Construction and Optimization of Talent Training Quality Based on Data Mining

Zihua Xu

School of Management, Yangtze Normal University, Chongqing 408100, China

Corresponding Author Email: 20100037@yznu.edu.cn

https://doi.org/10.18280/isi.250403

Received: 8 April 2020
Accepted: 13 July 2020

Keywords: random forest (RF), data mining, talent training quality (TTQ), data reconstruction

ABSTRACT

Thanks to the rapid development of information technology, it is now possible to mine out and fully utilize the valuable information from the big data on talent training quality (TTQ), using artificial intelligence and big data analysis. This paper develops a TTQ prediction model based on random forest (RF)-artificial neural network (ANN), aiming to improve and optimize the TTQ evaluation index system (EIS) in the big data environment, and realize efficient and accurate prediction of TTQ. Firstly, a scientific TTQ EIS was built up, including four primary indices (background evaluation, process evaluation, input evaluation, and result evaluation), and a number of secondary and tertiary indices. Through association matching, a data reconstruction model was constructed for the key evaluation indices of TTQ, the features of the data structure were analyzed, and the key features of each index were extracted and optimized intelligently. Finally, the backpropagation neural network (BPNN) was integrated with the RF algorithm into our prediction model. Experimental results demonstrate the accuracy and effectiveness of the proposed TTQ prediction model. The research findings provide a reference for applying the RF-ANN prediction model in other fields.

1. INTRODUCTION

Talent training quality (TTQ), which determines the employment rate, has always been a social concern. Therefore, most colleges have recently shifted their focus from improving the employment rate of talents to the methods of talent training [1-4]. Traditionally, the TTQ is evaluated against survey data on the expertise and career development of previous graduates. The data are mainly collected by conventional methods like questionnaire survey, interview, seminar, and employer survey. These methods face a variety of inevitable problems: the sample size is generally very small, the collected data are greatly affected by the subjective factors of the respondents, the surveyed items are not sufficiently diverse, and the statistical summary is a heavy work [5-8]. The development of big data technology has made it possible to mine out and fully utilize the valuable information from the massive data related to talent training, namely, occupational psychology test, job information, training plan, social demand, and recruitment, laying a solid basis for TTQ prediction models [9-11].

The prediction results of TTQ directly affect the model of talent training, as well as the philosophy and orientation of college management. The existing studies on TTQ evaluation mainly focus on the construction of evaluation index system (EIS) [11-14]. For example, Bloom [15] set up an EIS for the innovative TTQ in colleges, and compared the EIS with the traditional TTQ EISs, in terms of talent quality features, evaluation principles, subjects, and indices. Luo and Wang [16] combined the context, input, process, and product (CIPP) teaching evaluation model and the data platform into a TTQ EIS for three-year higher vocational colleges. Drawing on the philosophy of outcome-based education (OBE), Mitrofanova et al. [17] reverse engineered the talent training ideas of engineering colleges, refined the graduation requirements on the talents as per the curriculum standards, and finalized a TTQ evaluation plan based on the graduation requirements, curriculum standards, and the matching matrix between knowledge and skill indices.

The TTQ is affected by factors in various aspects. It is difficult to simulate the intricate relationship between these factors by traditional evaluation and prediction models. The difficulty can be overcome by artificial neural networks (ANNs) [18-23], which have been applied in many fields of education and teaching, such as psychological prediction, performance prediction, and employment quality evaluation. For example, Shi and Xu [24] preprocessed the massive data on talents in China with big data analysis techniques, including machine learning and clustering optimization, and sorted out the evaluation indices for the ability of talents in different specialties. Zhu [25] empirically analyzed the TTQ of the previous graduates from the mechatronics department of a college, created a TTQ prediction model for vocational and technical education based on fuzzy neural network (FNN), and improved the prediction accuracy through data supplementation and grayscale whitening of the equipotential dynamic model.

This paper attempts to improve and optimize the TTQ EIS in the big data environment, and realize scientific prediction of TTQ with the aid of advanced techniques like artificial intelligence, and big data analysis. To this end, a TTQ prediction model was established based on random forest (AF)-ANN. Firstly, a scientific TTQ EIS was built up, including four primary indices (background evaluation, process evaluation, input evaluation, and result evaluation), and a number of secondary and tertiary indices. Through association matching, a data reconstruction model was constructed for the
To effectively predict the TTQ, it is essential to build a scientific TTQ EIS in the first place. Referring to the existing studies on employment quality system and talent training model, a four-layer TTQ EIS was developed based on the multi-source data on education and teaching resources, job hunting, and recruitment, which evaluates the TTQ with four primary indices, namely, background evaluation, process evaluation, input evaluation, and result evaluation. The four primary indices reflect the goal construction, implementation process, resource allocation, and outcome/performance of the TTQ, respectively. Each primary index covers a number of secondary and tertiary indices. The hierarchical EIS is detailed as follows:

The first layer (TTQ evaluation goal): \( U = \{ TTQ \text{ evaluation results} \} \);

The second layer (evaluation items): \( U = \{ U_1, U_2, U_3, U_4 \} \) = \{ background evaluation, process evaluation, input evaluation, and result evaluation \} ;

The third layer (secondary indices): \( U_{11} = \{ U_{111}, U_{112}, U_{113} \} \) = \{ social recognition, social impact, government input \} ;

\( U_{21} = \{ U_{211}, U_{212}, U_{213}, U_{214} \} \) = \{ expenditure, course evaluation, college governance, specialty construction \} ;

\( U_{31} = \{ U_{311}, U_{312} \} \) = \{ college facilities, faculty \} ;

\( U_{41} = \{ U_{411}, U_{412}, U_{413} \} \) = \{ achievements, graduate development, social satisfaction \} ;

The fourth layer (tertiary indices): \( U_{11} = \{ U_{111}, U_{112}, U_{113} \} \) = \{ college type, number of students, enrollment \} ;

\( U_{12} = \{ U_{121}, U_{122}, U_{123}, U_{124} \} \) = \{ social service, received awards, radiation effect, quality project \} ;

\( U_{13} = \{ fund \text{ input} \} \) ;

\( U_{14} = \{ U_{141}, U_{142}, U_{143} \} \) = \{ teaching reform expenditure, daily expenditure, faculty construction expenditure \} ;

\( U_{22} = \{ course \text{ resources} \} \) ;

\( U_{23} = \{ U_{231}, U_{232}, U_{233}, U_{234} \} \) = \{ administrative staff, teacher management system, talent management system, teaching quality monitoring \} ;

\( U_{24} = \{ U_{241}, U_{242}, U_{243} \} \) = \{ college-enterprise cooperation, pilot specialties of “1+X certificate” program, specialty construction situation \} ;

\( U_{31} = \{ U_{311}, U_{312}, U_{313}, U_{314} \} \) = \{ infrastructure, teaching facilities, training base, book resources, informationization \} ;

\( U_{32} = \{ U_{321}, U_{322}, U_{323} \} \) = \{ team structure, scientific research level, team building \} ;

\( U_{41} = \{ achievements \} \) ;

\( U_{42} = \{ U_{421}, U_{422} \} \) = \{ employment, knowledge and vocational skills \} ;

\( U_{43} = \{ employer \text{ satisfaction} \} \).

3. DATA RECONSTRUCTION AND FEATURE EXTRACTION OF KEY EVALUATION INDICES

3.1 Data reconstruction

The intelligent extraction of key TTQ evaluation indices is to extract the key features of the evaluation indices with artificial intelligence and cloud data processing. The prerequisites for the intelligent extraction include big data mining and intelligent learning and training.

In this paper, the association matching is adopted to build the data reconstruction model of the key TTQ evaluation indices. During the feature extraction of such indices, an autocorrelation function was defined for each feature point:

\[
r(a, b) = \sum_{t} e \left[ L(a, b) - L(a + \Delta a, b + \Delta b) \right]^2
\]

where, \((\Delta a, \Delta b)\) is the sampling interval of key TTQ evaluation indices; \((a, b)\) is the extreme point of window function \(E\). Then, the feature partitions were accessed frequently to mine out the optimal information from the massive data in the teaching and education resource center, which is the data carrier of the EIS. Through fuzzy C-means (FCM) clustering, the objective function for the classification and query of TTQ evaluation key indicators can be constructed as:

\[
p = \max \left[ \int a(t) \frac{1}{\sqrt{x}} h \left( \frac{t-y}{x} \right) dt \right] = \max \left| E_a(x', y) \right|
\]

\[
W_i > \gamma_2 < W_0
\]

To realize the in-depth mining of the information of key TTQ evaluation indices on each layer, fuzzy feature extraction was adopted to scan the frequent pattern set on each layer of the data reconstruction model.

Inspired by the theory on formal concept analysis (FCA), suppose \( F_1 = (O_1, D_1, R_1) \) and \( F_2 = (O_2, D_2, R_2) \) are the nodes in the concept lattice set of the feature extraction interval of multi-source talent training data. Then, the data structure of a single index accessing to unified education and teaching resource center satisfies \( F_1 \sqsubseteq F_2 \Rightarrow R_1 \sqsubseteq R_2 \). Hence, the big data reconstruction of key TTQ evaluation indices can be expressed as:

\[
l(t) = \frac{1}{\sqrt{T}} \text{rect} \left( \frac{t}{\tau} \right) e^{-1 \frac{2 \pi N^2}{\tau^2} \left( \frac{t}{\tau} \right)^2}
\]

where, \( \text{rect}(t) \) is a rectangular function (\( \text{rect}(t)=1 \) and \( 0.5 \leq t \leq 0.5 \)); \( N = f_{\text{sam}} / f_B \) is the mean sampling frequency; \( f_0 = f_{\text{sam}} / B \) is the initial sampling time.

On this basis, the intelligent extraction algorithm for key TTQ evaluation indices can be developed through the data reconstruction of key TTQ evaluation indices, coupled with data mining algorithm.

3.2 Feature analysis on data structure

Through association rule mining, this paper mines the data structure features of the big data on key TTQ evaluation indices. Let \( c \) be the critical state of the data features in the education and teaching resource center. Then, the big data
distribution of key TTQ evaluation indices can be described by the following mathematical model:

$$c = f(c_i, I)$$  \hspace{1cm}  (4)

where, $I$ is the non-dimensional distributed time series in the standarized data storage model:

$$I = [l_1, l_2, ..., l_n]$$  \hspace{1cm}  (5)

Let $G(x_i, M_i(U))$, $i=1, 2, ..., I$, $j=1, 2, ..., J$, be the cluster centers for data association query. Then, the attribute set of the big data on key TTQ evaluation indices can be obtained as:

$$c_j = f_j(c_i, I, l_j)F_j = (O_j^e, D_j^e, R_j)$$  \hspace{1cm}  (6)

According to the power set of the above formula, the feature extraction return parameter $c_0$ of the big data on key TTQ evaluation indices could be obtained. The corresponding distributed nonlinear time series can be expressed as:

$$l(k, t_0) = 0, k \in [t_0, t_1]$$  \hspace{1cm}  (7)

Through the above analysis, the set of auxiliary information feature set $S = S_0 \cup S_1 \cup S_2$ was obtained for the key TTQ evaluation indices. The parameters in the feature set can be expressed as:

$$S_1 = \{(I, q) \in Q \times Q | l[k + t:1] = q[k + t:1], l_0 \neq q_0\}$$  \hspace{1cm}  (8)

$$S_2 = \{(I, q) \in Q \times Q | l[k + t:1] = q[k + t:1], W([l_1], q[1]) = 1, d_0 = d_0 = 1\}$$  \hspace{1cm}  (9)

### 3.3 Intelligent extraction and optimization of key indices

After the feature analysis on data structure, the deep learning of the multi-source data was carried out through natural language processing, aiming to construct a reasonable intelligent extraction model for key TTQ evaluation indices. Let $\beta$ be the minimum support threshold for the adjustment of user feedback. Then, the minimum support $\text{minSUP}$ can be expressed as:

$$\text{minSUP} = \beta \times |C|$$  \hspace{1cm}  (10)

For any key TTQ evaluation index, the extraction duration $Dur$ can be expressed as:

$$Dur = \max_{j=1}^{J} \left( \sum_{i=1}^{I} \{\text{Ext}(i), \text{Dist}(j)\} \right)$$  \hspace{1cm}  (11)

where, $\text{Ext}(i)$ is the extraction task of the $i$-th key TTQ evaluation index; $\text{Dist}(j)$ is the time needed to assign the extraction task to the $i$-th task processing unit. To improve the adaptability and reduce error of feature extraction, the eigenvalue must meet the following equation:

$$C(k) = 1 + \frac{\omega_0 e^{-\epsilon_k}}{(\gamma_k + U_k)k} = 0$$  \hspace{1cm}  (12)

where, $\omega_0$ is the weight coefficient of learning.

### 4. CONSTRUCTION AND OPTIMIZATION OF RF-ANN TTQ PREDICTION MODEL

#### 4.1 Construction of neural network (NN) model

The talents from different specialties differ in knowledge structure and training plan. To train high-level talents, the key TTQ evaluation indices obtained through selective feature extraction must be classified accurately for TTQ prediction. Here, the BPNN is combined with RF algorithm into a TTQ prediction model (Figure 1).

![Figure 1](image)

**Figure 1.** The workflow of RF-ANN TTQ prediction model

The proposed NN has an input layer, three hidden layers, and one output layer. The dimensions of the three kinds of layers are $1,500, 1,300$, and $1,000$, respectively. Let $\omega_j$ be the connection weight between the $j$-th input layer node and the $k$-th node on the first hidden layer; $x_j$ be the $j$-th eigenvalue of the $i$-th training data. Then, the linear transform $\phi_{2:k}$ of every node on the first hidden layer can be implemented by:

$$\phi_{2:k} = \sum_{j=1}^{j} \omega_{j} \cdot x_{j} + \epsilon_{k}$$  \hspace{1cm}  (13)

where, $\epsilon_k$ is the bias of the $k$-th node on the first hidden layer. The output data of the second and third hidden layers, and the output layer can be designed similarly as formula (13). To enhance the classification ability of the BPNN model for TTQ prediction results, the sigmoid function was adopted as the activation function:

$$\text{Sig}(\phi) = \frac{1}{1 + e^{-\phi}}$$  \hspace{1cm}  (14)

The sigmoid function normalizes the results of linear transform. Let $y_i$ be the results of linear transform of the $i$-th output layer node. Then, the output of the output layer can be obtained through softmax regression:
The results $y_i$ of linear transform of the $i$-th output layer node correspond to the probabilities for the training data of secondary indices to be assigned to one of the four primary indices.

The learning process of the NN relies on the forward transmission of information and the backpropagation of error. For the proposed NN model, the cross-entropy function was adopted as the cost function:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log(g(x_i))$$

(16)

where, $p(x_i)$ is the true distribution of training data; $g(x_i)$ is the probability distribution after BPNN prediction. For a BPNN with N samples, the total loss can be calculated by:

$$Loss = \frac{1}{N} \sum_{i=1}^{N} \Phi(p(x_i), g(x_i))$$

(17)

To minimize the loss, the weight and bias of the BPNN must be continuously adjusted. Then, the output of each node should be calculated at the minimum loss. The connection weight $\omega_{ij}$ and bias $\varepsilon_k$ can be respectively updated by:

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \frac{\alpha}{N} \sum_{i=1}^{N} \frac{\partial Loss}{\partial \omega_{ij}(t)}$$

(18)

$$\varepsilon_k(t+1) = \varepsilon_k(t) - \frac{\alpha}{N} \sum_{i=1}^{N} \frac{\partial Loss}{\partial \varepsilon_k(t)}$$

(19)

As shown in formulas (18) and (19), the connection weight $\omega_{ij}$ and bias $\varepsilon_k$ can be updated iteratively by finding their partial derivatives relative to the loss function of the training data on each secondary index.

4.2 RF optimization of the NN

The TTQ evaluation aims to select the talent with the best overall quality, based on the results of secondary and tertiary indices. The evaluation is a typical classification of unbalanced dataset.

Therefore, it is necessary to reduce the influence of the balance between evaluation indices on different layers over the prediction model, and control the impact of the average voting mechanism on the final classification results. For these purposes, the ranking method was employed to sort and weight the feature importance of indices at different layers, and the F-score algorithm was introduced to compute the weighted harmonic mean of the classifier. Figure 2 explains the workflow of optimizing the BPNN with RF algorithm, which requires data resampling through bootstrapping.

The original training set on TTQ secondary indices, which contains N samples, was subject to automatic Bootstrap aggregation (bagging). Then, training subsets were created by sample extraction with replacement. The probability of a sample in the training set not being extracted can be calculated by:

$$P_{out} = \left(1 - \frac{1}{N}\right)^N$$

(20)

If the sample size tends to infinity, the misalignment rate $OOB_{error}$ of the out-of-bag RF samples for error estimation and feature importance analysis is about 37%. In this paper, the feature importance of an index is computed based on the importance of its data samples (formula 12). The samples with relatively high importance were selected through the following steps:

Step 1. Add a random nose to the eigenvalue $C$ of the data samples of evaluation indices, create an out-of-bag data sample set, and record the out-of-bag error as $OOB_{error}$. Suppose there are $M$ decision trees in the RF. Then, the importance of eigenvalue $C$ can be computed by:

$$I_c = \sum_{i=1}^{M} \left( OOB_{error} - OOB_{error}' \right)$$

(21)

Step 2. Sort the training data on TTQ secondary indices by importance, select the top 70% features, and eliminate the last 30% features.
Step 3. Repeat Steps 1 and 2 until the total number of eigenvalues reaches the preset value, forming the final feature set of TTQ secondary indices.

The F-score algorithm was employed to compute the weighted harmonic average of Precision and Recall. The Precision, Recall, and Accuracy can be respectively calculated by:

\[ PRE = \frac{N_H}{N_H + N_{\text{error}1}} \]  
\[ REC = \frac{N_H}{N_H + N_{\text{error}2}} \]  
\[ ACC = \frac{N_H + N_L}{N_H + N_{\text{error}1} + N_H + N_{\text{error}2}} \]

where, \( N_H \) is the number of correct predictions of high-quality talents; \( N_H \) is the number of correct predictions of low-quality talents; \( N_{\text{error}1} \) is the number of incorrect predictions of high-quality talents; \( N_{\text{error}2} \) is the number of incorrect predictions of low-quality talents.

The F-score of the decision trees in the classifier can be calculated by:

\[ F = \frac{2 \times REC \times PRE}{REC + PRE} \]

For verification, the test set on the TTQ secondary indices was imported to each decision tree. Then, the predicted classes of the decision trees were compared with the actual classes.

5. EXPERIMENTS AND RESULTS ANALYSIS

Two 200 decision trees were set up for the RF-ANN TTQ prediction model. The BPNN parameters were configured as follows: the adjustment speed of connection weight was set to 0.01, and the maximum number of iterations to 1,200.

Figure 3 shows the RF ranking of TTQ secondary indices by feature importance. The last 30% of features in the ranking were removed, forming feature subsets. Each subset was imported to the BPNN prediction model.

Figure 4 provides the mean absolute error (MAE) between the prediction outputted from each feature subset and the actual value. It can be seen that the MAE first decreased to the minimum and then gradually increased. This means the prediction and classification performance of our model can be improved by removing redundant features, but suppressed by eliminating important features. Therefore, the proposed model can effectively extract features intelligently and recognize their importance.

The established TTQ EIS consists of 4 primary indices, 12 secondary indices, and 31 tertiary indices. The key features of each index were extracted intelligently. To verify the RF-optimization of the NN, the redundant features were removed recursively by the RF until the required number of features was reached. The cross-validation score of the optimized NN is displayed in Figure 5.

The established BPNN has three hidden layers. The number of nodes in each hidden layer directly bears on the distributed expression of the index features. Besides, the valuable information might be lost after noise addition. To reveal the effect of the number of hidden layer nodes on the prediction and classification accuracy of our model, the number of hidden layer nodes in the BPNN was adjusted as shown in Table 1. According to the prediction and classification results, the performance of our model was not desirable, when the hidden layers had relatively few nodes. This is manifested by the accuracies in Lines 2-5 of Table 1.

---

**Figure 3.** The RF ranking of TTQ secondary indices by feature importance

**Figure 4.** The MAE trend of our model

**Figure 5.** The cross-validation score of the RF-optimization of the NN
Table 1. The prediction and classification results at different number of hidden layer nodes

<table>
<thead>
<tr>
<th>Number of hidden layer nodes</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
<th>Accuracy</th>
<th>TP</th>
<th>FN</th>
<th>FP</th>
<th>TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>{15, 15, 20}</td>
<td>38.89%</td>
<td>27.93%</td>
<td>31.11%</td>
<td>96.78%</td>
<td>7</td>
<td>21</td>
<td>10</td>
<td>457</td>
</tr>
<tr>
<td>{18, 18, 15}</td>
<td>41.18%</td>
<td>27.93%</td>
<td>31.82%</td>
<td>95.99%</td>
<td>5</td>
<td>19</td>
<td>11</td>
<td>464</td>
</tr>
<tr>
<td>{12, 12, 15}</td>
<td>87.26%</td>
<td>35.41%</td>
<td>73.47%</td>
<td>92.19%</td>
<td>3</td>
<td>23</td>
<td>2</td>
<td>474</td>
</tr>
<tr>
<td>{15, 18, 12}</td>
<td>33.53%</td>
<td>61.11%</td>
<td>31.82%</td>
<td>93.99%</td>
<td>2</td>
<td>25</td>
<td>5</td>
<td>462</td>
</tr>
<tr>
<td>{18, 15, 12}</td>
<td>34.33%</td>
<td>36.41%</td>
<td>31.11%</td>
<td>93.19%</td>
<td>2</td>
<td>24</td>
<td>5</td>
<td>463</td>
</tr>
<tr>
<td>{12, 12, 15}</td>
<td>78.43%</td>
<td>17.52%</td>
<td>19.41%</td>
<td>97.19%</td>
<td>4</td>
<td>21</td>
<td>3</td>
<td>478</td>
</tr>
<tr>
<td>{15, 28, 12}</td>
<td>54.24%</td>
<td>36.81%</td>
<td>28.86%</td>
<td>96.96%</td>
<td>6</td>
<td>22</td>
<td>3</td>
<td>458</td>
</tr>
<tr>
<td>{18, 25, 15}</td>
<td>87.50%</td>
<td>14.81%</td>
<td>28.86%</td>
<td>96.51%</td>
<td>5</td>
<td>24</td>
<td>5</td>
<td>462</td>
</tr>
<tr>
<td>{18, 28, 15}</td>
<td>34.25%</td>
<td>19.52%</td>
<td>27.26%</td>
<td>97.31%</td>
<td>6</td>
<td>23</td>
<td>10</td>
<td>453</td>
</tr>
<tr>
<td>{15, 20, 25}</td>
<td>75.86%</td>
<td>25.32%</td>
<td>28.27%</td>
<td>97.14%</td>
<td>7</td>
<td>20</td>
<td>7</td>
<td>462</td>
</tr>
</tbody>
</table>

Note: TP, FN, FP, and FN are short for true positive, false negative, false positive, and false negative, respectively.

Table 2 compares the prediction and classification effects of our model with the same four models. The results show that the accuracy of our model surpassed 98% on average. Overall, our model achieved high results on Precision, Recall, Accuracy, and F-score, revealing its effectiveness and advantage in TTQ prediction and classification.

Table 2. The comparison of prediction and classification effects

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Precision</th>
<th>Recall</th>
<th>Accuracy</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>92.35%</td>
<td>97.80%</td>
<td>97.04%</td>
<td>96.44%</td>
</tr>
<tr>
<td>3-layer BPNN</td>
<td>94.42%</td>
<td>86.13%</td>
<td>90.36%</td>
<td>87.78%</td>
</tr>
<tr>
<td>SVM</td>
<td>96.62%</td>
<td>92.80%</td>
<td>96.20%</td>
<td>94.67%</td>
</tr>
<tr>
<td>CART</td>
<td>97.39%</td>
<td>95.39%</td>
<td>94.39%</td>
<td>95.33%</td>
</tr>
<tr>
<td>NBC</td>
<td>83.24%</td>
<td>87.14%</td>
<td>85.28%</td>
<td>86.67%</td>
</tr>
<tr>
<td>Our method</td>
<td>98.81%</td>
<td>98.05%</td>
<td>98.40%</td>
<td>97.89%</td>
</tr>
</tbody>
</table>

6. CONCLUSIONS

This paper mainly puts forward an RF-ANN TTQ prediction model. Firstly, a scientific four-layer hierarchical EIS was established, involving 4 primary indices and numerous secondary and tertiary indices. Through association matching, a data reconstruction model was constructed by key TTQ evaluation indices, followed by the feature analysis of data structure, as well as intelligent extraction and optimization of key features. Experimental results demonstrate that our model can extract key features of evaluation indices accurately, and achieve high robustness, reliability, and generalization ability. To predict TTQ accurately and efficiently, the BPNN was combined with RF algorithm into our prediction model. The proposed model was found to greatly outperform the traditional prediction models through experiments.

ACKNOWLEDGEMENT

This paper was Supported by Program for Chongqing Scholars and Innovative Research Team in University.

REFERENCES


