Research on Investment Portfolio Model Based on Neural Network and Genetic Algorithm

Wei Zhou
Zaozhuang University
Yuanjun Zhao
Shanghai Lixin University of Accounting and Finance
Weiwei Chen
Hunan University of Technology and Business
Yanghui Liu
East China Jiao Tong University
Rongjun Yang (sdzz1667100164@163.com)
Zaozhuang University
Zheng Liu
Shanghai University of Engineering Science

Research

Keywords: neural network, genetic algorithm, stock investment

Posted Date: August 13th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-57296/v1

Version of Record: A version of this preprint was published on November 2nd, 2020. See the published version at https://doi.org/10.1186/s13638-020-01850-x.

Abstract

With the maturity of neural network theory, it provides new ideas and methods for the prediction and analysis of stock market investment. The purpose of this paper is to improve the accuracy of stock market investment prediction, we build neural network model and genetic algorithm model, study the law of stock market operation, and improve the effectiveness of neural network and genetic algorithm. Through the empirical research, it is found that the neural network model can make up for the shortcomings of the traditional algorithm through the optimization of genetic algorithm.

Introduction

In the western countries with developed financial markets, the stock market has been running for more than 300 years, which to a certain extent reflects the development and trend of social economy in a specific period of time, and has a profound impact on economic development. In the early 1990s, China established its own securities market, and the listed companies in Shanghai and Shenzhen are increasing day by day. After nearly 30 years of accumulation, the market scale has reached a high level [1]. With the rapid development of the financial market, the enthusiasm of the public for financial investment is constantly improving, and the knowledge of financial investment is constantly increasing. However, the stock market investment is a high-risk and high-yield investment activity. The level of return and the degree of risk have a significant positive correlation. In other words, investors need to take a greater risk to get a higher return. In this context, a more reasonable stock market forecasting method can effectively reduce risk and increase returns. Therefore, it is of great significance to strengthen the research of stock market forecasting methods. In reality, the number of investors in the stock market is increasing, and investors also hope to obtain an effective analysis method to maximize the ratio of return to risk [2]. In recent years, with the general investors more and more understanding of the basic rules of stock market investment, as well as the increasing number of stock market investment researchers, stock market forecasting methods are also increasing [3]. It needs to be emphasized that at present, some stock market prediction methods emphasize ideal state, but due to the complex internal and external environment, various uncertain factors always impact the stock market investment market, which to a certain extent improves the prediction difficulty of the stock market, and greatly reduces the prediction effectiveness of the stock market investment [4]. Therefore, even the prediction methods with high popularity often fail in market prediction [5]. At present, the rapid development of science and technology provides a new way for stock market investment prediction and analysis, especially the growing maturity of neural network theory, which has been well applied in many aspects, such as signal processing, pattern recognition and so on. By analyzing the theory of neural network, it can be found that neural network has great advantages in self-adaptive and self-learning, has the characteristics of nonlinear approximation ability, and has a high degree of agreement with the stock market prediction [6]. Therefore, it is a good attempt to apply neural network to stock market prediction [7].

Related Work

2.1 Basic concept of genetic algorithm

In the process of social practice, people need to find the optimal solution in the complex system in order to solve the problem efficiently. However, because the solution space is relatively large, the correlation between the parameters and the target value is difficult to determine, and there are relatively many factors to be considered, so how to deal with the optimization problem must be highly valued. In many cases, people determine the approximate optimal solution by comparing and analyzing the random effective solution [8]. The essence of this method is to randomly extract the parameters of the domain of definition to obtain the optimal solution. This method is simple and easy, but it is only suitable for the field with small search space, but for the field with large search space, it can't solve the problem simply by exhaustive method More advanced optimization techniques are needed to solve the problem [9]. In contrast, the genetic algorithm with ‘survival of the fittest’ as the core has great advantages. By introducing competition mechanism into the algorithm, the search efficiency can be improved. The basic process of genetic algorithm is to determine a group of initial solutions in a random way, and then conduct individual search to obtain an independent solution, which is defined as a ‘chromosome’. Through the ‘fitness value’ index, the adaptability of chromosomes in the population can be effectively evaluated, and then whether to select them to enter the next stage can be judged [10]. According to the principle of survival of the fittest, on the basis of continuous crossing, selection and variation, the evolution selection of chromosomes forms a chromosome group with higher adaptability. After reaching a certain number of iterations, the chromosome convergence is completed and the optimal solution of genetic algorithm is obtained [11]. By analyzing the process, we can find that the whole process of genetic algorithm is essentially similar to the genetic principle in biological sense [12].

2.2 Running process of genetic algorithm

At the operational level, genetic algorithm is not complex. According to the above discussion, it is essentially an iterative process, that is, it starts from the initial group of individuals, and obtains the approximate optimal solution through continuous cross selection and mutation operation [13]. Overall, the basic elements of genetic algorithm are as follows:

(1) Chromosome coding. In the construction of genetic algorithm model, the first step is to determine the coding method, which is also the key and core problem of genetic algorithm. In the early stage of the development of genetic algorithm, binary coding is the most widely used coding method, which is the first choice of algorithm designers. Compared with other types of chromosome coding methods, binary long coding greatly reduces the difficulty of coding and decoding, and is conducive to the completion of cross selection, mutation genetic and other operations [14].

(2) Individual fitness evaluation. First of all, through comparative analysis, to understand the fitness of different individuals, so as to determine the individuals to be selected into the next generation, and gradually complete the construction of the next generation group; secondly, the algorithm designer needs to determine the fitness function of this paper according to the genetic selection needs, and use this as a tool to complete the local guidance search; finally, according to the fitness function determined previously

Number, analyze the differences of different individuals, and evaluate their fitness [15]. Generally speaking, the nature of the problem to be solved is different, and the field is different. The criteria to be referred to in the selection of fitness function should also reflect differences.

(3) Select the operator. In the process of building genetic algorithm model, operator selection is the key link, which is closely related to the effectiveness of genetic algorithm [16]. The designers of genetic algorithm model need to select the operators that meet the requirements scientifically and reasonably, so as to improve the performance of genetic algorithm to the greatest extent. Fundamentally speaking, probability rule is the core attribute of operators, aiming to select individuals with strong adaptability, and it is an effective tool to determine the next generation of population.

(4) Crossover operator. In the construction of genetic algorithm model, we need to choose a reasonable cross algorithm in order to produce new individuals different from the existing ones [17]. In many kinds of crossover algorithms, the single point crossover operator has the highest application rate and is also one of the basic operators.

(5) Mutation operator. In the process of building genetic algorithm model, in order to reflect the complexity and uncertainty of the environment, the principle of gene mutation in biology is used for reference [18]. Therefore, mutation operators will appear, but the probability of occurrence is relatively low, it can greatly improve the comprehensiveness of the genetic algorithm model and avoid falling into the problem of local search. In conclusion, it is necessary to introduce mutation operator into genetic algorithm model.

(6) Select the control parameters. Because genetic algorithm is a new science, and it has probability attribute, it emphasizes the experience of designers in parameter selection, or through some experiments to determine the parameters [19]. Among them, whether the genetic algorithm can be implemented efficiently depends on the population size to a large extent. If the population size is insufficient, the genetic algorithm lacks regularity, and the problem of local solution is relatively prominent, which has a negative impact on the final performance; on the contrary, if the population size is large, it will reduce the convergence speed, and have a greater negative impact on the efficiency of the algorithm. In the initial stage of genetic algorithm, individuals have significant randomness, and the applicable mutation rate should be large, so as to improve the diversity of population and create conditions for global search [20]. In the process of evolution, in order to maintain the stability of some high-quality characteristics, it is necessary to reduce the variation rate appropriately, but in the final stage of evolution, in order to avoid individual convergence, it is necessary to improve the variation rate again to ensure that individual diversity meets the requirements. Generally speaking, the crossing rate is between 0.2–0.95.

The operation flow of the basic genetic algorithm is as follows:

(1) In initialization, the evolutionary algebra is set to the maximum, and the initial population is determined according to the acquired random individuals;

(2) Individual evaluation, on the basis of determining groups, calculates the fitness of different individuals;

(3) According to the group characteristics, the optimal operator is determined;

(4) Cross operation is introduced in the operation process;

(5) In the application process of genetic algorithm, mutation operator is introduced into population to select individuals with different attributes from existing individuals to enter the next generation of population;

By cycling the above steps, if the condition of $t \leq T$ is reached, then continue to the second step cycle, the optimal solution of genetic algorithm model selects the individual with the largest fitness in the population, so as to get the best solution to the problem and achieve the final calculation goal.

2.3 Convergence theory of genetic algorithm

Markov chain analysis is an important part of genetic algorithm, and its core is the convergence theory [21]. It can be found that the convergence of traditional genetic algorithm is generally based on the Markov chain limit theory. In the practice of solving problems, the ultimate goal of genetic algorithm is to determine the global optimal solution. The essence of the whole process is random search, which has great uncertainty [22]. The operation process of genetic algorithm is continuously optimized under the expected value of the optimal solution, and it is regarded as the initial sequence. Through the convergence theory of genetic algorithm, its convergence can be effectively verified. Not only that, in order to achieve good convergence of genetic algorithm, we must focus on two parameters, one of which is the possibility of breaking away from the satisfactory solution set on the premise of determining the satisfactory solution; the other is the possibility of still not obtaining the satisfactory solution on the premise of not obtaining the satisfactory solution, and the convergence of genetic algorithm is formed on the basis of the above two parameters matching General theory [23]. The convergence research based on the two parameters is pure probability research, which is intuitive and simple in the convergence verification of genetic algorithm.

Methods

The purpose of this paper is to improve the accuracy of the stock market investment prediction. By combining the neural network model and the genetic algorithm model, we can predict the operation law of the stock market. This paper relies on the existing theoretical research results to optimize the real number coding scheme and improve the effectiveness of neural network algorithm and genetic algorithm. In this paper, the real number coding method is adopted, the sample segmentation is optimized, the training is strengthened, the training speed and convergence speed of neural network are improved, the local minimum value is obtained, so as to avoid falling into, the three-layer neural network is constructed to determine the global optimal solution, so as to effectively solve the problem.

3.1 Genetic algorithm to optimize the learning rules of neural network

In the training process, neural network learning rules need to be set in advance. However, whether the learning rules are reasonable or not is uncertain. Therefore, it is necessary to design and optimize neural network learning rules with the help of genetic algorithm, so as to improve the ability of neural network algorithm to solve complex problems and the adaptability of the algorithm to uncertain environment. Research results show how to design coding by learning rules is the core problem in the evolution process, so far there are no cases with application value. Therefore, the study of learning rules is only the initial stage, and its process is as follows:

(1) The effective coding method is determined, the learning rules are coded, and the matching between individual and single learning rules is realized [24];

(2) To construct a training set, the elements of the training set are determined firstly, and then the corresponding learning training is carried out according to the matching learning rules [25];

(3) Calculate the fitness of all learning rules;

(4) Select and determine learning rules that meet the requirements;

(5) Cross selection, individual variation processing, analysis of individual attributes, to determine the next generation of population;

(6) Repeat the above steps until the goal of evolution is achieved.

In this paper, after the optimization of genetic algorithm, the connection weight of neural network is improved. By solving the existing problems of neural network, the generalization function of neural network is enhanced. On this basis, the learning model of neural network is constructed, and the global optimal solution is obtained to achieve the ability of solving specific problems.

3.2 Basic idea of genetic algorithm optimization

Through the analysis of the operation logic of genetic algorithm and neural network algorithm, it can be found that the weight modification principle based on the progressive reduction of error is inevitable to have the problem of the minimum limit point. Because genetic algorithm is not suitable for local search, it is not conducive to the accuracy of genetic algorithm. Based on the past experience, this paper selects the mechanism of genetic algorithm to initialize the network weight. The core idea is to initialize the network weight within a specific range. According to the size of the population, the network weight can be divided into 100, 200, 300 or more. In addition, according to the evolution principle, to achieve the minimum network error as the goal, through continuous iterations, to obtain final weight that meets the requirements is the initial weight of algorithm training. Through the analysis of genetic algorithm and neural network algorithm, we can find that they have significant complementary advantages, which can improve the problem-solving ability. The application of genetic algorithm is inseparable from the matching coding method. The essence of weight learning is parameter optimization, which has continuity and complexity. For example, binary coding will lead to the length of the coding string does not meet the requirements, while real coding is relatively effective. Therefore, this paper optimizes the traditional algorithm, selects the real coding method, correlates the weights and thresholds, and transforms them into chromosomes in genetic space.

3.3 GA-BP algorithm design

Parameters play a decisive role in the performance of the algorithm model. In this paper, finite length coding is chosen. After the design of the algorithm coding scheme is completed, the parameter coding is transformed into the genetic algorithm coding, and then the function used to accurately evaluate the algorithm performance is determined, and the global search is completed in the parameter space. In this way, not only the space can be expanded, but also the target of regional search can be realized, and a balance state can be achieved between the two. In the initial stage of genetic search, due to the uncertainty brought by cross variation, the search scope has been expanded to a certain extent. After obtaining the high fitness solution, the crossover operation completes the search near the above solution. Therefore, through the genetic operation, we can determine the best combination of parameters to meet the requirements of practical application and solve the problem. In terms of algorithm implementation, the specific process is as follows:

(1) Randomly forming n codes and forming initial set s;

(2) Complete the coding in sequence, decode the coding, determine a parameter combination P reflecting BP model, determine BP, evaluate the BP and obtain its corresponding fitness value;

(3) According to the appropriate value determined in the previous step, determine n individuals, and enter the next generation to obtain the next generation group. In this step, some individuals may need to be selected multiple times;

(4) According to the probability P and the fitness value of different codes, the parent generation is determined, then cross inheritance is carried out, and the next population is entered after random pairing.

(5) According to the probability P and fitness, select the parent population that meets the requirements, insert new individuals through mutation inheritance, and achieve the goal of population iteration.

(6) By repeating the above steps repeatedly, the search target can be achieved on the premise of meeting the standard requirements.

$$w < x - 1 \quad (1)$$

$$w < \sqrt{(x+y)} + a \quad (2)$$

$$w = log_2 x \quad (3)$$

x is the number of input layer nodes, w is the number of hidden layer nodes, a is a constant between 0 and 10.

Results And Discussion

4.1 Neural network toolbox

As a highly complex and comprehensive algorithm model, neural network model has relatively high requirements for toolbox. Through the application of neural network toolbox, the goal of activation function can be realized. At the same time, through the algorithm training, the network designer can complete the specific subroutine, and based on this, promote the learning training, complete the corresponding call requirements to the greatest extent, and improve the effectiveness of neural network learning. In the process of building algorithm model, different types of algorithms are integrated into neural network toolbox, so as to improve the convenience of algorithm design.

4.2 Genetic algorithm toolbox

All kinds of algorithms of genetic algorithm ultimately need to act on chromosomes. Chromosomes are essentially vector types, which can be reduced to specific matrices, and matrix operations form operators. It can be seen that the basic data unit has the infinite feature of matrix dimension. Based on this recognition, users can ignore the low-level problems related to matrix algorithm, so as to improve the operation efficiency on the basis of programming. In the application process of genetic algorithm, the toolbox can provide the necessary algorithm with the characteristics of scalability and standardization. Through its matrix computing ability, it can improve the efficiency of genetic algorithm, reduce the difficulty of chromosome programming, and improve the difficulty of solving problems.

4.3 Sample data

4.3.1 Basic requirements for sample data

The number of samples is closely related to the accuracy of genetic algorithm, it is closely related to the complexity of mapping relationship, and also to a certain extent depends on data noise. With the increase of the complexity of mapping relationship, the number of training samples is also required to be higher, and with the increase of noise, the number of samples is also increased synchronously. In the aspect of sample selection, we need to adhere to the following three principles: (1) meet the requirements of sample quantity; (2) meet the requirements of sample accuracy; (3) meet the requirements of sample representativeness; (4) meet the requirements of sample distribution.

4.3.2 Acquisition of sample data

Stock price return is a common index in the quantitative analysis of stock market investment. However, if the stock market is positioned as a nonlinear dynamic system, the return is not the optimal price alternative transformation, and the factor of forecasting price should be fully considered. Not only that, with the increase of the number of training samples, the amount of calculation will be increased to a large extent, and the convergence speed in the training process will decline, which requires a longer convergence time. If the number of samples does not meet the requirements, the network will not be able to fit the corresponding stock index curve. In this paper, 'gzmt' is chosen as the representative of empirical analysis, and its stock number is 600519. In terms of data selection, this paper adheres to the basic principles of representativeness, continuity, uniform distribution and accuracy to improve the adaptability of this algorithm.

If we want to reduce the prediction error and improve the prediction accuracy, we need to choose a reasonable number of samples to meet the training requirements of neural network. On the basis of fully considering the prediction characteristics of neural network, the training samples are optimized to improve the convenience of detection.

4.3.3 Preprocessing of sample data

Before the formal learning, the effectiveness of the network largely depends on data processing. Data processing will affect its accuracy and speed. Generally speaking, the network training cannot directly apply the samples obtained, it needs to complete the necessary processing before it can be put into application. In other words, the acquired data samples usually need to be normalized before they can be applied to network training.

4.4 Closing price network model

4.4.1 Network model and sample design

In this paper, 277 historical closing prices are selected as data samples to build the basic data model of stock research. In order to ensure that the data samples meet the use requirements, the data samples must be normalized before they can be put into use. The input node of the network is $p = 5$, and the output node is $t = 1$, that is, the closing price data of five trading days before $t + 1$ is used for prediction. At present, the hidden layer of neural network cannot be determined. In this paper, on the basis of ensuring that the error meets the requirements, in order to reduce the calculation difficulty to the greatest extent, the best number of hidden layers is verified through experiments, so as to obtain the number of hidden layers in a reasonable range. The numerical results show that only through the L-M training method can a high-speed three-layer network be established. At present, 5-12-1 network structure is the most widely used, which uses 5 input nodes, 1 output node and 12 hidden layer nodes. Because BP network security has the characteristics of network generalization, 159 training samples and 122 test samples can be selected according to the data samples. Figure 1 is schematic diagram of stock market forecast. Figure 1 is schematic diagram of stock market forecast.

4.4.2 Simulation experiment and result analysis

The main goal of this paper is to optimize BP network and build an efficient and accurate operation model based on genetic network. In the process of forecasting the closing price of GZMT, Guizhou Province, after the introduction of the optimized network model, necessary testing and training are needed. Figure 2 is test chart of BP network model.

In this paper, based on the operation tools provided by neural network toolbox, combined with the algorithm discussed above, and through programming, the calculation of network closing price is completed. After completing the training, we need to rely on the test set to carry out the necessary tests, and then we can determine the samples that integrate with the stock index curve. Figure 3 is neural network fitting curve.

Through sample training, the number of iterations is determined when the target error meets the requirements. According to the network model constructed in this paper, the error curve can be calculated. Figure 4 is error curve.

By training 150 groups of sample data, the error can be reduced to the greatest extent, and most of the errors are reduced to about 0. Therefore, the learning training in this paper has basically achieved the expected goal. We found that the neural network model has high accuracy in prediction. Figure 5 is fitting curve of test sample.

After three times of training, the SSE of sample data can be determined as 11.5857e-004, which achieves a relatively good fitting effect, and the error analysis also shows that it achieves a good effect. Therefore, we choose L-M back propagation algorithm to carry out the learning and training, which not only has a good approximation effect, but also can achieve a high speed of operation. The most important thing is to avoid the local minimum problem, which needs to occupy a relatively large memory space. It should be emphasized that in the process of learning and training, it is necessary to track and observe the learning rate and target error, so as to reduce the impact on the convergence speed as much as possible. To sum up, this paper can effectively improve the efficiency of stock market prediction through artificial neural network algorithm, which has high application value.

4.5 Prediction and analysis of stock market based on GA-BP network model

4.5.1 GA-BP network model establishment and prediction realization

The specific process of GA-BP network model is as follows:

(1) Determine the initial function;

(2) Complete the fitness training;

(3) Obtain the initial population;

(4) Call the genetic function;

(5) The weight and threshold of neural network are determined;

(6) To construct the targeted network training;

(7) The network is determined by weight and threshold;

(8) Tracking and observing the network performance;

(9) Solve the prediction problem according to the network.

The neural network designed in this paper selects three-layer structure, inputs the closing price of T-4 day, T-3 day, T-2 day, T-1 day and t day, outputs the closing price of T + 1 day, and the learning rate is 0.5. After the following conditions are met, the learning is finished:

(1) Goal ≤ 1e-5 or epochs ≤ 1000;

(2) Select 100 iterations, and use GA-BP network training as shown in the figure below. Figure 6 is GA-BP network training. Figure 7 is training iteration diagram.

By comparing the output value with the actual value, it can be found that the neural network test sample in this paper achieves good prediction effect. Figure 8 is GA-BP fitting curve of test sample.

4.5.2 Test of network prediction results

Compared with other algorithms, it is difficult to evaluate through the test indicators in econometric technology. Therefore, this paper summarizes the most widely used centralized evaluation indicators based on the research results at home and abroad

Mean absolute error:

$$MAE = \frac{\sum |e_j|}{N} \quad (4)$$

Squared difference:

$$SSE = \sum (e_j)^2 \quad (5)$$

Mean square deviation:

$$MSE = \frac{\sum (e_j)^2}{N} \quad (6)$$

4.5.3 Comparison between GA-BP network model and improved model

According to the above table, it can be found that compared with the general neural network, the real coded genetic algorithm optimization model not only has high accuracy and good effect, but also can prevent the occurrence of local minima, and can improve the convergence speed, which has high practical value. Table 1 is comparison results of three algorithms.

Table1 Comparison results of three algorithms

test index	MAE	SSE	MSE
General BP algorithm	0.0385	0.5134	0.9942
Improved BP algorithm	0.0012	7.0750e-004	5.7992e-006
GA-BP algorithm	9.8222e-004	3.9378e-004	3.2277e-006

4.5.4 Evaluation of the improved network model of genetic algorithm

By analyzing the above results, we can find that genetic algorithm has the following characteristics;

(1) Through the application of real number coding, we can get the algorithm with mutation operator as the core, and can omit the cross selection. Compared with the standard genetic algorithm, mutation operators have higher performance in global or local aspects.

(2) In order to maintain the competition intensity of chromosomes, it is necessary to arrange the chromosome fitness first in the process of operator selection, so as to ensure that the chromosome with the smallest error is output, and to determine the selection probability to complete the selection of operators by using the rotating roulette method.

(3) GA is separated from the parameter and is only represented as a digital string of parameters. In the aspect of solution space processing, it is not centralized processing, but group points processing, so as to prevent local minimum.

(4) Determine the direction of evolution, and guide the cross selection and variation reasonably. All the crossed or mutated chromosomes can go through multiple crossing and mutating operations and retain the optimal solution. By using the guidance factor, the convergence rate can be improved, and the optimization process can be completed only by the value of the objective function. The optimization rules

depend on the probability, so using genetic algorithm to optimize multi extremum has some advantages, which can make up for the limitations of common algorithm.

(5) In the process of optimization, GA mode can complete the optimization process only through the objective function value. According to the above discussion, it can be found that through the application of genetic algorithm, it is beneficial to optimize multiple extremum, has greater advantages compared with other algorithms, can predict more accurately, and meet the expected requirements of improving the accuracy of stock market prediction.

Conclusion

With the development of intelligent technology, neural network has become a frontier interdisciplinary research field, which is conducive to improving the design level and comprehensive performance of neural network system. Based on the existing research results, this paper improves the application efficiency of the algorithm by optimizing the specific links of the neural network. It systematically explains the technical terms and relevant indicators of the stock market investment, explains the common prediction methods, and systematically discusses the research hotspot and difficulty. This paper systematically discusses the related theory and application process, and analyzes the similarities and differences between genetic algorithm and neural network technology, which can avoid the problem of limited minimum. Choose the stock market of our country as the research sample, and then get the feasibility of short-term forecasting stock market, which lays a theoretical foundation for this study. Scientific and reasonable selection of input parameters can not only reflect the scale and quality of information in the stock market, but also avoid the problem of learning and training difficulty or even non-convergence due to information overlapping. As the stock market changes with time, the network training samples also need to be adjusted accordingly, otherwise it will be difficult to guarantee the accuracy of prediction. Therefore, the preprocessing of the original data is an essential part. Through the empirical study, it is found that the neural network model can make up for the existing problems of the common algorithm through the optimization of genetic algorithm.

Declarations

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by National Social Science Foundation of China [grant number: 18CGL015].

Authors’ contributions

Wei Zhou designed research, performed research, analyzed data, and wrote the paper. All authors analyzed the data and were involved in writing the manuscript.

Abbreviations

Abbreviations is not used.

References

1. S. Asadi, E. Hadavandi, F. Mehmanpazir, M.M. Nakhostin, Hybridization of evolutionary Levenberg–Marquardt neural networks and data pre-processing for stock market prediction. Knowl.-Based Syst. 35, 245–258 (2012)
2. P.C. Chang, C.H. Liu, J.L. Lin, C.Y. Fan, C.S. Ng, A neural network with a case based dynamic window for stock trading prediction. Expert. Syst. Appl. 36(3), 6889–6898 (2009)
3. M. Jasemi, A.M. Kimiagari, A. Memariani, A modern neural network model to do stock market timing on the basis of the ancient investment technique of Japanese Candlestick. Expert. Syst. Appl. 38(4), 3884–3890 (2011)
4. M. Shaverdi, S. Fallahi, V. Bashiri, Prediction of stock price of Iranian petrochemical industry using GMDH-type neural network and genetic algorithm. Applied mathematical sciences 6(7), 319–332 (2012)
5. R.J. Kuo, C.H. Chen, Y.C. Hwang (2001). An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network. Fuzzy sets and systems, 118(1), 21–45
6. A.H. Moghaddam, M.H. Moghaddam, M. Esfandyari, Stock market index prediction using artificial neural network. Journal of Economics, Finance and Administrative Science 21(41), 89–93 (2016)
7. R.L. Naik, D. Ramesh, B. Manjula, A. Govardhan, Prediction of stock market index using genetic algorithm. Computer Engineering and Intelligent Systems 3(7), 162–171 (2012)
8. Z.H. Khan, T.S. Alin, M.A. Hussain, Price prediction of share market using artificial neural network (ANN). International Journal of Computer Applications 22(2), 42–47 (2011)
9. Y. Chen, B. Yang, A. Abraham, Flexible neural trees ensemble for stock index modeling. Neurocomputing 70(4–6), 697–703 (2007)
10. J.J. Wang, J.Z. Wang, Z.G. Zhang, S.P. Guo, Stock index forecasting based on a hybrid model. Omega 40(6), 758–766 (2012)
11. J. Zahedi, M.M. Rounaghi, *Application of artificial neural network models and principal component analysis method in predicting stock prices on Tehran Stock Exchange*, 438 (Statistical Mechanics

and its Applications, Physica A, 2015), pp. 178–187

12. C.F. Tsai, Y.C. Hsiao, Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. Decis. Support Syst. 50(1), 258–269 (2010)
13. W. Sun, Y. Zhao, L. Sun (2019). Big Data Analytics for Venture Capital Application: Towards Innovation Performance Improvement. International Journal of Information Management[J]. 2019, 50(1):557–565
14. L. Sun, Y. Zhao, W. Sun, Study on supply chain strategy based on cost income model and multi-access edge computing under the background of the Internet of Things. Neural Comput. Appl. (2019). https://doi.org/10.1007/s00521-019-04125-9
15. F. Wang, L. Ding, H. Yu et al. (2019). Big data analytics on enterprise credit risk evaluation of e-Business platform. Information Systems and e-Business Management. https://doi.org/10.1007/s10257-019-00414-x
16. R. Yang, L. Yu, Y. Zhao et al., Big data analytics for financial Market volatility forecast based on support vector machine. International Journal of Information Management[J] 50(1), 452–462 (2019)
17. Z. Liu et al., Research on Staged Pricing Model and Simulation of Intelligent Urban Transportation. IEEE Access[J] 7(11), 141404–141413 (2019)
18. M. Xie, H. Li, Y. Zhao, Blockchain financial investment based on deep learning network algorithm. Journal of Computational and Applied Mathematics[J]. 2020, 372(7) https://doi.org/10.1016/j.cam.2020.112723
19. S. Cai, X. Wang, Y. Zhao, Revenue Model of Supply Chain by Internet of Things Technology. IEEE ACCESS 7(1), 4091–4100 (2019)
20. V. Cho, MISMIS–A comprehensive decision support system for stock market investment. Knowl.-Based Syst. 23(6), 626–633 (2010)
21. F.C. Yuan, C.H. Lee, Using least square support vector regression with genetic algorithm to forecast beta systematic risk. Journal of computational science 11, 26–33 (2015)
22. Z. Chen, H. Cai, Y. Zhang, C. Wu, M. Mu, Z. Li, M. A. Sotelo. A novel sparse representation model for pedestrian abnormal trajectory understanding. Expert. Syst. Appl. 138, 112753 (2019)
23. B. Wu, T.T. Cheng, T.L. Yip, Y. Wang, Fuzzy logic based dynamic decision-making system for intelligent navigation strategy within inland traffic separation schemes. Ocean Eng. 197, 106909 (2020)
24. Z. Huang, X. Xu, J. Ni, H. Zhu, C. Wang, Multimodal Representation Learning for Recommendation in Internet of Things. IEEE Internet of Things Journal 6(6), 10675–10685 (2019)

Figures

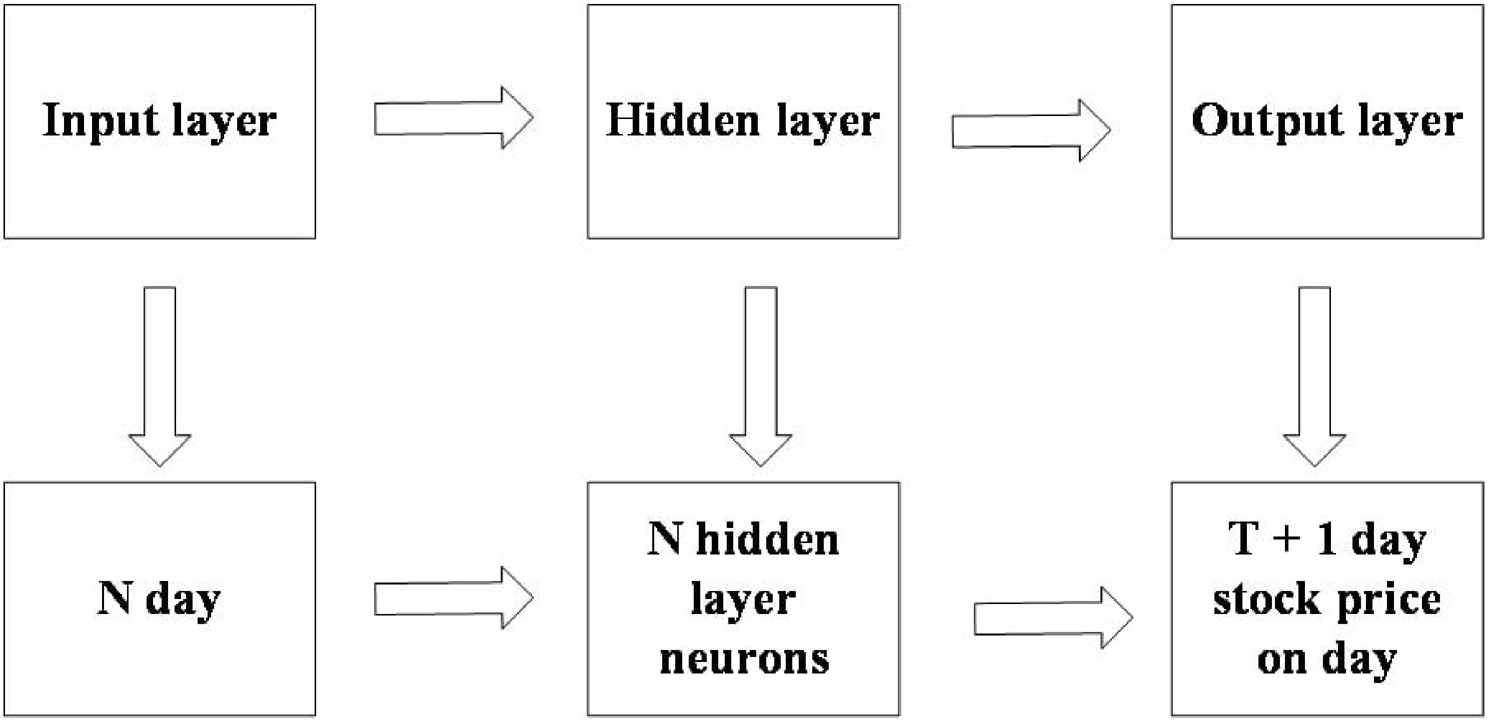

Figure 1

Schematic diagram of stock market forecast

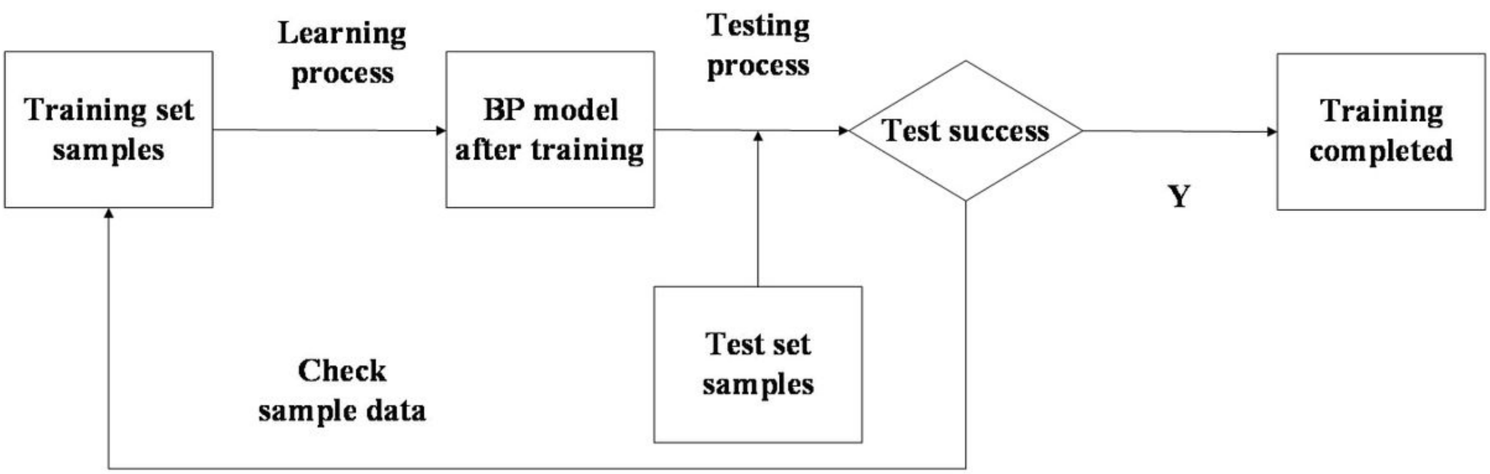

Figure 2

Test chart of BP network model

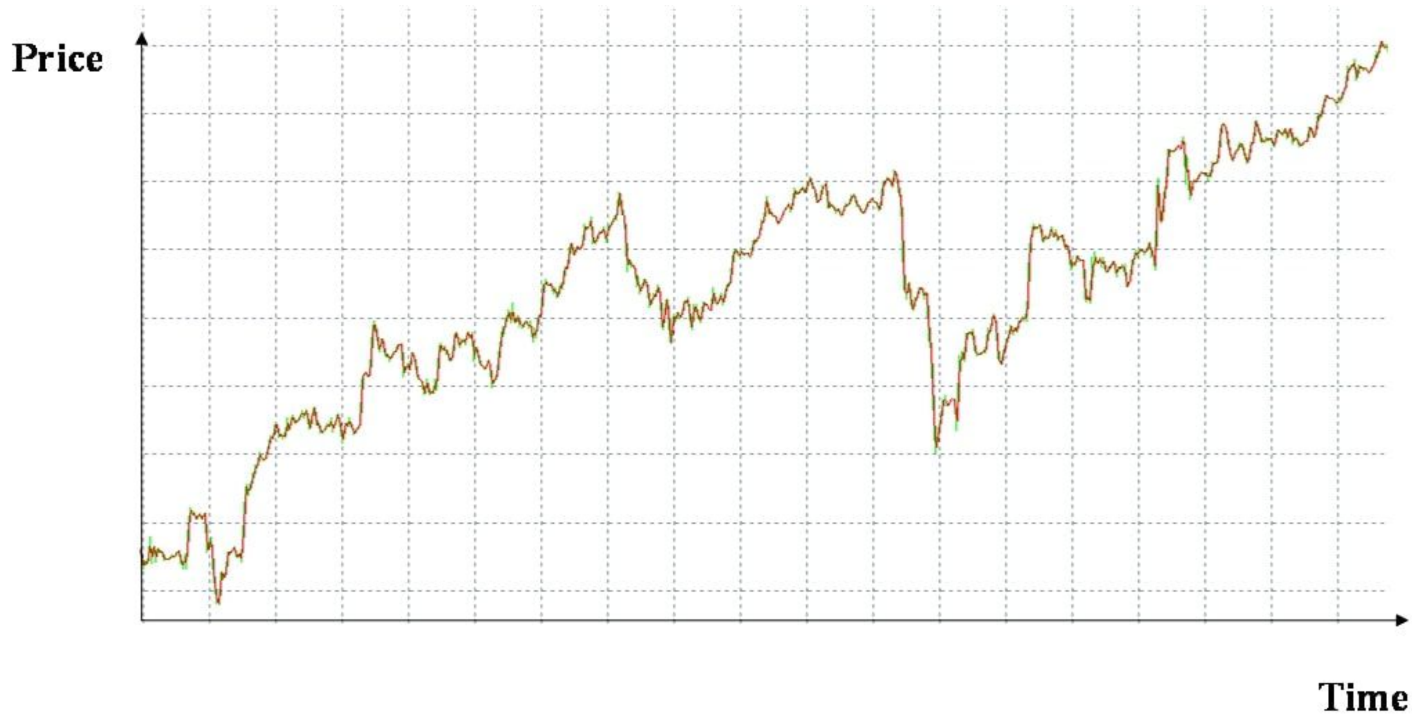

Figure 3

Neural network fitting curve

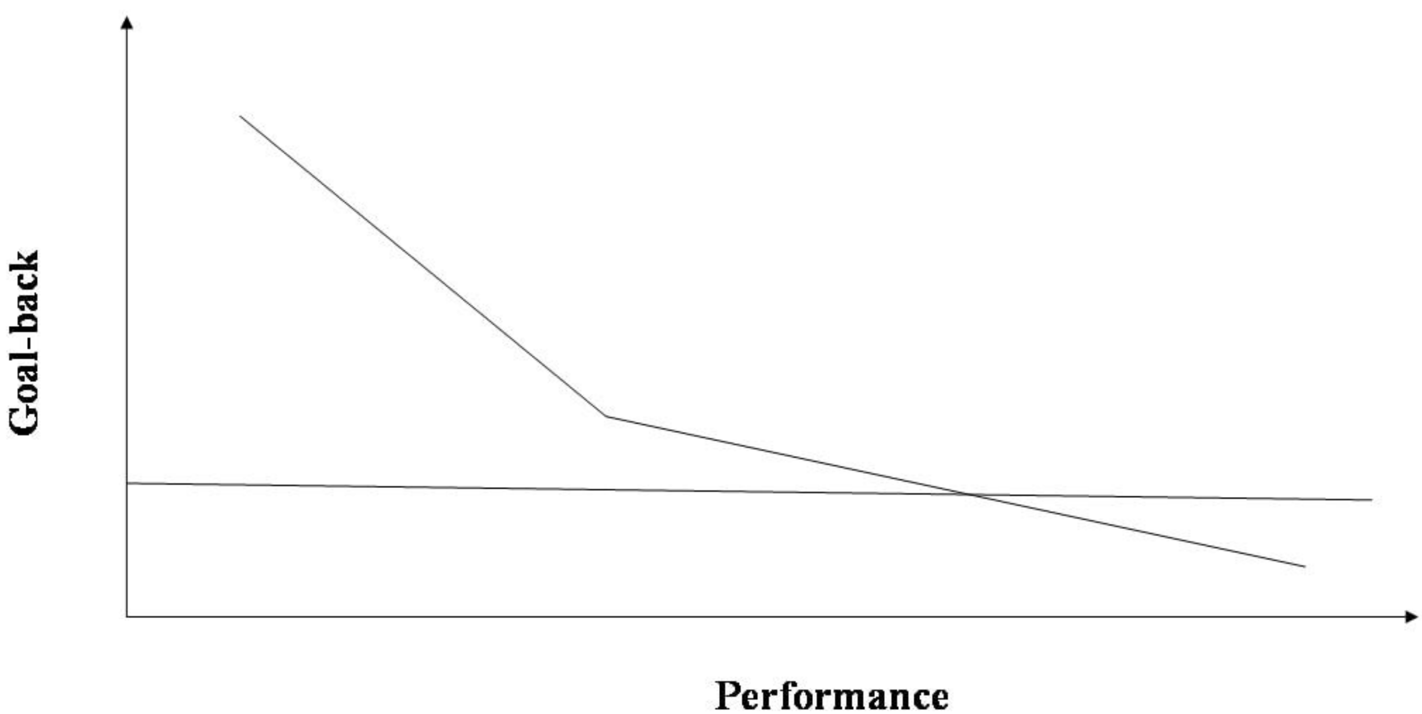

Figure 4

Error curve

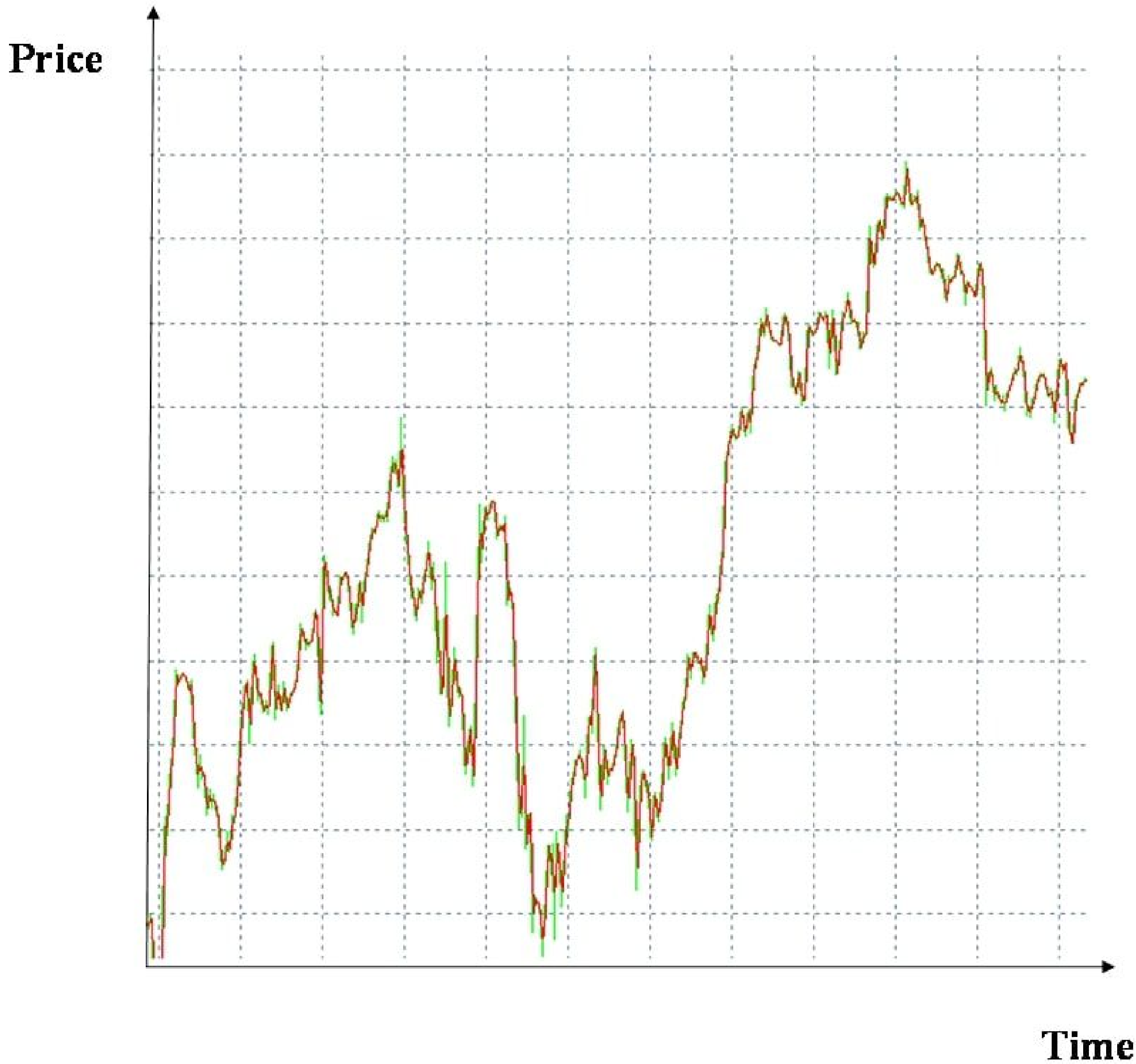

Figure 5

Fitting curve of test sample

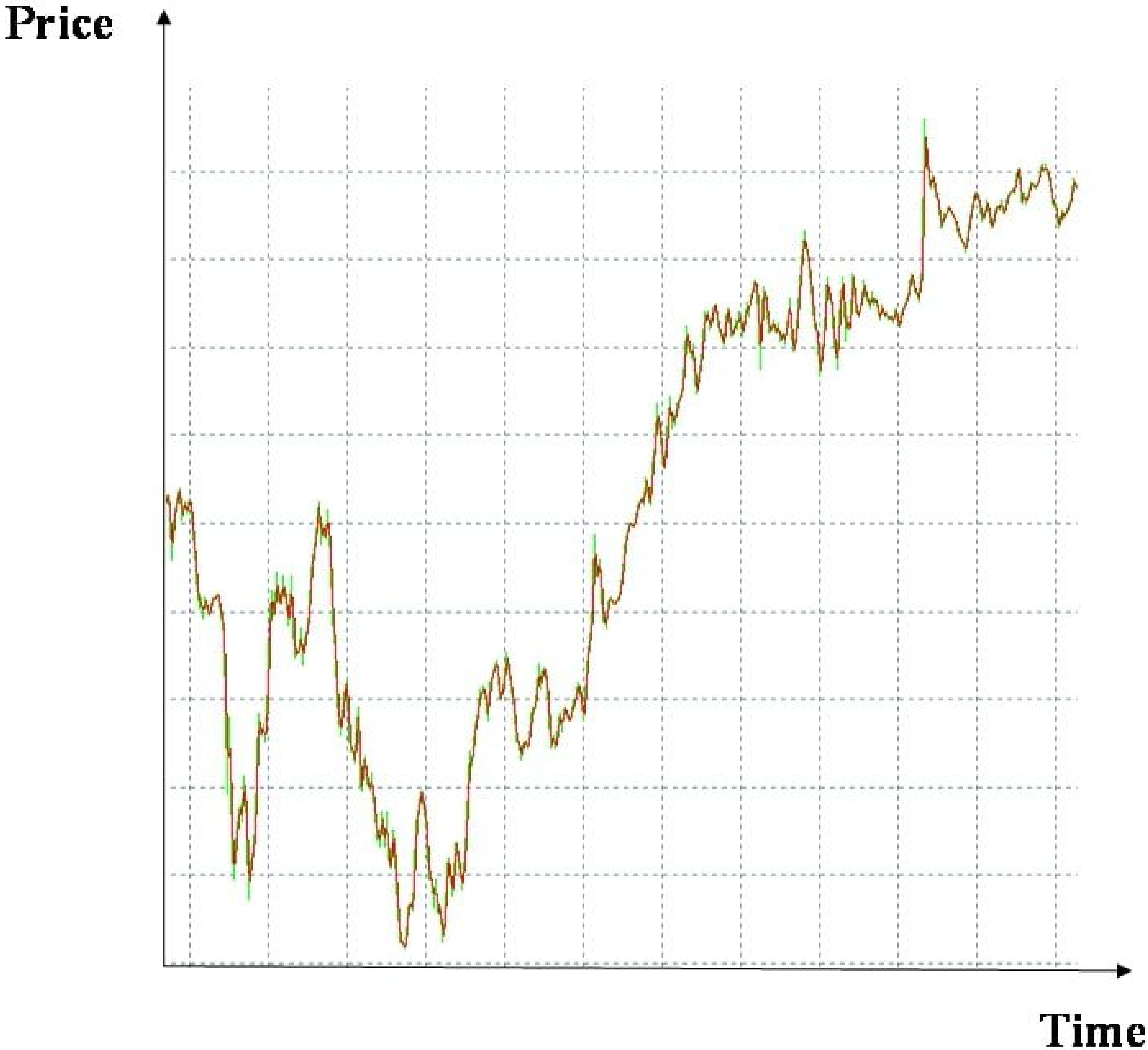

Figure 6

GA-BP network training

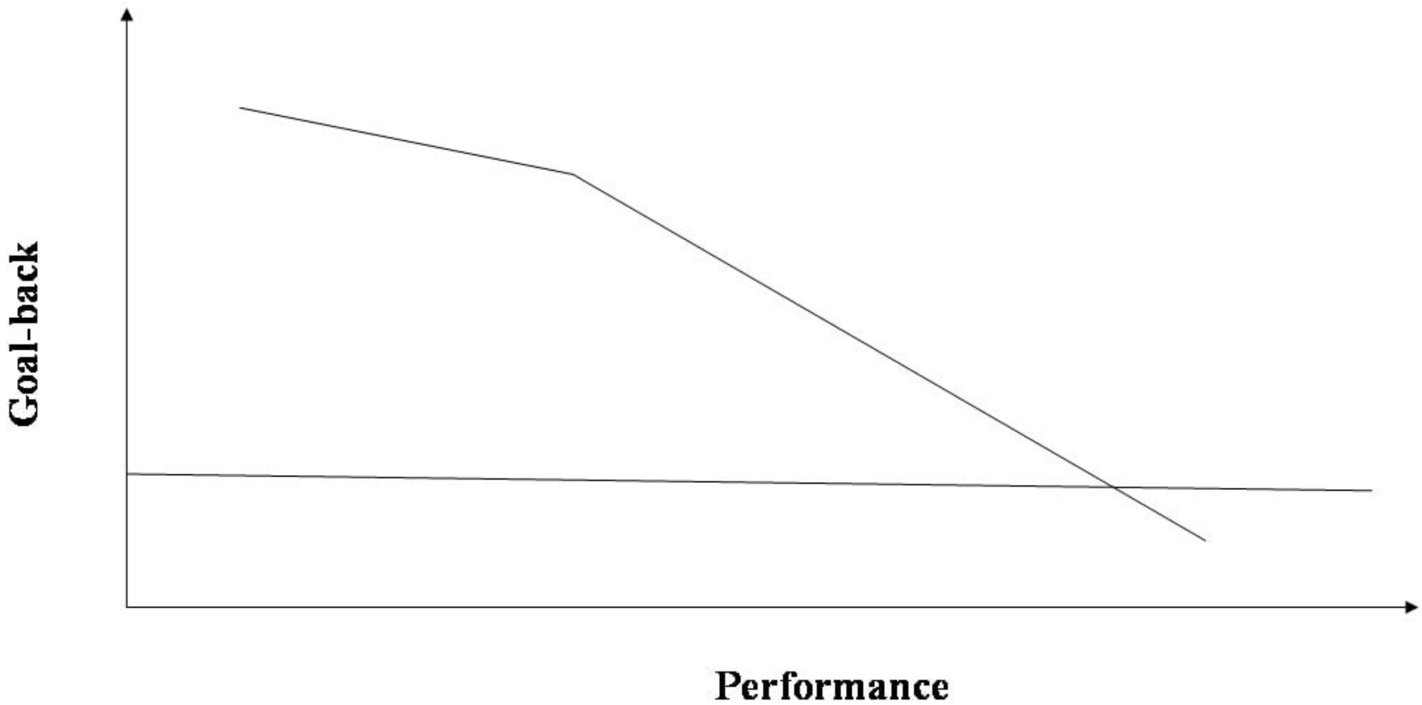

Figure 7

Training iteration diagram

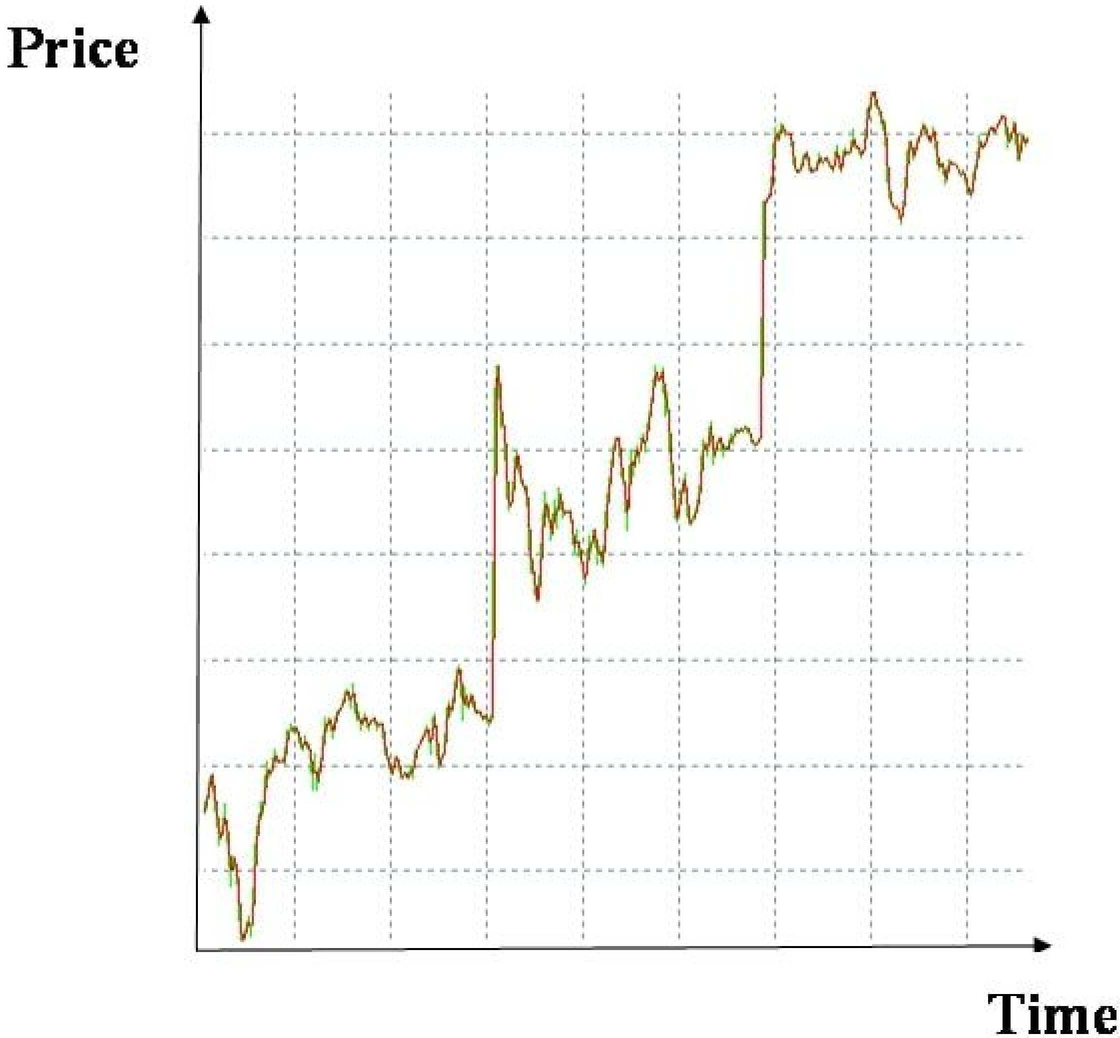

Figure 8

GA-BP fitting curve of test sample