An Approach to Teaching Quality Evaluation Based on ANN and Boosting

Ya-hui ZHAO¹, Xu CUI¹*, Rong-yi CUI¹, Zhe-feng YIN² and Lin-lin ZHAO¹

¹Dept. of Computer Science & Technology, Yanbian University, Yanji, China
²Educational Technology Center of Yanbian University, Yanji, China

*Corresponding author

Keywords: Artificial neural network, Boosting, Teaching quality evaluation, Non-linear dynamic model.

Abstract. An implementation scheme of intelligent evaluator for teaching of university based on artificial neural network (ANN) was proposed. The scheme realized non-linear mapping from factors to results of evaluation, and the current parameters are determined by the data produced in different periods, which makes the dynamic weights can fit to the evaluation of different periods. The evaluator integrates three sub-evaluators using the Boosting method to form the intelligent evaluator and intelligently evaluate the operational rules through the multi-evaluator. It is proved in practice that the method proposed in this paper can reflect the dynamic characteristics and the complex relationship between the factors and the results more effectively than the traditional evaluation methods, and produce better evaluation results.

Introduction

Modern quantitative statistical methods have been used in the evaluation of teaching quality in higher education and have achieved some effect in varying degrees [1,2]. However, these methods still exist a lot of imperfections, it is difficult to determine the weight of the evaluation indicators. Evaluating by experts experience, usually lead to the evaluation of subjective arbitrariness, lack of the objectivity and rationality in evaluation results. The traditional evaluation method is given an evaluation index system, in which each evaluation index is assigned a definite weight to show the importance of the indicator to the outcome [3,4]. The evaluator scores the individual indicators of the evaluator by observation and analysis. The total score is the weighted sum of the scores of each index, and finally, the evaluation result is deduced from the total score according to the pre-determined rule. There are two inherent flaws in this evaluation model: linearity and statics.

The artificial neural network is a non-linear mechanism to realize the mapping function of arbitrary functions and a powerful tool for solving non-linear and dynamic characteristics [5,6]. Boosting is a kind of effective method to improve the accuracy of any given learning algorithm [7,8], can be used to produce accurate evaluation results.

In this paper, we studied the intelligent evaluation method of teaching quality, by introducing the technique of artificial neural network to break the limitation of the above-mentioned two aspects, and used the method of Boosting to integrate multiple sub-evaluators, built the evaluator of high accuracy to reveal the complicated relations between factors and results as well as the dynamic performance of the relationship in given period when training sample set is constantly changing.

Construction of the Training Sample Set in Traditional Evaluation Model

Through investigation, analysis and induction, we built several factors influencing teaching effect, which called index in the evaluation system, and it is called evaluation factors in this article. The evaluation is made up of n factors, and present it as a set:

\[ F = \{f_1, f_2, ..., f_n\} \]  

(1)
Each evaluation vector component has a certain weight to constitute the weight vector:

\[ \mathbf{W} = (w_1, w_2, \ldots, w_n)^T. \]  

(2)

where \( w_i \) is the weight of index \( f_i \) in the evaluation factor set \( F \), indicating the importance of the index. Additionally, the scores of each indexes form the teacher \( j \)'s score vectors:

\[ \mathbf{X}_j = (x_{j1}, x_{j2}, \ldots, x_{jn})^T (j = 1, 2, \ldots, J). \]  

(3)

where \( x_{ji} \) is the score of index \( f_i \) for this example, and \( J \) is the total number of evaluation examples. Due to the traditional evaluation model is a linear evaluation model, the evaluate result of the teacher \( j \) is a linear combination of index scores:

\[ s_j = \mathbf{W}^T \mathbf{X}_j. \]  

(4)

In the traditional evaluation model, teacher \( j \)'s evaluation result \( e_j \) is as follows:

\[ e_j = E(s_j). \]  

(5)

where mapping function \( E \) maps the evaluation score in Eq.4 to the integers in \([0, 3]\) corresponding to the evaluation result "unqualified", "qualified", "good", "excellent".

**Design of Intelligent Evaluator**

**Composition of Intelligent Evaluators**

The evaluator includes three BP neural networks with the same structure, each of which is called a sub-evaluator (see figure 1). The three sub-evaluators work at the same time when evaluating the teaching quality of teachers, and using Boosting coordinated decision rules as the final evaluation result.

![Figure 1. Diagram of ensemble evaluator.](image)

**Sub-evaluator and Its Training Sample**

Each sub-evaluator BP neural network has \( n \) input nodes(corresponding to the evaluation index in \( F \)), \( m \) hidden layer nodes and 4 output nodes. The role of the input layer is to receive the score of each evaluation score of a teacher's teaching quality evaluation. The role of hidden layer is to extract the features of each evaluation factor, reflect the inherent feature of the given evaluation score, and the sigmoid function is used to perform the non-linear transformation. \( M \), the number of hidden nodes, has no theoretical explanation, may be determined according to the particular algorithm or experience [9,10]. Four output nodes corresponding to different evaluation results, and 4 output nodes \( y_i (i=1, 2, 3, 4) \) is processed by following rule to get normalized output \( y'_i \):

\[ y'_i = \frac{\exp(y_i)}{\sum_{k=1}^{4} \exp(y_k)}. \]  

(6)

To example \( j \), target output vector is combined by the normalized output:
\[ y_j^* = (y_{j1}^*, y_{j2}^*, y_{j3}^*, y_{j4}^*)^T. \]

The matching between the target output vector and the evaluation values of the example j shown in Eq.5 is shown in table 1:

<table>
<thead>
<tr>
<th>Artificial evaluation value ( e_j )</th>
<th>Evaluator output vector ( y_j^* )</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>(1 0 0 0)</td>
<td>Excellent</td>
</tr>
<tr>
<td>2</td>
<td>(0 1 0 0)</td>
<td>Good</td>
</tr>
<tr>
<td>1</td>
<td>(0 0 1 0)</td>
<td>Qualified</td>
</tr>
<tr>
<td>0</td>
<td>(0 0 0 1)</td>
<td>Unqualified</td>
</tr>
</tbody>
</table>

The training sample j is:

\[ d_j = \langle X_j, y_j^* \rangle. \]  

where \( X_j \) is the score vector (see Eq.3) of example j, as input of the neural network, each component value of which is evaluation factors' score; \( y_j^* \) is the target output of example j (see Eq.7 and table 1). The training dataset based on samples is:

\[ D = \{ d_1, d_2, \ldots, d_J \}. \]

It is possible to determine whether the output of the evaluator is valid by the following output entropy:

\[ H = -\sum_{i=1}^{4} y_i \log y_i. \]

**Training and Running of Intelligent Evaluator**

The training strategy of the evaluator is as follows:

Step1: Select \( J_A \approx J/3 \) samples randomly from the sample set D to form a sample set \( D_A \), and train sub-evaluator \( E_A \) on sample set \( D_A \); \( E_A \) is only required to be a weak classifier;

Step 2: Generate a random 1-bit binary number \( R \)

1) if \( R=1 \): select samples from \( D \setminus D_A \) randomly to \( E_A \) one after another for testing, when the first wrong sample is encountered, add it to the training set \( D_B \) and repeat this step;

2) if \( R=0 \): add the sample which is classified correctly by \( E_A \) to the training set \( D_B \);

3) stop generating when the \( J_B \approx J/3 \);

Step3: Use \( D_B \) to train sub-evaluator \( E_B \);

Step4: Select the samples in \( (D \setminus D_A) \setminus D_B \) and test with \( E_A \) and \( E_B \): if the results of \( E_A \) and \( E_B \) are different, add the sample to \( D_C \), and use \( D_C \) to train the sub-evaluator \( E_C \).

When evaluating teachers' teaching quality, for the new evaluation example \( x \), the three sub-evaluators run according to the following rules of Boosting and provide the evaluation results:

(1) If the results of \( E_A \) and \( E_B \) are the same, the evaluation of \( x \) follow the result.

(2) If the results of \( E_A \) and \( E_B \) are different, the evaluation of \( x \) follow the result provided by \( E_C \).

**Operation Effect and Analysis**

**Data Collection**

The evaluation system has practically worked and tested for three years, used the scores of end-of-term exams of our university from 2013 to 2015 and the assessment data from students and peers, also collected the samples that the evaluation results are unqualified, as the training data for training the evaluator repeatedly.

The data collected is shown in table 2 as follow:
Table 2. Data sizes in evaluation.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of teachers</th>
<th>Number of courses</th>
<th>Number of students’ evaluation</th>
<th>Number of peers’a evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>987</td>
<td>1324</td>
<td>148920</td>
<td>2961</td>
</tr>
<tr>
<td>2014</td>
<td>1021</td>
<td>1343</td>
<td>149070</td>
<td>3063</td>
</tr>
<tr>
<td>2015</td>
<td>1094</td>
<td>1371</td>
<td>152050</td>
<td>3240</td>
</tr>
<tr>
<td>Total</td>
<td>3102</td>
<td>4038</td>
<td>450040</td>
<td>9264</td>
</tr>
</tbody>
</table>

2/3 of the students’ and peers’ evaluation were used for training, and 1/3 were used for testing.

**Training and Testing**

The training process is as follows:

(1) Constructing the scoring vector (see Eq.3);
(2) Determining the target output values (see Eq.4-5), through the corresponding relationship in table 1, to determine the target output vector;
(3) Preparing training samples (see Eq.8-9);
(4) Training sub-evaluators.

In testing phase, the validity of the output of the sub-evaluator needs to be determined. The testing process is as follows:

(1) Testing the validity of the output of sub-evaluator: if the output entropy (see Eq.10) \( H > 0.5 \), the data is unavailable to the evaluation results, needs to be put into informative data set \( D' \); otherwise enter the decision process;
(2) Decision process: judge the evaluation results according to Boosting rules given in above section.

During the testing phase, putting invalid test data into informative data set \( D' \) and retraining (informative data feedback training) can further improve the judgment ability of the evaluator. With the increase of feedback training frequency of informative data sets, the proportion of invalid output data tends to decrease (as shown in figure 2).

![Figure 2. Effect of training with feedback of informative data.](image_url)

After the training of the informative feedback training method, the test results are shown in table 3:
Table 3. Results of evaluation testing.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of testing data</th>
<th>Percentage of effective output data (%)</th>
<th>Percentage of the results affected by a particular factor in traditional mode (%)</th>
<th>Percentage of the results affected by the specific factors in proposed mode (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>50627</td>
<td>87.1</td>
<td>90.5</td>
<td>10.5</td>
</tr>
<tr>
<td>2014</td>
<td>50711</td>
<td>88.2</td>
<td>90.6</td>
<td>3.6</td>
</tr>
<tr>
<td>2015</td>
<td>51763</td>
<td>79.2</td>
<td>91.4</td>
<td>5.1</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>84.8</td>
<td>90.8</td>
<td>6.4</td>
</tr>
</tbody>
</table>

where "percentage of effective output data" is the percentage of data of three sub-evaluators that are satisfied $H<0.5$ in Eq.10 after the feedback training of the informative data.

The right two columns in the table "percentage of the results affected by a particular factor in traditional mode" and "percentage of the results affected by the specific factors in proposed mode" are used to examine the link between the evaluation result and evaluation factors, in which the evaluation factor is selected as "proper break between classes", and averagely 90.8% of the evaluation results in the traditional way changed when the scores were changed from 0 to 100, and the situation was similar every year. However, only 6.4% of the evaluation results in proposed mode have changed, moreover, the variation amplitude of each year is different, indicating that the influence of this factor on the evaluation result is not very strong in non-linear mode, and it possesses dynamic characteristics.

Conclusion

In this paper, we release the method of teaching quality evaluation, and overcome the defects of traditional evaluation based on linear static model effectively, making the evaluation result more objective and reasonable. The innovation points of this paper are as follows:

1. Using artificial neural network model to established non-linear mapping mechanism from factors to result.
2. Realizing the dynamic adjustment of correlation strength distribution between key factor set and evaluation result.

The future research work need to solve the problems of theory and practice. It mainly includes:

1. To determine the completeness of teaching quality factors, we need to seek reasonable and effective solutions within the education theoretical framework.
2. Using new theories and methods in the field of intelligent information processing, the big data produced by the development and reform of higher education in our country can be mined and processed effectively, and make the intelligent evaluation of teaching quality tends to be perfect.

Acknowledgement

This research was supported by Higher Education Association of Jilin Province, China.

References


