The Application of BP Neural Network Optimized by Genetic Algorithm in Students' Comprehensive Quality Evaluation

Ren Rongrong [1], Fu Jia [1], Chao Jinbo[2]i, Yuan Huilin [1], Li Jing[3]

1 Northeastern University at Qinhuangdao, Hebei, China
2 Yanshan University, Qinhuangdao, Hebei, China
3 TBEA Shenyang Transformer Group Co.
high-voltage switch Institute, Shenyang, Liaoning, China

***Abstract* — This paper evaluates students' comprehensive quality by using a number of indicators. The students' comprehensive quality indicators are nonlinearity which influence each other. The BP network model is good at solving nonlinear problem, so a BP neural network assessment model optimized by genetic algorithm has been established to judge the quality of students. The method uses historical data to train BP neural network. The results implemented by MATLAB show that, neural network possesses memorizing and learning capability, and its function can be achieved true judgment. Compared with BP neural network model, the model has the advantages of less number of iterations, convergence speed and strong generalization ability.**

Keywords - BP neural network; Genetic algorithm; Optimized; Matlab

I. INTRODUCTION

The modern theory suggests that it's not scientific to evaluate students only based on learning scores. In order to avoid the disadvantage of judging only by examination, increasingly schools implement the evaluation of students' comprehensive quality. The comprehensive evaluation is almost focus on the indicators of moral, intelligence, health, aesthetic and labor (Yan Taishan 2005) .On the basis, this paper adds an item of personality development. It's necessary to pay attention to students' will and interest development, as the juvenile stage is the key stage for character formation. Schools should cultivate the students to be ideals, morality, culture and discipline, rather than the "reading machines".

In comprehensive evaluation, there are a number of factors that associate with each other, so it is difficult to express the evaluation results with a linear equation. At present, there are several researches on students' evaluation. For example, Zhou yanjun put forward the students' comprehensive quality evaluation model based on genetic algorithm(Zhou Yanjun 2008); Zhang xinliang put forward the model based on BP neural network(Zhang Xinliang 2012); Zhang wensheng put forward the algorithm design based on BP neural network(Zhang wensheng 2009); Wang xiaoxu put forward the model based on SOFM neural network(Wang xiaoxu 2011); Xiao mingwei put forward the model based on Analytic Hierarchy Process(AHP)(Xiao mingwei 2011); Lv hongbin put forward the model based on Fuzzy Evaluation(Lv hongbin 2007); Zhang ying put forward the model based on Fuzzy Analytic Hierarchy Process(FAHP)(Zhang ying 2007). Xiaofeng Li put forward the high school students' comprehensive quality evaluation model based on Rough and ANN Methods (Xiaofeng Li 2014).

Several problems can be found through literature analysis, the existing studies are mostly based on one algorithm to evaluate the students, but only one algorithm has some limitations. For example, BP neural network has the disadvantage of low convergence speed and falling easily into local minimum; SOFM neural network has the problem that the distortion will occur when the high-dimensional mapping to the low-dimensional. And it's widely used in the classification case, however, this paper is not a classification problem; In AHP, the data statistics will be large with the increasing of evaluation indicators. We need to compare the importance of each two indicators, and it's difficult to determine which is more important. Moreover, the compare results cannot easily pass the consistent test. Because of the complexity of the objective things and one-sidedness of the understandings, the eigenvectors (weights) obtained by constructed judgment matrix are not necessarily rational; FAHP improved the judgment matrix consistency problem of AHP, but the method to some extent affected by more subjective factors, so the index weight remains to be further improved.

The emergence of neural network provides a new approach for system evaluation. The neural network has many good qualities, such as self-adaptive and self-organization. It's good at making decision from the approximate, uncertain, and even contradictory knowledge environment, avoiding calculating the weights and correlation coefficient artificially (Xinhong Z 2002).

BP network is a highly complex nonlinear dynamical system to identify the model, the use of BP network overall quality of students to make the results more objective evaluation (Zhang qian 2013). But as explained, BP neural network has the disadvantage of low convergence speed and falling easily into local minimum. On that basis this paper establishes the students' comprehensive quality evaluation model based on BP neural network optimized by genetic algorithm. The genetic algorithm is a well-known random search and global optimization method based on the idea of natural selection and evolution. Rather than relying on the

gradient information, it searches the optimal solution by simulating the natural evolution process (Hongming 2013). The model combines the self-adaptive and self-organization of BP neural network with rapid global search ability of genetic algorithm, solves the problems of slow convergence speed and easily falling into local minimum. The practical simulation results prove that the method has stronger practicability.

II. BASIC PRINCIPLE OF BP NEURAL NETWORK OPTIMIZED BY GENETIC ALGORITHM

Usually, the BP algorithm adjusts the weights between neurons through some learning rules. In the learning process, the topological structure and network learning rules are constant. However, a neural network information processing function depends not only on the strength of the connections between neurons, but also on the network topology (neurons connection), characteristics of input and output neurons, and the weights and thresholds of neurons (Xiaofeng L 2004).

BP neural network has great dependence on the weights and thresholds, but if the initial values are closer to the true ones, the time of network training will be shorten obviously. This paper uses the basic genetic algorithm to optimize BP neural network weights and thresholds, leads the weights and thresholds close to the real values, rather than a simple random assignment. Then, the paper adopts the trained BP network as the model of students' comprehensive quality evaluation. Simulation results demonstrate that the genetic algorithm has a significant effect to accelerate the convergence speed.

The algorithm that optimizes BP neural network weights and thresholds with basic genetic algorithms is described as follows:

1) Encode BP neural network weights and thresholds with the real, and produce an initial population $W = (W_1, W_2, \cdots W_p)^T$ randomly, which has P individuals. Each individual is a string of real numbers, containing the connection weights of the input layer and hidden layer, hidden layer thresholds, hidden layer and output layer weights and the output layer thresholds.

2) Use the reciprocal of an error function as the fitness function. The error produced by the mean square error function of the desire output and the actual output of the network, determine the fitness of each individual by fitness function. Smaller the error is, greater the fitness is.

3) Use the roulette bet method as the selecting method. In the method, the selected probability of each individual is proportional to the fitness. The selected probability increases with increasing of fitness. Select the part of higher fitness individuals as the parents, and eliminate the lower ones.

4) Handle the parents to produce progeny by crossover and mutation operators. If the offspring's fitness is higher than the parent's, the parent's individual will be eliminated, and the new offspring will become the parent. Keep the number of parent individuals as a constant.

5) Repeat step (2) ~ (4), a new round of selection, crossover and mutation will be executed to the new group, until the termination condition is satisfied.

6) Decompose the optimal individuals of current groups into the connection weights and thresholds of BP neural network, and use the optimal weights and thresholds as the initial values of BP neural network.

7) Train BP neural network with preset algorithm parameters.

8) Stop training when the training objective is satisfied.

The procedure of BP neural network optimized by genetic algorithm is shown as Figure 1.n the training objective is satisfied.

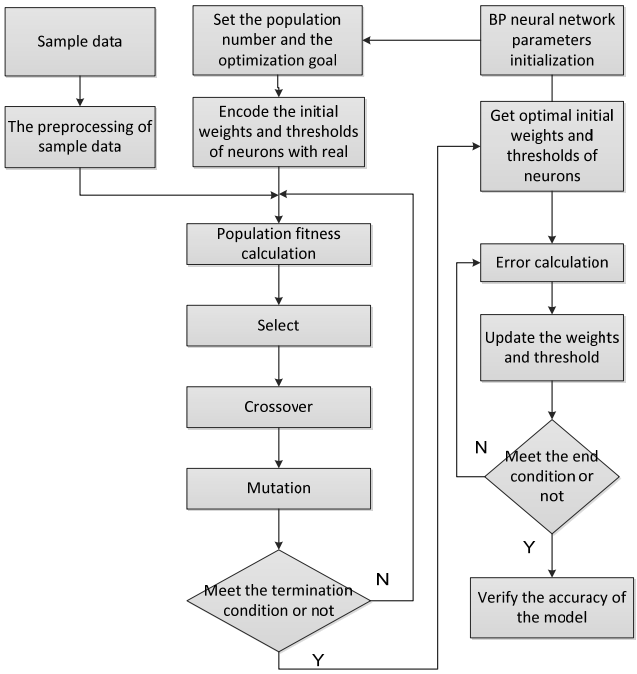

Fig 1. The procedure of BP neural network optimized by genetic algorithm

III. BP NEURAL NETWORK MODEL OPTIMIZED BY GENETIC ALGORITHM

A. BP Neural Network Model

BP neural network can theoretically approximate any nonlinear continuous function under the condition of reasonable structure and appropriate weights. It makes use of error gradient descent algorithm to minimize the mean square error between the output value of network and the actual output value. The network is composed of input layer nodes, hidden layer nodes (hidden layer can be one or more), and output layer nodes. The structure of BP neural network model is shown as Fig 2.

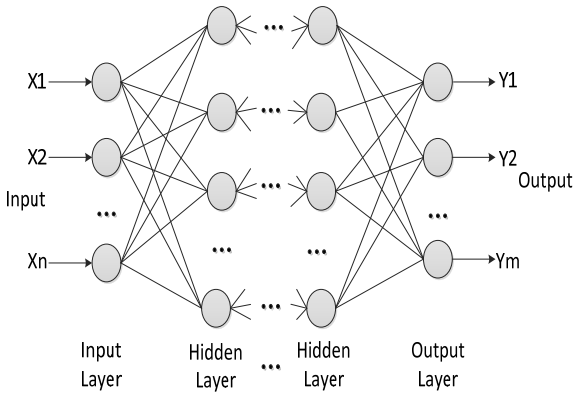

Fig.2 The structure of BP neural network model

There is no universal theory guidance for selecting how much hidden layers and hidden layer nodes of each layer, but after a lot of practice, predecessors have accumulated some experience. Theoretical analysis shows that, BP network with a single hidden layer can map the arbitrary continuous nonlinear function. Only when the learning is not a continuous function (such as a saw tooth wave), two hidden layers are needed. Increasing the hidden layer number can improve the training accuracy and reduce the errors, but the network will become complicated and time-consuming. In fact, the nodes of the hidden layer can be increased to improve the training accuracy, when the increased number of nodes is not significantly to reduce the error, trying to increase the number of hidden layers (Liqun H 2006). So this network uses three layers network, namely the input layer, the hidden layer and the output layer.

1) Design the input and the output layer

The design of input and output layer is according to the specific problems, there are 2 evaluation indexes and one evaluation result, so the input layer is set to n=19, the output m=1.

2) Design the hidden layer

In the BP network, hidden layer nodes are used to extracting and storing the samples inherent law from the samples, so setting the numbers of hidden layer nodes depend on the training sample size and complexity. There is no exact formula for the calculation of the hidden layer nodes, but the summary based on network structure obtained an empirical formula. The hidden layer nodes number $l=\sqrt{n+m}+\alpha$, n and m are the input and output layer nodes numbers, α is a constant between [1, 10]. After repeated testing, this paper determines the hidden layer node number is 11.

3) Design the driving function

In the BP neural network which is established in this paper, hidden layer transferring function uses bipolar S type function.

$$f(\mathrm{x})=\frac{1-e^{-x}}{1+e^{-x}} \tag{1}$$

The output layer transferring function uses linear function.

$$f(\mathrm{x})=x \tag{2}$$

The structure shown as Fig 3:

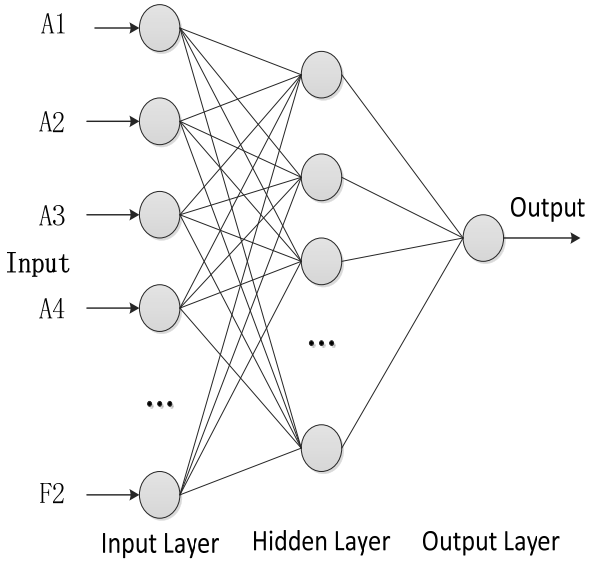

Fig 3. The structure of BP neural network model of students' comprehensive quality evaluation

B. The Coding Method Selection

Genetic algorithm needs to encode the initial weights and thresholds of BP network, and BP network weights values and thresholds values composed of many real values. So it will be the real if uses binary coding. This will affect the learning accuracy, so this paper adopts real coding method.

Assume each chromosome in the population contains S gene, then

$$S=R\times S_1+S_1\times S_2+S_1+S_2 \tag{3}$$

Where R is the number of input layer nodes; S_1 is the number of hidden layer nodes; S_2 is the number of output layer nodes. So, the chromosome length $S=19\times 11+11\times 1+11+1=232$, a total of 232 parameters need to be optimized.

C. Fitness Function Design

In the genetic algorithm, the value of fitness function directly determines the direction of evolution population. The fitness function is designed generally based on the objective function of optimization. In the network designed in the paper, the objective function makes the mean square error of all the sample output to the minimum, namely the minimum of equation (4).

$$E_{MSE} = \frac{1}{p}\sum_{p=1}^{p}(\mathrm{d}_p - \mathrm{y}_p)^2 \tag{4}$$

Where P is the number of training samples; d_p is the desired output of the p samples; y_p is the actual output of the p sample.

As the individuals with higher adaptive value are selected to evolve in genetic algorithm, the adaptive value function uses the reciprocal of sample output variance.

$$f = \frac{1}{E_{MSE}} \tag{5}$$

D. Design of Genetic Operation

Genetic operations include selection, crossover and mutation.

1) Selection

Selection operation is to choose better individual from initial population to do crossover and mutation operation. Therefore, the individuals with high fitness will gain the high selection probability. The selection strategy ensures the good genetic inheritance to the next generation (Cao chengzhi 2010), the selection probability of each individual is shown as follows.

$$p_i = f_i \Big/ \sum_{i=1}^{N} f_i \tag{6}$$

2) Crossover

Crossover is an operation that exchanges part chromosomes between a pair of parent individuals with a relatively large probability and then produces two new individuals. This operation can enlarge the diversity of solution and ergodicity of searching space (Feng Yu 2014).

The crossover adopts the real crossover method, namely the chromosome marked i and chromosome marked j cross at Position r.

$$a_{ir} = a_{ir}(1-c) + \mathrm{a}_{jr}\, c \tag{7}$$

$$a_{jr} = a_{jr}(1-c) + \mathrm{a}_{ir}\, c \tag{8}$$

The c is a random number between [0, 1].

3) Mutation

Mutation is a method to generate new individual by changing one or more genes on chromosome. The operation can offer new genes which the generation don't have or find the genes that losing in the selection process. It can provide new content for the population and enhance the diversity of the population.

Select the j genes of the individual marked i to compile operation.

$$a_{ij} = \begin{cases} a_{ij} + (\mathrm{a}_{ij} - \mathrm{a}_{\max})\,\mathrm{r}_2(1 - g/G_{\max})^2 & r_1 \geq 0.5 \\ a_{ij} + (\mathrm{a}_{\min} - \mathrm{a}_{ij})\,\mathrm{r}_2(1 - g/G_{\max})^2 & r_1 < 0.5 \end{cases} \tag{9}$$

Where $\mathrm{a}_{\max}$ and $\mathrm{a}_{\min}$ are the upper and lower bounds of gene a_{ij}; r_1 is a random number between [0, 1]; r_2 is another random number; g is the current iterations; $G_{\max}$ is the maximum number of evolution.

E. Design of Termination Condition

In the BP neural network algorithm , when the optimal individual fitness reaches a given threshold, or the best individual fitness and population fitness will not rise any more, even or the number of iterations reaches a preset algebra, the algorithm will be terminated.

In this paper, the algorithm selects the default algebra as 100 generation, and terminates when the iterations reaches a preset algebra.

F. BP Neural Network Prediction

Take the weights and thresholds optimized by genetic algorithm as the BP neural network weights and thresholds, then use them to predict.

IV. THE EXAMPLE RESEARCH OF STUDENTS' COMPREHENSIVE QUALITY EVALUATION

This paper evaluates students' comprehensive quality by using a number of indicators. The indicators contain moral, intellectual, health, aesthetic, labor and personality development. The detail evaluation indicators show as Fig.4.

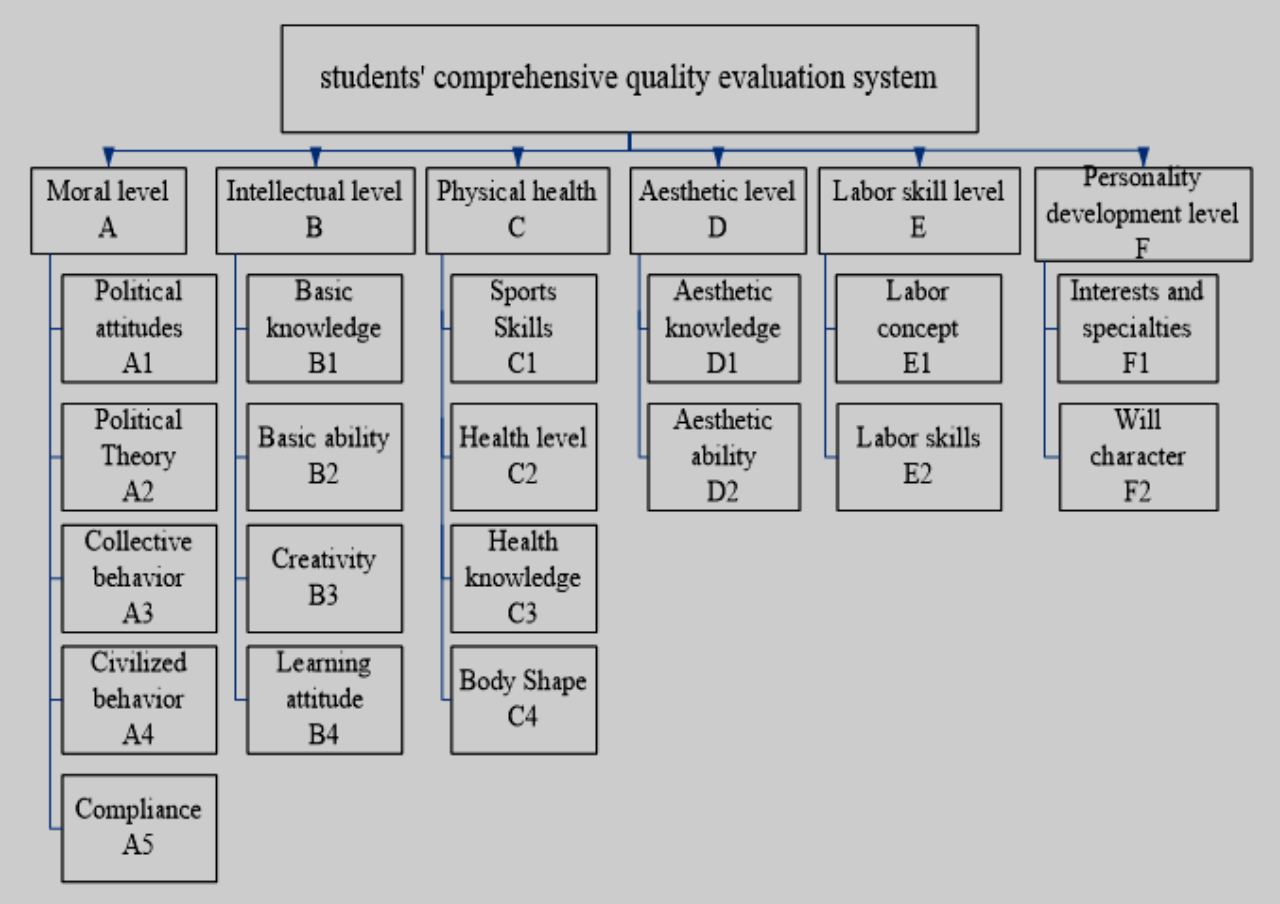

Fig.4 The system of the students' comprehensive quality evaluation

This paper adopts the standard score, a relative quantity gained through the transformation of original score, to make calculation. It reflects the position of a student in the overall students. So, whether the test is hard or easy and the whole original score is high or low, the overall standard score will not change. Therefore this paper uses the standard score which between [0, 1].The historical data samples of students' comprehensive quality evaluation are shown in Tab 1:

TABLE I. SAMPLE DATA

A_1	A_2	A_3	A_4	A_5	…	F_2	Evaluation objectives
0.95	0.95	0.9	0.85	0.95	…	0.95	0.92
0.95	1	0.95	0.9	0.95	…	0.85	0.93
0.9	0.95	1	0.9	0.85	…	0.9	0.91
1	0.9	0.9	0.95	0.95	…	1	0.91
…	…	…	…	…	…	…	…
0.65	0.6	0.6	0.8	0.7	…	0.6	0.635
0.65	0.85	0.8	0.7	0.55	…	0.6	0.678
0.6	0.7	0.5	0.55	0.6	…	0.55	0.57
0.7	0.65	0.55	0.5	0.5	…	0.6	0.59

A. Training and Results

100 groups of data as the training sample are used to train BP neural network model, network structure is BP (11, 19, 1). The BP neural network is trained and optimized by genetic algorithm with Matlab, the results are shown as Fig 5 and Fig 6.

The results show that the convergence speed and error are greatly improved by genetic algorithm optimization.

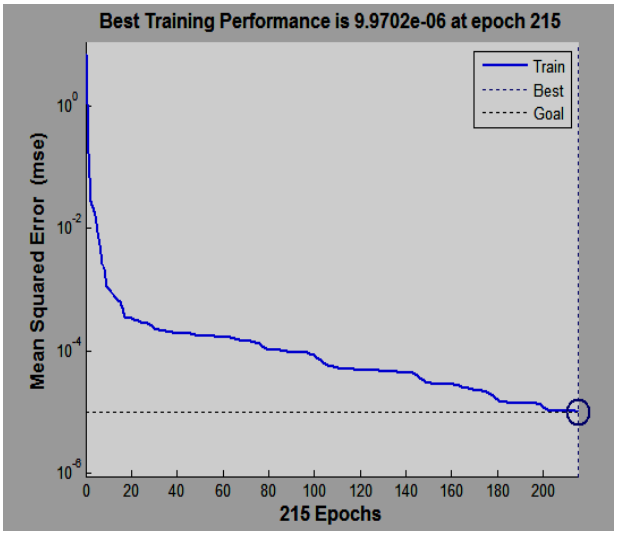

Fig.5 The training results of the BP neural network

A. *Simulation and Prediction*

After training the BP neural network, 6 groups of data which is not trained are used to test the network. The costract of test results and specialist evaluation results is shown as follows.

The results show that the BP neural network optimized by genetic algorithm has good performance in students' comprehensive quality evaluation. The evaluation is very similar to the specialist evaluation, the errors are smaller than 2 percent. The results proved that the network was trained successful and had better generalization ability.

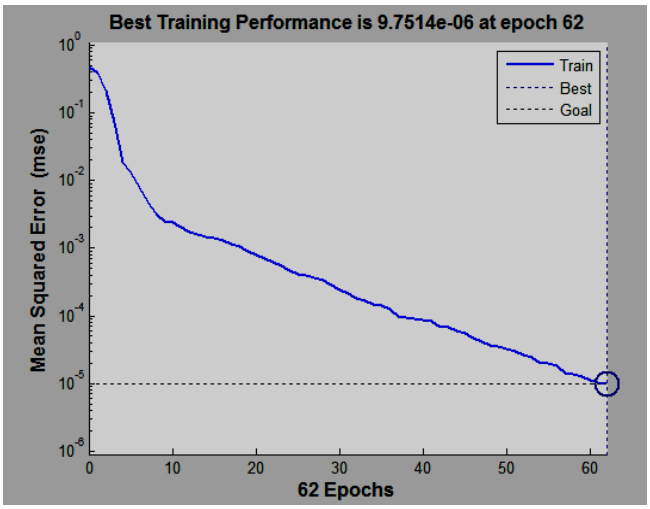

Fig.6 The training results of the BP neural network optimized by genetic algorithm

TAB.2 TEST RESULT OF BP NEURAL NETWORK OPTIMIZED BY GENETIC ALGORITHM

A_1	A_2	A_3	A_4	A_5	…	F_2	specialist evaluation	test results	error
1	0.95	1	0.95	0.85	…	0.9	0.9300	0.9411	-1.11%
1	0.95	0.9	0.95	0.95	…	0.9	0.9200	0.9241	-0.41%
0.9	0.9	0.85	0.9	0.85	…	0.85	0.8600	0.8783	-1.83%
0.85	0.9	0.85	0.85	0.9	…	0.8	0.8500	0.8665	-1.65%
0.65	0.8	0.8	0.75	0.55	…	0.65	0.6900	0.7011	-1.11%
0.65	0.6	0.55	0.55	0.6	…	0.55	0.5700	0.5787	-0.87%

V. CONCLUSION

This paper establishes the system of students' comprehensive quality evaluation, and the indicators contain moral, intellectual, health, aesthetic, labor and personality development that can evaluate the students in various aspects. This paper elaborates the advantages and disadvantages of the method based on BP neural network. Through introducing the genetic algorithm, a new students' comprehensive evaluation model based on BP neural network optimized by genetic algorithm is proposed and applied to the students' comprehensive evaluation. The simulation results show that the BP neural network optimized by genetic algorithm has good application in students' evolution. The method improves the convergence speed greatly, and it does not fall easily into local minimum.

ACKNOWLEDGMENT

The authors are grateful to the valuable and constructive comments from the editor and anonymous referees. This work is sponsored by the National Natural Science Fund (No. 60821063) and Science and Technolgy Plan Project of Hebei Province (No. 13214716).

REFERENCES

[1] Yan Taishan 2005.Application of BP neural network in the student quality comprehensive assessment system.China Science and Technology Information,13:81-82.

[2] Zhou Yanjun 2008. Students comprehensive quality evaluation based on genetic algorithm.Journal of Yangtze University(Natural Science Edition) Sci & Eng V, 5(4):317-319.

[3] Zhang Xinliang,Min Jiayuan 2012. Application of BP neural network in the evaluation of comprehensive quality of College Students. The fifth educational reform and Management Engineering Symposium.

[4] Zhang Wensheng 2009.BP neural Network Based on the Overall Quality of the Student Evaluation of Algorithm Design. Computer Knowledge and Technology.

[5] Wang Xiaoxue,Wang Linshan 2011. Comprehensive Evaluation of Students Based on SOFM Neural Networks. Journal of Hebei Normal University (Natural Science Edition), 239-243.

[6] Xiao Mingwei 2011. Research on comprehensive quality evaluation of middle school students based on Analytic Hierarchy Process. Liaoning Normal University.

[7] Lv Hongbin 2007. Study on College Students' comprehensive quality evaluation system based on fuzzy evaluation. Shandong Normal University.

[8] Zhang Ying, Feng Yanfang 2007. FAHP Evaluation on General Quality of Undergraduate. Journal of Wuhan University of Technology, 707-710.

[9] Xiao-feng Li, Ling-yuan Liu 2014.The Evaluation of College Students' Comprehensive Quality Based on Rough and ANN Methods. Electrical Engineering, 241:483-494.

[10] Xinhong Z, Pie Z 2002. Neural networks based integrated evaluation method for MIS Journal of Systems Engineering, 17(5):63-68.

[11] Zhang qian, Wang ling, Ma ying 2013. Directors of Comprehensive Quality System of College Students in China: Based on BP Neural Network. Management Science and Engineering, 7(3):99-102.

[12] Hongming Yang, Jun Yi, Junhua Zhao, Zhaoyang Dong 2013. Extreme learning machine based genetic algorithm and its application in power system economic dispatch.Neurocomputing, 102:154-162.

[13] Xiaofeng L, Jiupin X 2004.The establishment of BP adaptive artificial neural network learning algorithm and its application. Journal of Systems Engineering,24(5):1-8 .

[14] Liqun H 2006. Artificial neural network tutorial. Beijing University of Posts and Telecommunications press, 34-46.

[15] Cao chengzhi 2010. Artificial intelligence technology. Beijing Tsinghua University press, 326-415.

[16] Feng Yu, Xiaozhong Xu 2014. A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. Applied Energy, 134: 102 – 113.