[1]Liqin He
[2,*]Chaojie Hu
[3]Ling Nie
[4]Chunxia Li
[5]Honglian Liu

Design of English Teaching Capability Evaluation Model Under Big Data Analysis

Journal of Electrical Systems

Abstract: - Traditional methods of evaluating English teaching capability involve a considerable degree of subjective human judgment, leading to classification errors in big data information. To improve the comprehensiveness and accuracy of English teaching capability evaluation, it is necessary to construct a corresponding evaluation model based on big data. This paper employs the k-means clustering analysis algorithm to devise a system structure design for English teaching capability, applies constrained parameter big data structure analysis, and proposes utilizing a quantitative recursive approach to evaluate the teaching capabilities of big data information models based on cluster analysis. The simulation results demonstrate that the method designed in this paper can enhance the comprehensiveness and precision of English teaching capability evaluation, thereby contributing to the advancement of information fusion analysis capabilities.

Keywords: Big Data Analysis, English Teaching, Teaching Capability Evaluation, Quantitative Recursive Analysis, K-means Clustering Analysis Algorithm.

I. INTRODUCTION

In the context of the ongoing advancement of internet big data analysis, data mining and data analysis technologies are extensively utilized across diverse domains, especially in corporate marketing activities. Internet companies, by analyzing consumers' psychology and habits, can more accurately target market demands. Currently, as the depth of English teaching intensifies, assessing English teaching capability has garnered growing significance. When assessing English teaching capability, the core indicators that underpin proficiency are the capacity to store and effectively apply English vocabulary, meaning that vocabulary usage can explain the capacity for language learning. Manual evaluation, as a traditional method of assessing English teaching capability, has the advantage of a flexible evaluation process. However, due to the presence of subjective human judgment, it might lead to significant evaluation errors. When assessing English teaching capability, it is imperative to establish standardized and consistent evaluation criteria. Due to various constraints affecting measurement of English teaching capability levels, constructing constrained parameter models and big data analysis models is beneficial for improving the predictive capacity and effectiveness of the evaluation. Consequently, this research endeavors to establish an evaluation model for English teaching capability, drawing upon the sophisticated techniques of big data clustering analysis and information fusion. To substantiate the efficacy of the proposed methodologies, rigorous simulation experiments have been conducted.

II. ENGLISH TEACHING CAPABILITY EVALUATION SYSTEM STRUCTURE DESIGN

A. Improvement Ideas for K-means Clustering Analysis Algorithm

The K-means clustering analysis algorithm is included in data mining clustering analysis techniques. The instability of the algorithm, as well as the significant impact of outliers and noise, primarily stem from the technique employed for selecting the center. In order to tackle these challenges, this paper introduces an innovative approach for selecting initial clustering centers. Specifically, it involves classifying the extracted S sample sets using hierarchical agglomerative clustering technology, ensuring that the classes of clustering centers are inconsistent, and that each class's sample set is representative. The new method for selecting initial clustering centers is as follows: first, extract S sample sets from the original data, then hierarchically agglomerate and classify each sample set as k′ classes, remove classes with ≤1 members to eliminate outliers, thus ensuring the representativeness of

[1] Jiangxi College of Traditional Chinese Medicine, Fuzhou 344000, Jiangxi, China
[2] Guangzhou Huashang College, Guangzhou 511300, Guangdong, China
[3] Jiangxi College of Traditional Chinese Medicine, Fuzhou 344000, Jiangxi, China
[4] Jiangxi College of Traditional Chinese Medicine, Fuzhou 344000, Jiangxi, China
[5] Jiangxi College of Traditional Chinese Medicine, Fuzhou 344000, Jiangxi, China
*Corresponding author: Chaojie Hu
Copyright © JES 2024 on-line : journal.esrgroups.org

clustering centers; next, agglomerate the sample sets into k classes, with clustering centers also being K; finally, re-agglomerate the S ×k′ clustering centers generated from the S samples, with the optimal initial class center being K. The advantage of this method is that it proposes an automatic selection method for initial clustering centers, which selects initial clustering points with high representativeness, thus improving the efficiency of the entire dataset partitioning. [1-2]

B. Design of the English Teaching Capability System Structure

The evaluation of English teaching capability also involves teachers' assessment of students. As internet technology continues to rapidly evolve, the methods students utilize to access the internet have grown progressively simpler and more convenient. In the English teaching system, the learning tasks arranged for students include traditional text reading as well as online discussions, live program design development, etc. Teachers can set their teaching goals in advance, and students can also obtain their tasks through the system. The system will timely collect and categorize data on students' task completion and interaction system, providing firsthand data for data mining algorithms. Both teachers and students can receive feedback from the system periodically for timely corrections. [3]

As shown in Figure 1, the system theoretically designs a dynamic English teaching capability evaluation model. The input layer includes many evaluation parameters, with multiple interactive behavior data constituting the process parameters, specifically including students' task completion time, number of modifications, usage time, etc. The evaluation content of task parameters is the time and difficulty level of the tasks completed by the students. The teacher's assessment specifically refers to the grading of students after task completion. When calculating the comprehensive score, it is necessary to select methods and evaluation index weight libraries, with the method being the predefined evaluation algorithm, thus obtaining the final evaluation results. Teachers and students are now able to receive statistical analysis results and improvement suggestions transmitted by the system, facilitating teachers in enhancing their teaching methods and further developing their educational capabilities. Simultaneously, the system's next round of new evaluations will begin, forming a virtuous cycle.

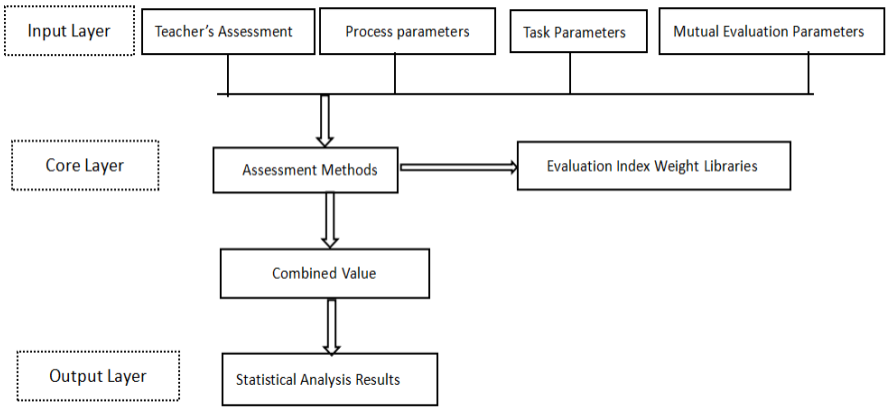

Figure 1: English Teaching Capability Evaluation Process

III. English Teaching Capability Evaluation Model Based on Big Data Analysis

A. In-depth Big Data Analysis of Restrictive Factors in Assessing English Teaching Capability

Enhancing the effectiveness of English teaching capability evaluation can be accomplished by developing an information sampling model that is grounded in the constraining parameters specific to English teaching abilities. The methods for analyzing English teaching capability are information fusion and time series analysis, where the former is nonlinear. The constraint parameter indicators of English teaching capability constitute a set of nonlinear time series, characterized by a high-dimensional distribution space across various parameter metrics. The indicators of English teaching capability are diverse and comprehensive, encompassing factors such as the breadth of English vocabulary, proficiency in oral communication, and the level of investment in teaching facilities, to name a few. The differential equation of the information flow model for the constraint parameters of English teaching capability is as follows: [4-5]

$$x_n = x(t_0 + n\Delta t) = h[z(t_0 + n\Delta t)] + \omega_n \tag{1}$$

In equation (1), $h[z(t_0 + n\Delta t)]$ and ω_n represent the multivariate value function and error residual function of English teaching capability evaluation, respectively. Due to the high-dimensional feature distribution space of

parameter indicators, the method for calculating the evaluation solution vector is the relevance information fusion method, thus obtaining the characteristic training subset Si, which satisfies the following conditions: [6]

$$\begin{cases} \sum_{L} = diag(\delta_1, \delta_2, ..., \delta_t), \delta_i = \sqrt{\lambda_i}, \forall i \neq j \\ US_i = V - v_s \end{cases}$$

Assuming $x_{n+1}= \mu x_n(1-x_n)$ as the conjugate solution of the statistical information model for English teaching capability evaluation, then the conjugate solution should meet the condition of characteristic decomposition for initial values, expressed as:

$$U = \{u(t) | u(t) \in X, \|u\| \leq d, t \in I\} \tag{2}$$

In equation (2), the number set(Ii) $i \in N=\{x_1, x_2, \cdots, x_m\}$. Given the multivariate nature of variables within the English teaching capability evaluation system, where the feature distribution sequence is represented as x(n), the computational formula for assessing measurement values within the information flow framework for evaluating English teaching capability, relying on prior statistical measurement values, is as follows:

$$\begin{gathered} c_{1x}(\tau) = E\{x(n)\} = 0 \\ c_{2x}(\tau) = E\{x(n)x(n+\tau)\} = r(\tau) \\ c_{kx}(\tau_1, \tau_2, ..., \tau_{k-1}) \equiv 0, k \geq 3 \end{gathered} \tag{3}$$

When the distribution condition of English teaching capability levels is a three-dimensional continuous functional, it indicates that the assessment framework for English teaching capability yields a convergent outcome, with its condition constraint equation being: [7]

$$\psi_x(\omega) = \ln \Phi_x(\omega) = -\frac{1}{2}\omega^2\sigma^2 \tag{4}$$

The Leveraging the established information flow model for English teaching capability data, the sequential arrangement of the big data distribution model precisely derives the anticipated value of English teaching proficiency.

B. Quantitative Recursive Analysis of Teaching Capability Evaluation

The approach utilized for analyzing the data information model pertaining to English teaching capability evaluation involves quantitative recursive analysis. Consequently, the control objective function of the English teaching capability prediction model is formulated as follows: [8]

$$\max_{x_{a,b,d,p}} \sum_{a \in A} \sum_{b \in B} \sum_{d \in D} \sum_{p \in P} x_{a,b,d,p} V_p \tag{5}$$

$$s.t. \quad \sum_{a \in A} \sum_{d \in D} \sum_{p \in P} x_{a,b,d,p} R_P^{bw} \leq K_b^{bw}(S), b \in B \tag{6}$$

The method employed for the quantitative recursive evaluation of English teaching capability utilizes the grey model, assuming $\{x_i\}_{i=1}^N$ represents the historical distribution data of English teaching capability levels. Assuming that the disturbance characteristic value remains constant within the model, the probabilistic density functional formula utilized for evaluating English teaching capability can be expressed as follows:

$$u_c(t) = Kx_c(t) \tag{7}$$

In the high-dimensional feature distribution space, the continuous function used to evaluate English teaching capability is: $u : I \times IR^d \rightarrow IR$, which, after K-1 ($k \geqslant 1$) iterations, with N(k)<L, is suitable for obtaining the grey sequence of English teaching capability evaluation. By conducting quantitative recursive analysis on the output indicators of the English teaching capability evaluation model, K nearest sample values of the big data information flow are obtained: [9-10]

$$P_{1J} = \sum_{d_i \in KNN} Sim(x_i, d_j) y(d_i, c_j) \tag{8}$$

Employing the method of big data information fusion, formulate the objective function for clustering data pertaining to English teaching capability:

$$J_m(U,V)=\sum_{k=1}^{n}\sum_{i=1}^{c}\mu_{ik}^{m}(d_{ik})^2 \tag{9}$$

To obtain quantitative recursive feature results, a rigorous quantitative analysis is conducted on the correlation distribution sequence pertaining to the evaluation of English teaching capability:

$$x_n=a_0+\sum_{i=1}^{M_{AR}}a_i x_{n-i}+\sum_{j=0}^{M_{MA}}b_j\eta_{n-j} \tag{10}$$

In the formulas above, a_0 represents the sampling amplitude of the evaluation variable, x_{n-i} represents the time series, b_j and represents the oscillation decay value of the evaluation numbers.

C. Application Analysis of English Teaching Capability Evaluation Model

For the constructed constraint parameter indicator analysis model, this paper introduces an evaluation approach that integrates big data fuzzy K-means clustering with information fusion techniques, leveraging the quantitative recursive analysis method for enhanced accuracy and comprehensiveness. Thus, English teaching capability evaluation can be seen as solving for the clustering objective function, i.e., solving for the estimated value of the constraint vector of the English teaching capability evaluation model. When is minimized, the entropy characteristic value of the English teaching capability constraint feature information is: [11]

$$P_{loss}=1-\frac{1-p_0}{\rho}=\frac{p_0+\rho-1}{\rho}=\sum_{n=1}^{N}p_{K,n} \tag{11}$$

Taking into account the evaluation disturbance item of English teaching capability, the English teaching capability evaluation model is simplified to the following least squares equation:

$$z(t)=x(t)+iy(t)=a(t)e^{i\theta(t)}+n(t) \tag{12}$$

In Equation (12), x(t) represents the actual component of the temporal sequence of data distribution, whereas y(t) corresponds to the imaginary component of the series indicating constraints.

The method for random amplitude processing of teaching ability is the surrogate data method, obtaining $x'(k)$, based on the empirical distribution data of English teaching capability, the utilization rate of English teaching resource distribution is: [12]

$$U_{util}=\gamma\overline{X} \tag{13}$$

By establishing a hierarchical tree for the English teaching capability evaluation model and analyzing its principal component features within the context of big data, the similarity of English receptiveness distribution can be determined using the fuzzy proximity fill method. The formula for this method is as follows:

$$Sim_1(d_i,d_{1j})=\frac{\sum_{k=1}^{M}W_{ik}\times W_{1jk}}{\sqrt{\sum_{k=1}^{M}W_{ik}^2}\cdot\sqrt{\sum_{k=1}^{M}W_{1jk}^2}} \tag{14}$$

In Equation (14), the symbol d_{ij} denotes the index associated with the vector of the K-means clustering center, whereas d_i signifies the feature vector representative of the prior distribution within the evaluation framework. Through the clustering and integration of evaluation indicator parameters, an expression pertaining to the integration of teaching resources can be derived by employing the linear correlation feature fusion approach:

$$P(w|x)=P(x|w)/P(x) \tag{15}$$

From equation (15), when the quantitative recursive feature satisfies $N(i)\bmod L<m$, the distribution probability characteristic of teaching ability is: $p(i)=\left[\frac{N(i)}{L}\right]$, dividing the data flow $X(i)$ during the English teaching capability evaluation process into submatrices X_{ij} , with the size of submatrices being $N_{ij}\times m$, and there are $p(i)$ submatrices. By reclustering and integrating the indicator parameters within the English teaching capability evaluation model, and subsequently adjusting the corresponding English teaching methods, it is possible to enhance the overall English teaching capability. [13]

IV. SIMULATION EXPERIMENT

This paper evaluates the effectiveness of the English teaching capability evaluation method utilizing big data analysis through the design and conduct of simulation experiments. For the simulation experiment, six distinct types of English class content were chosen, and a comparative analysis was conducted between the evaluation effects of traditional English teaching capability evaluation methods and the method proposed in this paper, thereby ensuring the validity of the experiment. [14].

A. Parameter Setting

The purpose of setting parameters in this paper is to guarantee the efficacy of the proposed method. With reference data $|Y_x|$ outside the value range [1000, 1250], D_p, D_D, S_j, F_i are set to 40×10^4, 12.8, 800, 1260, respectively. In order to maintain consistency in the variables being compared throughout the experiment, the paper established the relevant experimental data as presented in Table 1.

Table 1: Experimental Data

Experiment Number	1	2	3	4	5	6
English Teaching Duration/min	20	35	40	45	50	50
English Teaching Content	English Grammar	English Vocabulary Explanation	Sentence Analysis	Oral Practice	Writing Practice	English Dialogue

B. Experiment Validity Test

The validity of the experiment was ensured through calculation by SPSS, where the validity value must be greater than 0.95. Substituting this validity value yields:

$$Y(k)=\frac{y[u(k-1)]}{1+y^2(k-1)}=\frac{1\times[9.7\times(5-1)]}{1+1\times(5\text{-}1)}\geq 0.95$$

This indicates the validity of the experiment.

C. Experiment Data Error Correction

Since the comparison of teaching ability evaluation involves both models and methods, there is a significant error. To uphold the validity of the experimental data, it is imperative that the coefficient substitution ratio between the measured values and the true values remains within a reasonable range. With the change in the number of experiments, the ratio of coefficient substitution also changes, but its value range always remains [2.45, 6.55]. Extreme value determination can avoid non-zero errors, i.e., the method chosen for setting zero point values is the time truncation method. [15]

D. Comparative Analysis of Results

From Table 2, it is evident that the accuracy and utilization rates of English teaching capability evaluated by traditional methods are lower than those evaluated by the method designed in this paper. As depicted in Figure 2, the evaluation refusal values pertaining to the method introduced in this study consistently surpass 7500, thereby demonstrating the enhanced accuracy of the English teaching capability evaluation model developed through the integration of big data information fusion and fuzzy clustering analysis.

Table 2: Evaluation of Effectiveness Comparison

Evaluation Cycle	Traditional Method		Proposed Method in This Paper	
	Accuracy (%)	Utilization (%)	Accuracy (%)	Utilization (%)
1	85.61	88.65	98.51	98.24
2	84.16	88.13	97.36	97.96
3	87.43	80.37	96.82	99.35
4	89.54	79.16	98.83	96.84

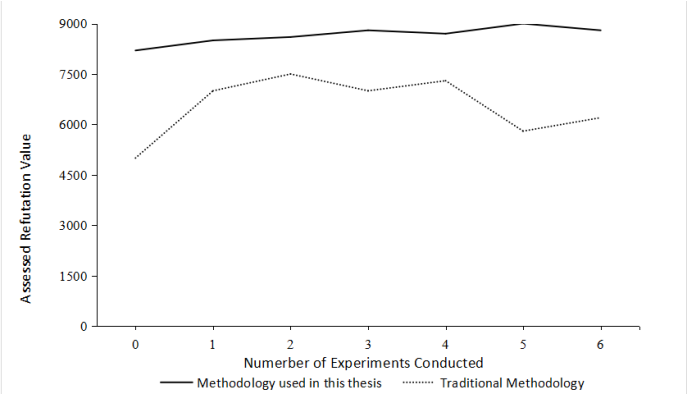

Figure 2: Experimental Comparison Results

V. CONCLUSION

Due to the many limitations of traditional manual English teaching capability evaluations, this paper introduces an enhanced evaluation model that integrates big data information fusion with cluster analysis for a more comprehensive assessment. Among the critical indicators for assessing English teaching capability, the proficiency in storing English vocabulary and the capability to apply it effectively stand out as the most significant.. Therefore, based on these indicators, this paper tested and compared the evaluation effects of the proposed method with traditional methods through simulation experiments. The findings indicate that the optimized evaluation method employed in this study is beneficial for enhancing the comprehensiveness and precision of the assessment of English teaching capability.

ACKNOWLEDGMENT

This work was supported by the Research Project of Vocational Education Teaching Reform in Jiangxi Province “Research and Practice on Improving Digital Teaching Ability of English Teachers in Higher Vocational Colleges in the Context of” the Three Educations Reform” (Project Number: JXJG-23-44-3).

REFERENCES

[1] Jin X. Evaluation Model of Higher Vocational Teachers’ Professional Practice Ability Based on AHP-Neural Network. Journal of Beijing Polytechnic College, 2017.16(4): 38-42.
[2] Jiao-Rong W U, Lin Z Y, Hui Y. The Undergraduate Practical Teaching Evaluation Model and Application Based on Ability Training. Education Teaching Forum, 2017(7): 45-49.
[3] Wang F. An analysis of college English writing teaching reform based on the era of big data. Journal of Jiamusi Vocational Institute, 2017(2): 211-212.
[4] Li Ying. Research on the Application of Visualized Big Data Analysis Model in Personalized English Teaching. Campus English, 2017(10): 40-41.
[5] Ge Guoliang. Research on English Teaching Ability Evaluation Algorithm Based on Big Data Fuzzy K-Means Clustering. Modern Electronics Technique, 2017, 40(20): 31-33.
[6] Wang Xuemei. Research on the Employment Evaluation Model of College Graduates Based on Hierarchical Analysis Method in the Context of Big Data. Times Education, 2016(6): 20-21.
[7] Song Yujuan. Discussion on the Development Path of English Teaching in the Era of Big Data. Business Manager, 2016(30).
[8] Zhang Huidan. English Ubiquitous Learning Ecosystem Based on Big Data Analysis. Teaching and Management, 2017(9): 107-109.
[9] Sun Yan. English Teaching Ability Evaluation Based on Feedback Model. Modern Electronic Technique, 2017, 40(20): 25-27.
[10] He Mengyi. Prospects for the Application of Big Data Analysis in College English Teaching Evaluation. Archives, 2017(29): 191-191.
[11] Liu Dong, Wang Lifeng, Zhang Huaifeng. Design of Statistical Analysis Model Based on Big Data. Software Guide, 2016, 15(7): 28-30.
[12] Yin Minhong. Analysis of English Teaching Level Evaluation Model. Journal of Heilongjiang Ecological Engineering Vocational College, 2012(3): 119-120.
[13] Nie Ying. Research on Self-Improvement Ways of College English Teachers’ Classroom Teaching Ability. Contemporary Educational Practice and Teaching Research: Electronic Edition, 2017(4): 87-87.
[14] He Sheng. Research and Application of Neural Network Genetic Algorithm in the Employability Evaluation of Vocational College Students . Information Systems Engineering, 2016(8): 143-144.
[15] Xu Liqing. Research on the College Teaching Quality Evaluation System Based on Data Mining. Information Technology and Informatization, 2016(8):79-82.