Model of Normal University Students’ Education Quality Based on Artificial Neural Networks

Shaofeng Wang¹, Yue He² and Haitao Lin¹*  
¹Yuxi Normal University, 134# Fenghuang Road, Yuxi, 653100, China  
²China National Tobacco Corporation Yunnan Provincial Company, 263# Tuodong Road, Kunming, 650041, China  
*Email: drhtlin@yxnu.edu.cn

Abstract: As a main form of checking out learning and teaching effects, exam has traditionally been used in schools. Accurate and objective assessment of students’ learning effects through examination performance is an important part of evaluation of students’ education. This paper proposes a comprehensive evaluation model of evaluating examination performance based on quality of examination paper, explores the application of artificial neural network technology in students’ comprehensive education evaluation, focusing on BP neural network, some works can also be introduced by Python. The process is to compose input vector of BP neural network using the four evaluating indicators of quality of examination paper, namely reliability, validity, difficulty and discrimination, and raw score, namely the students’ score of examination paper. Take valuation namely the quantitative value of learning effect as the output vector of BP neural network. Design a reasonable network structure and training sample, put the training sample in the network for processing till until systematic error meets the specified requirements. By so, the obtained network model is the required comprehensive evaluation model of examination performance. At last, the paper analyzes the simulation of the evaluating model’s feasibility using MATLAB software, obtains satisfying results, and proves the feasibility of the proposed network model.

1. Introduction

Education evaluation [1,2] is a higher level requirement after the completion of educational informative construction. It makes value judgments of educational evaluation through systematic information search, analysis and explanation using feasible scientific methods under the guideline of certain educational values or target, thus to provide foundation for continue education optimization and educational decision making. Educational evaluation has developed into a rather complete theoretical system through the efforts of researchers since its emergence in the 1930s. Its evaluation objects not only cover the earlier learning effect of students, but also cover the education planning, education activities, even the whole education process. The applied methods in evaluation [3-7] include linear programming, dynamic programming, data envelopment analysis, analytic hierarchy process method, regression analysis, factor analysis, cluster analysis, Markov chains and so on. 

As a main form of checking out learning and teaching effects, exam has traditionally been used in schools. Accurate and objective assessment of students’ learning effects through examination performance is an important part of evaluation of college students’ education. The objectivity of examination performance, not only relates to the objectivity of evaluation of students' learning and abilities, but also relates to objectivity of the evaluation of teachers' teaching abilities and teaching
effects. However, many objective factors affect examination performance such as paper quality, invigilation, scoring and the ability level of students. An examination paper consists of a large number of questions, and the quality of questions is crucial to the quality of examination paper, while the quality of examination paper also has a direct impact on students’ examination performance [8]. So the students’ learning effect cannot be reflected accurately only by the judgment of examination performance.

The development of modern statistical analysis technology and computer information processing technology provides a scientific method to analyze quality of examination paper and evaluate examination performance. BP neural network is strong in nonlinear processing capability. An evaluation model based on BP neural network is not only great in objectivity, also is convenient for computer programming. This paper studies on the comprehensive evaluation of exam performance based on the quality of examination paper, which is actually a complex integrated decision making. In the end, this paper simulates the feasibility of the evaluation model using MATLAB software.

2. Research Methods

2.1. Overview of BP Neural Network Model [9-11]

Artificial neural network models involve perception, BAM network, BP network, Boltzmann machine, RBF neural network, Hopfield neural network etc. Among them, the most widely used is BP network model. It is a theoretically sophisticated artificial neural network model. BP algorithm is a kind of mentor learning algorithm, belonging to feed forward neural network.

In the network do not have internal state. The study of samples is achieved by adjustment of weight \(w_{ij}\) and threshold \(\gamma_i\). BP neural network is composed of input layer, hidden layer and output layer. The number of nodes in traditional BP network is determined by the number of input and output parameters. The hidden layer can be one layer or multiple layers. The learning process of BP network is composed of forward propagation and back propagation. In the process of forward propagation, the input mode after passing the input layer is processed by the hidden layer neurons then is transferred to the output layer. The status of neurons in this layer only has effects on those of its following neighbour. If the expected output is not obtained in the output layer, then the model is transferred into back propagation. At this moment, error signal is transferred from the output layer to the input layer, and the interlayer connection weights and thresholds are adjusted along the way so that the error decreases until it reaches the precision requirements. The algorithm is to find the minimum value of the error function in essence. It repeats training of multiple samples, uses the fastest descent method to make the weight follow the direction of negative gradient of error function and converge to the smallest point.

The nodes of the three layers of BP network can be represented as follow: input node is \(x_j\), hidden layer node is \(y_i\), output node is \(o_l\). The network weight between input node and hidden layer node is \(w_{ij}\), network weight between hidden layer node and output node is \(T_{iy}\). When the expected output of output node is \(t_l\), the calculating formula of BP network model is as follows.

1. Calculating formula of output \(o_l\) of output node
   ① Input of input node: \(x_j\)
   ② Output of hidden layer node: \(y_i = f \left( \sum_j w_{ij} x_j - \theta_i \right)\), in which, the connection weight is \(w_{ij}\), node threshold is \(\theta_i\).
   ③ Output of output node: \(o_l = f \left( \sum_i T_{iy} y_i - \theta_l \right)\), in which, the connection weight is \(T_{iy}\), node threshold is \(\theta_l\).

2. The modification formula of output layer(between hidden layer node and output node)
1. The expected output of output node: \( t \)
2. Error control

Error of all samples: \( E = \sum_{k=1}^{P} e_k < \varepsilon \), error of one sample: \( e_k \), in which, \( P \) is the number of samples, \( n \) is the number of output nodes.
3. Error formula: \( \delta_i = (t_i - o_i) \cdot a_i (1 - a_i) \).
4. Modification of weight: \( T_k(k+1) = T_k(k) + \eta \delta_i y_i \), in which, \( k \) is iteration.
5. Modification of threshold: \( \theta_i(k+1) = \theta_i(k) + n' \delta'_i \)

(3) The modification formula of hidden layer(between input node and hidden layer node)
1. Error formula: \( \delta'_i = y_i (1 - y_i) \sum_k \delta_k T_k \).
2. Modification of weight: \( w_{ij}(k+1) = w_{ij}(k) + \eta' \delta'_i x_j \).
3. Modification of threshold: \( \theta_i(k+1) = \theta_i(k) + \eta' \delta'_i \).

(4) Generally the transfer function is \((0,1)\), system function is \( f(x) = \frac{1}{1+e^{-x}} \).

(5) Error function

The formula for error of sample P is \( E_P = \sum_{k=1}^{P} \frac{(t_{PK} - o_{PK})^2}{2} \). In the formula, \( t_{PK} \) is the expected output and \( o_{PK} \) is the calculating output of network.

The learning process of evaluation model of college students’ educational neural network is composed of forward propagation and back propagation. In the process of forward propagation, enter information \( x_j \), it passes from input layer through hidden layer into output layer. If the output information \( o_i \) is not the expected output, then error back propagation is adopted by turning the error signal back and adjusting the weight \( w_{ij} \) in each level, so that the error signal is narrowed to its minimum. When the error reaches to its minimum or all samples gain the expected output, the learning process is finished.

2.2. Basic Idea

The basic idea of applying BP neural network in comprehensive evaluation of examination performance is to compose input vector of BP neural network using the four evaluating indicators of quality of examination paper, namely reliability, validity, difficulty and discrimination, and raw score, namely the students’ score of examination paper. Take valuation namely the quantitative value of learning effect as the output vector of BP neural network. Design a reasonable network structure and training sample, put the training sample in the network for processing till until systematic error meets the specified requirements.

3. Evaluation Model of College Students’ Education Based on BP Neural Network [12-17]

3.1. Indicators of Comprehensive Evaluation Model of Examination Performance

As a main form of checking out learning and teaching effects, exam has traditionally been used in schools. Accurate and objective assessment of students’ learning effects through examination performance is an important part of evaluation of college students’ education. The objectivity of examination performance, not only relates to the objectivity of evaluation of students' learning and abilities, but also relates to objectivity of the evaluation of teachers' teaching abilities and teaching effects. However, many objective factors affect examination performance such as paper quality, invigilation, scoring and the ability level of students. Among them the quality of examination paper
has a direct impact on students’ examination performance. And the students’ learning effect cannot be reflected accurately only by the judgment of examination performance. Besides, the difficulty of papers should be measured by scientific standards. The following indicators are adopted as objective measure standards widely in current examination quality at home and abroad.

(1) Reliability
Reliability refers to the degree of whether the examination results are authentic and objective reflection of the actual level of the examinee. The higher the reliability is, the more reliable and stable examination results are. In this paper, reliability is calculated by Celebrook formula (also known as $\alpha$ coefficient method), which is shown as below.

$$\alpha = \frac{m}{m-1} \left(1 - \frac{\sum_{i=1}^{m} S_i^2}{S_y^2}\right).$$

In the formula, $S_i^2$ is the variance of examinee’s score of $i$, $S_y^2$ is the variance of examinee’s total score, $m$ is the number of questions in the paper. $r = \frac{\sum_{i=1}^{m} (X_i - \bar{X})(A_i - \bar{A})}{nS_xS_a}$.

In the formula, $X_i, A_i$ stand for the scores of the current test and criterion test respectively, $\bar{X}, \bar{A}, S_x, S_a$ are arithmetic mean value and standard deviation of all scores of the two tests respectively. $n$ is the number of examinees.

(2) Difficulty
Difficulty refers to the degree of difficulty of the examination questions, is an indicator of appropriateness between the examination questions and the level of students’ knowledge and abilities. In educational measurement, the difficulty of objective questions is generally measured by ratio of the number of examinees who offers no answer or wrong answer and the total number of examinees. The difficulty of subjective questions is calculated by the formula

$$H = 1 - \frac{\bar{X}}{A}.$$

In the formula, $\bar{X}$ stands for the average score of a certain question of the entire students. $A$ stands for the provisioned full score of this question.

(3) Discrimination
Discrimination refers to discernibility of the questions in differentiating the examinees’ actual level of ability. Commonly used grouping formula is $D = \frac{(H - L)}{F}$.

In the formula, $H$ stands for the average score of a certain question in high score group, $L$ stands for the average score of this question in low score group, $F$ stands for the provisioned full score of the question.

3.2. Comprehensive Evaluation of BP Neural Network Design
The comprehensive evaluation model of examination performance is a three-layer BP neural network composed of multiple input unit, single hidden layer unit and single output unit. Its structure is shown as Figure 1.
The number of input nodes in BP network model of comprehensive evaluation equals with the number of sub-index of each object to be evaluated, the number of hidden layer nodes is determined by trial-and-error method. There is only one node in output layer, which stands for the overall evaluation indicator of the number i objects to be evaluated. And the transition function is
\[ f(x) = \frac{1}{1 + e^{-x}}. \]

This paper takes reliability, validity, difficulty, discrimination, and raw score as the input of a neural network, takes value as the output of the neural network. As shown in figure 1, the five nodes in input layer x1, x2, x3, x4, x5 represents reliability, validity, difficulty, discrimination, and raw score respectively. There are 11 nodes in hidden layer, and one node in output layer. z is the value falling into the range of [0,1].

3.3. Determination of Training Sample and Network Training

A learning sample is composed of input sample and output sample. The input sample is
\[ x(i) = \{x_1, x_2, x_3, x_4, x_5\}. \]

The output sample z is the overall index of comprehensive evaluation, and determined by the formula
\[ z = \sum_{j=1}^{5} w_j x_j, \]
where \( x_j \) is the value falling into the range of [0,1].

Based on the above principle and methods, this paper designs a training sample as shown in the following table using the following methods: in accordance with the range of evaluation indicators, combined with the qualitative relationship among evaluation indicators, a group of typical index value is chosen. Together with part of the actual data obtained from students’ examination of a school, the correspondent value can be achieved, thus comes out the expected training sample, shown as Table 1.

| Sequence Number | Reliability | Validity | Difficulty | Discrimination | Raw Score | Value |
|-----------------|-------------|----------|------------|----------------|-----------|-------|
| 1               | 0.63        | 0.79     | 0.1        | 0.2            | 35        | 0     |
| 2               | 0.63        | 0.79     | 0.1        | 0.2            | 45        | 0.1   |
| 3               | 0.63        | 0.79     | 0.1        | 0.2            | 55        | 0.2   |
| 4               | 0.63        | 0.79     | 0.1        | 0.2            | 65        | 0.3   |
| 5               | 0.63        | 0.79     | 0.1        | 0.2            | 75        | 0.4   |
| 6               | 0.94        | 0.97     | 0.3        | 0.5            | 35        | 0.2   |
| 7               | 0.94        | 0.97     | 0.3        | 0.5            | 45        | 0.3   |
4. Simulation Experiment

By applying comprehensive evaluation model of examination performance based on neural network, this paper makes a comprehensive evaluation of examination performance of students in recent three years studying a course named Data Structure of a certain school. In the experiment, the neural network toolbox in MATLAB is used for simulation. Therefore, the values in Table 2 are obtained.

Table 2. Simulation Results of Examination Samples.

| Sequence Number | Reliability | Validity | Difficulty | Discrimination | Raw Score | Value   |
|-----------------|-------------|----------|------------|----------------|-----------|---------|
| 8               | 0.94        | 0.97     | 0.3        | 0.5            | 55        | 0.4     |
| 9               | 0.94        | 0.97     | 0.3        | 0.5            | 65        | 0.5     |
| 10              | 0.94        | 0.97     | 0.3        | 0.5            | 75        | 0.6     |
| 11              | 0.98        | 0.99     | 0.5        | 1              | 35        | 0.4     |
| 12              | 0.98        | 0.99     | 0.5        | 1              | 45        | 0.5     |
| 13              | 0.98        | 0.99     | 0.5        | 1              | 55        | 0.6     |
| 14              | 0.98        | 0.99     | 0.5        | 1              | 65        | 0.7     |
| 15              | 0.98        | 0.99     | 0.5        | 1              | 75        | 0.8     |
| 16              | 0.94        | 0.97     | 0.7        | 0.5            | 35        | 0.2     |
| 17              | 0.94        | 0.97     | 0.7        | 0.5            | 45        | 0.3     |
| 18              | 0.94        | 0.97     | 0.7        | 0.5            | 55        | 0.4     |
| 19              | 0.94        | 0.97     | 0.7        | 0.5            | 65        | 0.5     |
| 20              | 0.94        | 0.97     | 0.7        | 0.5            | 75        | 0.6     |
| 21              | 0.63        | 0.79     | 0.9        | 0.2            | 35        | 0       |
| 22              | 0.63        | 0.79     | 0.9        | 0.2            | 45        | 0.1     |
| 23              | 0.63        | 0.79     | 0.9        | 0.2            | 55        | 0.2     |
| 24              | 0.63        | 0.79     | 0.9        | 0.2            | 65        | 0.3     |
| 25              | 0.63        | 0.79     | 0.9        | 0.2            | 75        | 0.4     |
| 26              | 0.76        | 0.87     | 0.53       | 0.22           | 35        | 0.3     |
| 27              | 0.76        | 0.87     | 0.53       | 0.22           | 45        | 0.4     |
| 28              | 0.76        | 0.87     | 0.53       | 0.22           | 55        | 0.5     |
| 29              | 0.76        | 0.87     | 0.53       | 0.22           | 65        | 0.6     |
| 30              | 0.76        | 0.87     | 0.53       | 0.22           | 75        | 0.7     |
| 31              | 0.44        | 0.55     | 0.45       | 0.16           | 35        | 0.1     |
| 32              | 0.44        | 0.55     | 0.45       | 0.16           | 45        | 0.2     |
| 33              | 0.44        | 0.55     | 0.45       | 0.16           | 55        | 0.3     |
| 34              | 0.44        | 0.55     | 0.45       | 0.16           | 65        | 0.4     |
| 35              | 0.44        | 0.55     | 0.45       | 0.16           | 75        | 0.5     |
| Sample | Value | Raw Score |
|--------|-------|-----------|
| Sample 2 | 0.44  | 0.66      | 0.45 | 0.16 | 46 | 0.40294 |
|         | 0.44  | 0.66      | 0.45 | 0.16 | 48 | 0.40925 |
|         | 0.44  | 0.66      | 0.45 | 0.16 | 51 | 0.41788 |
|         | 0.44  | 0.66      | 0.45 | 0.16 | 61 | 0.44165 |
|         | 0.76  | 0.87      | 0.53 | 0.22 | 46 | 0.41217 |
| Sample 3 | 0.76  | 0.87      | 0.53 | 0.22 | 51 | 0.46163 |
|         | 0.76  | 0.87      | 0.53 | 0.22 | 61 | 0.55836 |
|         | 0.76  | 0.87      | 0.53 | 0.22 | 70 | 0.64776 |

Figure 2. Contrast Curves of Simulation Results.

Figure 2 is the contrast curves of simulation results of the three samples in table 2. As can be seen from the figure, value of each sample increases as the increase of raw score, which is in consistence with the quantitative relationship between raw score and value.

5. Conclusion
This study indicates that the model is reasonable. According to people's understanding, with the same raw score, the values should be increasing in comparing the three sets of samples, namely, value of sample 2 should be higher than value of sample 1 and lower than value of sample 3. However, figure 2 fails to show this trend, which indicates that the model is inadequate. As for the reason, it is mainly because the model is established on training samples lack of accuracy, which will be focused in the future study.

Acknowledgments
This work was supported by Yunnan Social Sciences Research Projects (YB2020066).

References
[1] Hu Z F, Li F 2000 *Educational Measurement and Evaluation (Volume 1)* (Guangdong: Guangdong Higher Education Press) 168 p

[2] Zhang Y Z 2002 *Research on the Methods of Higher Educational Evaluation*. (Shanghai: Fudan University Press) 202 p

[3] Ju L, Xi H F 2010 Research of the Graduate Education Quality Evaluation Body in China. *J. Chongqing University of Posts and Telecommunications (Social Science Edition)* 15 257

[4] Chen X, Wang X J, Gu A H, Sun Y T 2010 Implementation and Improvement of Education Quality Evaluation System in Open and Distance Education— Practice and Exploration of Shanghai TV University. *J Tianjin Radio & Television University* 63 1022

[5] Zhang Z Y 2010 The Overall Construction of Higher Education Quality Evaluation System.J Panzhihua University 10 531

[6] Cheng X 2012 Human Capital Evaluation of Professional Degree Postgraduates and Appraisal of Its Education Quality *J Shandong University of Finance and Economics* 12 1341

[7] Ma Y H, Guo J 2011 Construction of Assessment Indicator System of Teacher Education Technology Capacity Training in Primary and Secondary School *J China Educational Technology* 12 923

[8] Zhao L X, Chen W Y, Guo Z J. 2004 *The Quantitative Evaluation on Quality of Test Paper* *J South China Agricultural University (Social Science Edition)* 3 141

[9] Hu S R 2002 *Introduction to Neural Network* (Beijing: National University of Defense Technology Press) 232 p

[10] Wang H Y, Shi G D 2002 *Artificial Neural Networks Technology and Its Application*. (Beijing: China Petrochemical Press) 202 p

[11] Yan T S 2006 An Evaluation Method of Graduate Prihects' Quality Based on Artificial Neural Network *J Pioneering With Science & Technology Monthly* 1 137

[12] She L 2001 *A Study of Evaluating Teaching Quality Based on Improved BP Neural Network*. (Changsha: Central South University) 2033 p

[13] Tang J H, Zhang X Q, Cheng L 2013 Evaluation of Teacher Education Technical Ability Training Based on BP Neural Network *J Computer Technology and Development* 6 1252

[14] Tan L Z, Wang Y, Gao J. 2015 Research on Entrepreneurial Evaluation System of College Students Based on BP Neural Network *J Zhejiang Sci-Tech University* 68 3489

[15] Wen S, Liu W, Yang Y, Zhou P, Yan Z, Guo Z, Chen Y, and Huang T, 2020 Multi-label image classification via feature/label co-projection, *IEEE Transactions on Systems, Man and Cybernetics: Systems*, doi: 10.1109/TSMC.2020.2967071

[16] Wen S, Wei H, Huang T, and Zeng Z 2018 Memristive fully convolutional networks: an accurate hardware image-segmentor in deep learning, *IEEE Transactions on Emerging Topics in Computational Intelligence* 2 324-334

[17] Wen X 2000 *Design of Application of MEATLAB Neural Network*. (Beijing: Science Press) 200 p