Intelligent Science and Technology, Xidian University, Xi'an, China;

³School of Electronic Engineering, Xidian University, Xi'an, China

^azhanshu15@sina.com

Keywords: Data Mining, BP Neural Network, Stock Price, Data Prediction, Applications.

Abstract. As an indispensable part in financial industry, stock investment has gradually become more and more important in people's life. The major fluctuation in stock market directly affects the stability of financial markets and the healthy development of the economy. The successful stock price and trend prediction will help investors to yield profit and provide timely and reasonable guidance to market regulators. Therefore, how to forecast stock market is a crucial and worthy topic in financial area. Neural network simulated how brain functions, can automatically extract features from the historical data and related information to predicate future trend, with the primary characteristics of self-organization and adaption. Therefore, neural network is suitable for processing prediction in the field of nonlinear time series. This paper presents a novel data mining algorithm based on the BP neural network and its applications on stock price prediction. Taking improving the efficiency of the traditional back propagation network, we revise the perspective error, and modify incentive function and the adaptive learning rate in order to avoid the local minimum and decrease training time. Moreover, to eliminate the influences caused by noisy data, we combine the fuzzy rough set theory to pre-process the original data. The experimental result verifies feasibility of our model, without the local minimum and slow convergence speed, compared to the other state-of-the-art approaches. Therefore, our algorithm on the predications of the exact stock price value and future trend are applicable and satisfactory.

Introduction

Before introducing our proposed algorithm, we firstly demonstrate the prerequisites for reasonable stock price prediction as the follows. (1) Market behavior reflects all information. This theory points out all the factors that affect the market price, including politics, economy, psychology, information such as the company's performance are reflected in the price, and digested by the price in advance. The time and effort on the study of these factors are not necessary, and investors should try to conduct research on the price change itself and the regularity. (2) Once the price trend is formed, it tends to last for some time. The trend is the core component of the technical analysis. When a trend just happened, it will be timely and accurately exposed. The trend of the premise is once formed; it is often along with the existing trend and continues to evolve, until at a certain time. And then this trend turned into another trend. (3) The history will reappear and repeat. In reality, the technical analysts believe that as long as the people's behavior, psychological and environmental characteristics are connected with the similar situation of market movements in the history, we can predict the market price trend after associated comparison and analysis.

The existing stock prediction techniques could be generally summarized as listed two categories. (1) Fundamental analysis. Fundamental analysis is also called basis analysis, which is based on the basic principles of the financial management and investment, to determine the stock value and price. The basic elements such as the movements of macroeconomic indicators, economic policy, industry

development, the product market and the company performance can affect and indicate people's reaction and judgment to the future market. The reasonable price change will be reflected in the market. (2) Technical analysis. Technical analysis is only from the stock market data to analyze the trend of future stock price changes. Securities market data can reflect in a variety of forms, such as the securities market price, trading volume, price and quantity changes. [1-2]. Besides the fundamental and technical analysis, we will introduce an innovated model on revised BP neural network and data mining technique for stock prediction in this paper. Artificial- neural- network- based price forecast model is helpful to strengthen securities investment decisions to be more scientific. As the development of financial theory and artificial neural network technology are getting mature, the artificial neural network technology is widely used in the financial forecast tasks. The rest of the paper is organized as the following sections. We firstly analyze the BP neural network theoretically and point out drawbacks. Then, to optimize the traditional BP algorithm, we incorporate the manipulations of error, incentive function modification and the adaptive learning rate adjustment to propose the enhanced BP network. To overcome the condition with noisy data, we combine the fuzzy rough set theory to serve as the data pre-processing operation. Later, we simulate algorithm compared with others and draw the conclusion.

The Proposed Methodology

The Traditional BP Neural Network. The BP neural network is to use an error back propagation algorithm for training multilayer feedforward neural network. It consists of input layer, hidden layer and output layer. Hidden layer is for a layer or layers, and each layer of neurons called nodes or unit. To design the network topology, the key task is to find out the optimal number of hidden layer and the hidden layer nodes. Theory has been proved that if the learning function is discontinuous function, we only need two hidden layer structured BP network, in addition with a single hidden layer of the BP network, which can realize the arbitrary continuous function mapping. In the figure one, we demonstrate the primary structure of the traditional BP neural network.

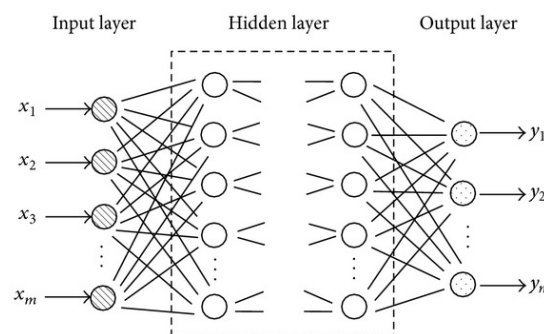


Fig. 1 The Structure of the Traditional BP Neural Network

As the layer with most complexity, hidden layer should be paid special attention to. Hidden layer is between input layer and output layer. As internal representation of the input mode, namely the type of input mode is different from other types of input mode characteristics were extracted, and will draw out the characteristics of the passed to the output layer, while the output layer will judge the final category of the input mode. Therefore, the hidden layer is also referred to be the feature extraction tool that will assist finalizing the adjustment of the weight values inside of the network.

Theoretically, the principle of the BP neural network could be described as the follows. (1) The positive work signal transmission. Input signal starts from the input layer through the hidden layer to the output layer, the output signal could be obtained from the output side. In the signal transmission during the process of forward network, the weights are fixed and each layer of the neurons state affects only the next layer of neurons state. If the output layer does not achieve the desired output, the signal will turn into the signal back propagation process. (2) Back propagation error signal. In the process of back propagation of error signal, the network's weights are adjusted by the error feedback by modifying weights of the network and the actual output is more close to the desired output.

Suppose the input layer of the network contains I nodes, the hidden layer contains J nodes, the final output layer contains K nodes, respectively. The network input, actual output and desired output are defined in the following formula one where the $p = 1, 2, 3, \dots, P$ and P represents the sample size.

$$X_p = (x_{p1}, x_{p2}, \dots, x_{pI})' \quad O_p = (o_{p1}, o_{p2}, \dots, o_{pK})' \quad T_p = (t_{p1}, t_{p2}, \dots, t_{pK})' \quad (1)$$

In addition, the w_{ij} represents the weight from input node i to hidden node j , the w_{jk} denotes weight from hidden node j to the output layer k , $(o_{p1}, o_{p2}, \dots, o_{pJ})$ represents the output of the hidden layer. The outputs of the hidden layer and output layer are shown in the equation 2 and 3, respectively.

$$o_{pj}' = f\left(\sum_{i=1}^I w_{ij}' x_{pi}\right), \quad j = 1, 2, \dots, J \quad (2)$$

$$o_{pk} = f\left(\sum_{j=1}^J w_{jk} o_{pj}'\right), \quad k = 1, 2, \dots, K \quad (3)$$

The incentive function for the network is selected to be the Sigmoid as the formula 4.

$$f(x) = 1 / (1 + \exp(-x)) \quad (4)$$

To train and adjust weight of the network, we need back training process. The error measurement standard adopted for the BP is square error defined in equation 5.

$$E_{overall} = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^K (t_{pk} - o_{pk})^2 \quad (5)$$

Based on the prior discussed procedures, we could determine each neuron weights' adjustment formulas for the output and hidden layers as the equation 6~7.

$$\Delta w_{jk} = \eta \sum_{p=1}^P \left(-\frac{\partial E_p}{\partial net_{pk}} \times \frac{\partial net_{pk}}{\partial w_{jk}} \right) = \eta \sum_{p=1}^P (t_{pk} - o_{pk}) o_{pk} (1 - o_{pk}) o_{pj}' \quad (6)$$

$$\Delta w_{ij}' = \eta \sum_{p=1}^P \left(-\frac{\partial E_p}{\partial net_{pj}'} \times \frac{\partial net_{pj}'}{\partial w_{ij}'} \right) = \eta \sum_{p=1}^P \left(\sum_{k=1}^K \delta_{pk} w_{jk} \right) o_{pj}' (1 - o_{pj}') x_{pi} \quad (7)$$

The Network Optimization Approaches. Before optimizing the traditional BP neural network, we firstly summarize the corresponding drawbacks as the follows. (1) Lack of structural parameters of the guidance in theory, the choice of network structure is a unified and complete theoretical guidance and usually can only be chosen according to the experience. (2) Network generalization ability of learning is closely related to the sample representativeness. However, typical samples selected for the training set is a very difficult problem. (3) Traditional BP algorithm uses the local search optimization method, but global optimization is used for - solving the complex nonlinear function, which is likely to fall into local extremum. (4) Due to the complexity of solving optimization of objective function, it is bound to be under the condition of the neuron output, which tends to be 0 or 1 and forms some flat areas. In these areas, the weight error change is very small which results the training process almost to halt.

To eliminate mentioned drawbacks, we propose the following countermeasures regarding of perspectives of error, incentive function modification and the adaptive learning rate. The incentive function of traditional neural network is usually adopting Sigmoid function which holds no variable parameters. So the function model is relatively fixed, and therefore its steepness, location and mapping range are also fixed. The Sigmoid function could not automatically find the optimal function body and often leads to fall into local minimum point during training with slow convergence speed. Moreover, its generalization ability is not strong. To assist the incentive function getting out of the saturation region, we propose the piecewise function for the potential optimization as formula 8.

$$f(x) = \begin{cases} 1, & \gamma x > \pi / 2 \\ [\sin(\gamma x + 1)] / 2, & -\pi / 2 < \gamma x < \pi / 2 \\ 0, & \gamma x < -\pi / 2 \end{cases} \quad (8)$$

Standard BP algorithm adopts the formula 5 to serve as the error feedback function. However, in the process of training the network, when the input works in a saturated zone incentive function, the error function gets convergence and tends to be a constant, which causes the paralysis, the network

can't completely finish the training. To avoid a paralysis phenomenon, we usually revise the error function based on Baum's theory as the equation 9.

$$E_{revised} = \sum_i \left[(1 + y_i) / 2 \log \left((1 + y_i) / (1 + d_j) \right) + (1 - y_i) / 2 \log \left((1 - y_i) / (1 - d_j) \right) \right] \quad (9)$$

In the traditional BP neural network, the adjustment law of the weights across layers refers to the formula 10, where the η represents the learning rate.

$$w_{k+1} = w_k - \eta \frac{\partial E}{\partial w} \quad (10)$$

The selection of the η value will influence the overall performance of the neural network in a large extent. The adaptive learning algorithm could adjust the η automatically according to the features of the network. When a big learning rate can make the network stably study, we will make its error continues to fall. And then we increase the learning rate, the study will keep in greater learning rate; Once the learning rate is too large that cannot continue to reduce error, we will reduce the learning rate to make the learning process more stable. In the figure 2, we demonstrate the basic concepts of the adaptive learning algorithm for neural network.

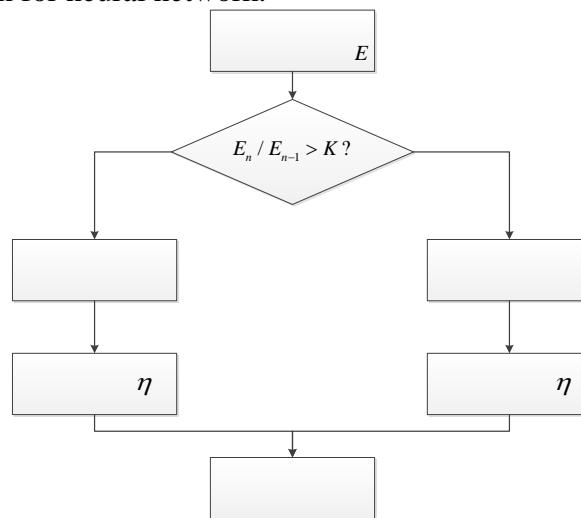


Fig. 2 The Flowchart of the Adaptive Learning Algorithm for the Neural Network

The Fuzzy BP Neural Network. In general cases, we are always considering the complete sets of data to train the neural network and execute the prediction functionality. However, as for the stock price, the data sets are generally incomplete for the reason that there are always existing noisy outliers. To eliminate the influences caused by the noisy outliers and restore the reality for trend prediction, we need to adopt fuzzy rough theory to serve as the data pre-processing step [6-7].

Fuzzy set is described by the membership function and the key to process of fuzzy sets is to accurately determine the membership function. We firstly define the fuzzy set as equation 11.

$$\mu_A : U \rightarrow [0,1], x \rightarrow \mu_A(x) \quad (11)$$

We put the price information area as a defined fuzzy set, and on each trading day stock value has its own membership. For the membership value, the more information shows in the defined area, the higher possibility it is and vice versa. Such information collection area of the definition holds the characteristic of ambiguity which conforms to the reality to facilitate further process of useful data. Uncertainty and fuzziness of fuzzy theory is based on the concept of boundary that is a fuzzy concept with fuzzy boundary. Each uncertain concept consists of a pair, which is called upper approximation and lower approximation precise. The definition of this pair is the following: In a set of a given knowledge base, for each subset and equivalent relation, the collection can be divided according to basic set description. The formula 12~14 defines the corresponding classification rules.

$$R_-(X) = \bigcup \{Y \in U / R : Y \subseteq X\} \quad (12)$$

$$R^-(X) = \bigcup \{Y \in U / R : Y \cap X \neq \emptyset\} \quad (13)$$

$$BN_R(X) = R^-(X) - R_-(X) \quad (14)$$

$R_-(X)$ and $R^-(X)$ represent the upper approximation and the lower approximation. In a decision system, there is a certain degree of dependence or association between each condition attributes and the reduction, which means without losing information, it can be most simply described as the decision system of decision attribute dependency or correlation degree of condition attributes set. With the prior filtering operation with the fuzzy rough theory, we could clean the price data that may influence the prediction on the trend and achieve better prediction accuracy.

Simulation and Verification

In order to assess the effectiveness and overall efficiency of our stock price prediction algorithm, in this section, we simulate our methodology compared with the method1 [3], method2 [4] and method3 [5]. We take the stock price of IFLYTEK CO., LTD. from December 9th, 2011 to the later 14 trading days. In the table one, we demonstrate the real values and the predicted values by different algorithms of the stock price. The blue tag denotes the predicted data that is closest to the real value and the green tag represents the second optimal prediction compared with other three approaches. To show the visualized result, we draw the predicted values on the figure 3 and the box plots on the figure 4. From the result, we could conclude that in most of cases, our algorithm holds better prediction performance especially on the reflection of price developmental trend.

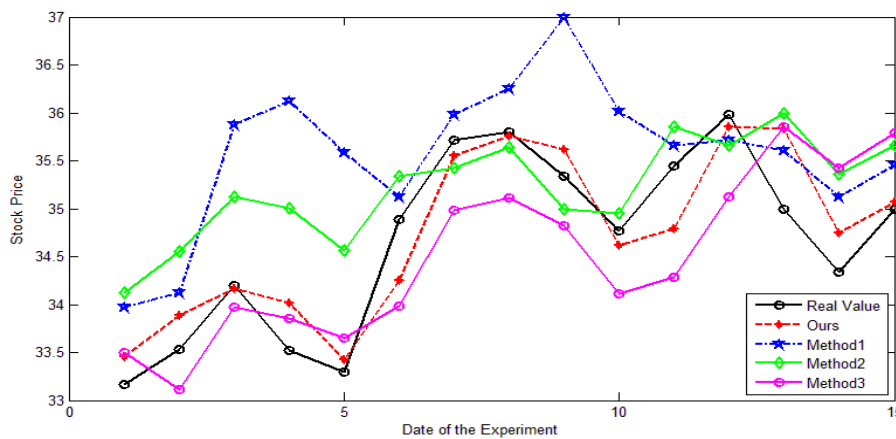


Fig. 3The Comparison Experiment on the Stock Price Prediction

Table 1The Numerical Statistical Data of the Stock Price Prediction Result

Date	Real Value	Ours	Method1	Method2	Method3
1	33.17	33.46	33.97	34.12	33.50
2	33.53	33.89	34.12	34.55	33.11
3	34.20	34.17	35.88	35.12	33.97
4	33.52	34.02	36.12	35.01	33.85
5	33.30	33.42	35.59	34.56	33.65
6	34.89	34.25	35.12	35.34	33.98
7	35.71	35.55	35.98	35.42	34.98
8	35.80	35.76	36.25	35.64	35.11
9	35.34	35.62	36.99	35.00	34.82
10	34.77	34.62	36.02	34.95	34.11
11	35.45	34.79	35.66	35.85	34.29
12	35.98	35.86	35.72	35.66	35.12
13	35.00	35.83	35.61	35.99	35.86
14	34.34	34.75	35.12	35.36	35.42
15	35.00	35.07	35.46	35.66	35.79

Conclusion

In this paper, we propose the novel data mining algorithm based on the BP neural network and its applications on stock price prediction. The stock market is a highly complicated nonlinear dynamic

system and its change not only holds the intrinsic regularity, but also by the market together with the influence of the economic and non-economic factors. Neural network possesses the characteristics of self-organization, adaption and can automatically extracted from historical data and knowledge from related financial activities therefore it is very suitable for solving the problem of stock prediction in the field of nonlinear time series. Our methodology starts from analyzing the primary theoretical basis of the BP neural network to serve as the foundation of the study, and later, we revise the traditional network training and organizing pattern by optimizing the error, incentive function and incorporate the concept of self-adaptive learning algorithm. To deal with the condition of redundant data noise, we introduce the fuzzy rough set theory into the data pre-processing step to eliminate the influence of noisy data for improving the prediction accuracy. The experimental analysis verifies the effectiveness of the proposed algorithm compared with the other state-of-the-art approaches. In the later, we plan to take more data sets into consideration to test the robustness and feasibility of our algorithm.

References

- [1] Wang, Ping, et al. "Stratified analysis of the magnetic Barkhausen noise signal based on wavelet decomposition and back propagation neural network." *Sensors and Actuators A: Physical* 201 (2013).
- [2] Cireşan, Dan, et al. "Multi-column deep neural network for traffic sign classification." *Neural Networks* 32 (2012): 333-338.
- [3] Liu, Chih-Feng, Chi-Yuan Yeh, and Shie-Jue Lee. "Application of type-2 neuro-fuzzy modeling in stock price prediction." *Applied Soft Computing* 12.4 (2012): 1348-1358.
- [4] Zarandi, M. H., Esmail Hadavandi, and I. B. Turksen. "A hybrid fuzzy intelligent agent - based system for stock price prediction." *International Journal of Intelligent Systems* 27.11 (2012): 947-969.
- [5] Hsu, Chih-Ming. "A hybrid procedure for stock price prediction by integrating self-organizing map and genetic programming." *Expert Systems with Applications* 38.11 (2011): 14026-14036.
- [6] Guo, Zhen-hai, et al. "A case study on a hybrid wind speed forecasting method using BP neural network." *Knowledge-based systems* 24.7 (2011): 1048-1056.
- [7] Ding, Shifei, Chunyang Su, and Junzhao Yu. "An optimizing BP neural network algorithm based on genetic algorithm." *Artificial Intelligence Review* 36.2 (2011): 153-162.