ISSN 2522-6398 Vol. 6, Issue 2: 57-62, DOI: 10.25236/FER.2023.060210

The Evaluation of College English Writing Teaching Ability in English Teaching Courses Based on Big Data

Zhurong Liu

School of International Tourism and Culture, Guizhou Normal University, Guiyang, 550025, China
93984012@qq.com

Abstract: *Under the background of big data information fusion, the use of stepwise regression analysis method to evaluate the English writing teaching ability of college teachers usually results in inaccurate classification of data flow and great errors of constraint index parameters in information fusion. In this paper, k-means fuzzy clustering algorithm is proposed. By constructing parameter constraint index system and big data analysis model, various nonlinear English teaching parameters are clustered and integrated to achieve the extraction of entropy characteristics of English constraint information. On this basis, the recursive quantitative analysis method is used to intuitively analyze various parameters of English teaching ability and the law of data change in time series. Finally, the correlation of constraint parameters and disturbance data of various English teaching parameters is analyzed by Matlab simulation algorithm. Research shows that the use of k-means algorithm to carry out information fusion of English writing teaching ability data can further improve the efficiency and accuracy of English teaching ability assessment of teachers.*

Keywords: *Big Data Information Fusion; K-means Clustering Algorithm; College English Writing; Teaching Ability Assessment; English Teaching Courses*

1. Introduction

Using information processing technology and big data analysis technology for teaching evaluation and resource information scheduling has positive and important significance to improve the quantitative management and planning ability of teaching process. To this end, this paper studies the assessment of English writing teaching ability based on big data analysis. Because the English writing teaching ability evaluation is more restrictive factors, quantitative test and analysis on the need for English teaching level, constructs a model of parameter constraint English writing teaching level and large data analysis model, using big data information fusion and clustering approach to teaching English writing capability assessment, build teaching ability evaluation of the objective function and statistical analysis model, quantitative prediction ability to improve English teaching ability evaluation. This paper presents a fuzzy k-means clustering based on large data and information fusion estimation methods, the teaching of English writing ability to implement English teaching ability index parameter clustering and integration, prepare the corresponding teaching resource allocation plan, implement teaching English writing ability evaluation quantitative planning, realize the teaching of English writing ability to accurately assess [1].

2. Methods and Analysis

2.1 The Implementation Principle of K-Average Clustering Algorithm for College English Writing Teaching Ability Evaluation

The evaluation of college English teachers' writing teaching ability is a cluster analysis of factors such as professional accomplishment, teaching resource organization ability, classroom teaching ability, scientific research and innovation ability. After the weighted average of the teaching ability of the above aspects, the normalized mapping processing of different data is completed, and the data information is used as the data object calculated by the K-means clustering algorithm [2]. According to the individual similarity of different data, the samples of English teachers' writing ability are divided into k categories, and the center positions of k cluster arrays are determined. Then calculate the fitness

ISSN 2522-6398 Vol. 6, Issue 2: 57-62, DOI: 10.25236/FER.2023.060210

of each particle in different clusters, evaluate the distance of different particles to the center of their own clusters [3] [4], and record the maximum fitness of all particles. The maximum position of the particles can be recorded as Gbest. Then, iteratively operates on multiple cluster arrays until the Gbest position of each particle is the optimal position it has experienced, that is, the end of the algorithm when the particle's iterative update is optimal [5].

2.2 The application of k-means clustering algorithm in English writing teaching ability assessment

2.2.1 The systematic framework of English writing teaching ability evaluation based on k-means clustering

According to the collected data samples of English writing teaching ability, the main process of k-means clustering using Matlab simulation software is (as shown in figure 1) :

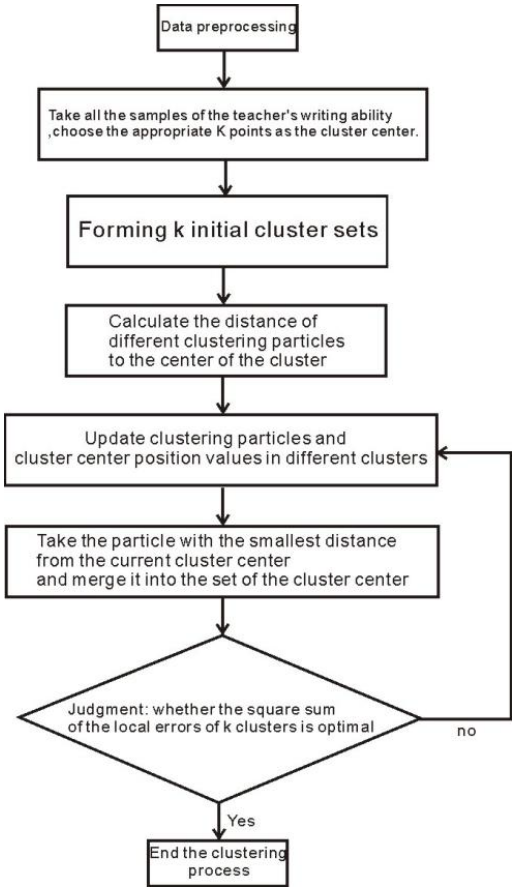

Figure 1: The process of k-means clustering of English teaching elements by Matlab simulation software

According to the above the mean K-mean clustering algorithm implementation process, can draw ability of English writing teaching class cluster data, will not be in accordance with a certain behavior path for successive iterations, but through the clustering center around the particles, into the distance of its recent clustering groups, to achieve professional ability of organization, teaching resources, classroom teaching, scientific research and innovation ability factor of iteration and the optimal clustering[6].

2.3 Construction of K-means Clustering Analysis Model for English Writing Teaching Ability Evaluation

2.3.1 Information Sampling Model for Evaluating Constrained Parameters of English Writing Teaching Ability

Using the time series and information fusion analysis method, the constraint parameter distribution model of English writing teaching ability evaluation is constructed to complete the statistical analysis of teachers' professional quality, teaching resource organization ability, classroom teaching ability,

ISSN 2522-6398 Vol. 6, Issue 2: 57-62, DOI: 10.25236/FER.2023.060210

scientific research and innovation ability [4]. The English teaching ability constraint parameter information flow model constructed by the differential equation expression with high dimensional feature distribution is [7]:

$$x_n=x(t_0+n\triangle t)=h[z(t_0+n\triangle t)]+\omega_n \quad (1)$$

(x: time series of evaluation values of English teachers' writing teaching ability; h: multi-window function of teacher's teaching ability evaluation;ω_n: constrained parameter error; t: time).

Assume that the English teaching ability evaluates the information model of the constraint parameters. There are $x_n+1=\mu x_n(1-x_n)$ common solutions. The common solution can perform feature decomposition on the initial values of k cluster arrays, and the decomposition condition is

$$U=\{u(t)|u(t)\in X,||u||\leq d,(t\in I,I_i=(x_1,x_2,\ldots,x_n))\} \quad (2)$$

In the process of analyzing the parameters of various English writing teaching ability and analyzing the statistical characteristics, it is necessary to construct a data flow model of English writing teaching ability evaluation based on its distribution sequence x(n). The model provides a more accurate time series for the evaluation of English writing teaching ability:

$$C_1x(\tau)=E\{x(n)\}=0 \quad C_2x(\tau)=E\{x(n)x(n+\tau)\}=r(\tau) \quad C_kx(\tau1,\tau2\ldots,\tau k-1)\equiv0,k\geq3 \quad (3)$$

2.3.2 Quantitative Recursive Analysis of English Writing Teaching Ability Evaluation

Using recursive quantitative analysis method, the evaluation and analysis of English writing teaching ability is carried out. And the information flow model of the big data evaluation, the control of a variety of constraint parameters objective function, the formula of the English writing teaching ability evaluation objective function is [8]:

Target function $\max_{x_{a,b,d,p}} \sum_{a\in A}\sum_{b\in B}\sum_{d\in D}\sum_{p\in P} x_{a,b,d,p} V_p$; constraint condition

$$\sum_{a\in A}\sum_{b\in B}\sum_{d\in D}\sum_{p\in P} x_{a,b,d,p} R_p^{bw} \leq K_p^{bw}(S), b\in B \quad (4)$$

When the grey system prediction model is used, the recursive queuing analysis and evaluation of English teachers' writing teaching ability can obtain the approximate solution of English teaching ability evaluation and prediction. Assume that the data of the distribution characteristics of English teachers' teaching ability is $\{x_i\}_{i=1}^N$, under the condition that the initial eigenvalue fluctuation of the K-means clustering algorithm is small, the probability density of the teaching ability evaluation function element is:

$$U_c(t)=Kx_c(t) \quad (5)$$

In the process of multiple iterations of the constraints of English writing teaching ability, the characteristic distribution probability of high-dimensional functions will also change accordingly. After k-1 iterations, the continuous function of the English teacher's teaching ability evaluation statistical model is transformed from u: I×IRd to u:IR(k≥1), and the gray-order continuous function satisfies the condition N(k)<L [9] . After using the recursive quantitative analysis method, the approximate sample value of the output information stream k of the output English writing teaching ability evaluation data is:

$$P_{1,J}=\sum_{d_i\in KNN} Sim(x,d_i)y(d_i,C_j) \quad (6)$$

The big data K-means clustering objective function is

$$J_m(U,V)=\sum_{k=1}^{n}\sum_{i=1}^{c}\mu_{ik}^m(d_{ik})^2 \quad (7)$$

According to the above, the distribution data of English teachers' teaching ability characteristics $\{x_i\}_{i=1}^N$ is obtained. Arrangement is a sequence of correlation distribution for the evaluation of English writing teaching ability. K-means clustering algorithm is used to optimize and quantify K-values, and finally the feature extraction results of quantitative recursive time series for English teachers' writing teaching ability evaluation are obtained:

ISSN 2522-6398 Vol. 6, Issue 2: 57-62, DOI: 10.25236/FER.2023.060210

$$x_n = a_0 + \sum_{i=1}^{M_{AR}} a x_{n-i} + \sum_{j=0}^{M_{MA}} b_j \eta_{n-j} \tag{8}$$

(a: the initial sampling maximum value of English writing teaching ability evaluation; x_{n-1}: vectorless time series; b_j: oscillation attenuation of English writing teaching ability evaluation).

2.3.3 Implementation Process of K-means Clustering Algorithm Applied in English Writing Teaching Ability Evaluation Model

Firstly, aiming at the above-mentioned English writing teaching ability constraint parameter model, the recursive analysis of the influence factors of various English teaching ability is carried out [10]. Then, using K-means clustering algorithm, the constraint parameters of English writing teaching ability evaluation are transformed into the least squares optimization problem of K clustering objective functions. The least squares optimization solve problem refers to the estimated consistency value of the constraint vector β of the English writing teaching ability evaluation, so that the value of the dependent variable Y and the independent variable X||Y-Xβ|| is the smallest [11].

Approximate sample vector for the evaluation of English teacher writing ability d_i. K-means clustering analysis is carried out into the constraint parameter model, and finally the least squares solution of the transformed English writing teaching ability evaluation is:

$$z(t)=x(t)+iy(t\}=a(t)e^{i\theta(t)}+n(t) \tag{9}$$

(x (t): the real part of the evaluation sequence of English writing teaching ability; the imaginary part of multiple evaluation data time series)

Using the obtained English writing teaching ability evaluation time series, Gaussian linear random processing of its mean value and variance is obtained, and the substitute data x'(k) of different clustering targets is obtained. Then, the original time series data and the substitute data are tested and compared statistically, and the distribution data of the k-th writing teaching ability evaluation is disturbed and analyzed, and the sub-class set of the k-th clustering objective function is obtained. This can determine the utilization rate $Uutil=\gamma^{\bar{X}}$ of various English writing teaching ability evaluation indicators. Then, using the top-down hierarchical decision tree greedy algorithm, the central time series feature quantity of English writing teaching ability evaluation is analyzed to determine the subset similarity of different teaching ability evaluation resource distribution [12]:

$$Sim_1(d_i, d_{1_j}) = \frac{\sum_{k=1}^{M} W_{ik} \times W_{1jk}}{\sqrt{\sum_{k=1}^{M} W_{ik}^2} \sqrt{\sum_{k=1}^{M} W_{1jk}^2}} \tag{10}$$

(d_i: Evaluate the prior prediction feature vector for English teaching ability; d_{1j}: 1 iteration center vector for large data K-means clustering.)

Combined with the feature fusion algorithm of information correlation, the parameters of the English writing teaching ability evaluation time series are clustered and integrated [13], and the output expression of the teaching ability evaluation resource information fusion is

$$P(\omega|x)=P(x|\omega)/P(x). \tag{11}$$

If quantitative recursive analysis is used to classify the probability density of English teaching ability assessment resources, the English writing teaching ability can be evaluated as a big data stream X(i). It is divided into P(i) N_{ij}×m sub-matrix X_{ij} with the same value, in which $p(i)=\left[\frac{N(i)}{L}\right]$,

$$N(i)modL<m \tag{12}$$

Through the clustering integration of multiple English teaching ability evaluation index sequences, it can complete the evaluation and optimization of teachers' English teaching ability, and promote the resource allocation of English writing teaching and the effective improvement of teachers' English teaching skills [14].

ISSN 2522-6398 Vol. 6, Issue 2: 57-62, DOI: 10.25236/FER.2023.060210

2.4 Matlab simulation experiment on English writing teaching ability assessment

2.4.1 Data acquisition and preprocessing

English writing teaching ability of college teachers is selected as a data training sample to preprocess such factors as teachers' professional quality, teaching resource organization ability, classroom teaching ability, scientific research and innovation ability. Pandas, a data analysis and processing library, is used to read the collected English. CSV data files, find and delete missing values and outliers without correlation, or use Matplotlib graph software to draw scatter graph of clustering factors, and delete those data that are far away from the clustering center [15]. After the processing and analysis of all strings with numerical values, visual analysis of the data of English writing teaching ability can be drawn. There are mainly two distribution forms: centered around the clustering center and divergent away from the clustering center. The overall distribution trend of data factors is shown in figure 2:

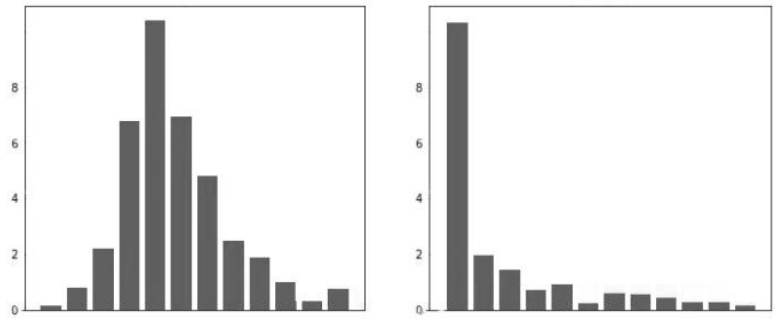

Figure 2: Visual distribution trend of English writing teaching ability data

It can be seen from the above figure that teachers' writing teaching ability and writing teaching behavior path are both random and not subject to the constraints of time and place, so the writing teaching obtained may not meet the teaching expectation. Therefore, for the collection of clustering factors and data preprocessing of teachers' writing teaching ability, it is necessary to remove the collected redundant data samples before further clustering iteration can obtain the ideal values of various English teaching constraint parameters [16].

2.4.2 Matlab simulation experiment and discussion on English writing teaching ability evaluation

The statistical results of the sampling data of the teaching ability are evaluated, and information fusion and index clustering processing are carried out. Assume that the relevance distribution parameters of English writing teaching resources are $\xi_{c1}^{d2}=3/5$, $\xi_{c2}^{d2}=2/5$, $\xi_{c3}^{d2}=2/5$, $\max g_{c1}(d_2)=6/5$, $\max g_{c2}(d_2)=3/8$, and $\max g_{c3}(d_2)=1/10$. The initial adaptive step size isρ=0.95, the sampling frequency is f0=600Hz. The critical value of writing teaching ability evaluation decision is Dx=2, and the correlation distribution parameter of teaching resource characteristics is B=1.13. According to the above data parameters, the Matlab numerical analysis and simulation software is used to reconstruct the constraint parameters of the English writing teaching ability evaluation, and the optimal test results of the teaching ability evaluation index are obtained, as shown in Table 1:

Table 1: Accuracy Test Results of Teaching Ability Evaluation Indicators

Iteration cycle	1	2	3	4
Evaluation accuracy /%	96.53	97.04	98.19	98.68
Utilization rate /%	99.01	98.22	97.75	96.43

According to the above Table 1, it can be concluded that using Matlab numerical analysis and simulation software, the big data K-means clustering parameters of English teachers' writing teaching ability evaluation, information fusion and data clustering processing are carried out. This will give the best similar solution to the assessment of teaching ability [17]. This indicates that the algorithm has low redundancy when evaluating English teachers' writing ability, and the minimum error accuracy is high.

3. Conclusion

This paper mainly analyzes the elements of English teacher writing teaching ability based on information fusion feature algorithm and K-means clustering algorithm. Firstly, the data flow model of

ISSN 2522-6398 Vol. 6, Issue 2: 57-62, DOI: 10.25236/FER.2023.060210

teaching ability evaluation is constructed, the objective function of different constraint parameter evaluation indicators is determined, and the quantitative recursive time series of teaching ability evaluation is extracted [18]. Then, the feature quantity of the time series is analyzed, the subset similarity of different teaching ability evaluation resource distribution is determined, and the information fusion algorithm is used to calculate the output expression of the teaching ability evaluation resource. Finally, using Matlab numerical analysis and simulation method, the big data reconstruction of various constraint parameters is completed, and the big data time domain waveform and optimal solution of teaching evaluation resources are obtained. At the same time study also has limitations: without experience, it is difficult to determine the number of category K value, thus affecting the clustering quality.

References

[1] Huo Yujia. Analysis of teaching quality evaluation system based on big data mining [J]. Computer products and circulation, 2019(06):232.
[2] Chen Mingxuan, Geng Nan. Research on evaluating effective teaching supported by big data [J]. Journal of distance education, 2019(03):95-102.
[3] Yu Yang, Yu tao, Wang xinghua. Research on English writing teaching based on big data -- a case study of millions of the same titles in 2017 [J]. Journal of northeast agricultural university (social science edition), 2019, 17(02): 82-88.
[4] Gao guoliang. Research on English teaching ability evaluation algorithm based on big data fuzzy k-means clustering [J]. Modern electronic technology, 2017, 40(20): 31-33.
[5] Wang Lu and Zhao Min. Exploration and Practice of Teaching Sales Strategy and Art Curriculum Politics [J]. Journal of Anhui Open University, 2022, No. 3.
[6] Xia Yongquan, Sun Jingru, WU Xin-wen, Zhi Jun, Wang Bing, Xie Xanxiang. Study on star point clustering based on improved k-means clustering algorithm [J]. Acta graphica sinica, 2019, 40(02): 358-363.
[7] Yao Jialin, Tan Zhao, Yu Songnan, Ju Yinyin, Liu Sujuan. Study on learning behavior analysis based on k-means clustering algorithm and naive bayesian classifier [J]. Computer knowledge and technology, 2019, 15(11): 17-20.
[8] Li Shuang, Chen Ruirui, Lin Nan. Hadoop framework k-means clustering algorithm for big data mining [J]. Computer engineering and design, 2008, 39(12):3734-3738.
[9] Liu Hongfei, Xu Qiang, Xu Hongguo, Bao Cuizhu, Wang Guojun. Determination method of vehicle lateral stability based on k-means cluster analysis [J]. Journal of hunan university (natural science edition), 2018, 45(08): 48-53.
[10] Yang Changqing. Attribute selection algorithm combining k-means clustering and low-rank constraints [J]. Chinese journal of information technology, 2008, 32(07):91-98.
[11] Zheng Xin. Application of analytic hierarchy process model in teaching evaluation [J]. Science and technology economic guide, 2008, 26(17):149.
[12] Yang Juan, Qu Chuanhui. Improved k-means clustering algorithm [J]. Ship electronic countermeasures, 2017, 40(06):91-93.
[13] Wang Peng. Research on English teaching evaluation based on fuzzy comprehensive evaluation model [J]. China packaging, 2017, 37(11):80-83.
[14] Chen Ziyang, Wang Xiaohua. Application of k-means clustering method in high school student performance analysis [J]. Mathematics learning and research, 2017(21): 160-161.
[15] Sun Yan. Evaluation of English teaching ability based on feedback model [J]. Modern electronic technology, 2017, 40(20):25-27.
[16] Xiao L, Dai Q, Liu S, Xie Q, Shu H. Teaching reform of higher vocational "Computer Application Fundamentals" course based on computational thinking [J]. Nanfang Agricultural Machinery 2021; 52(15):191-192+196.
[17] Xue Yingxi, Xu Hongwen, Li ling. Global optimal K mean clustering algorithm based on sample density [J]. Computer engineering and application, 2008, 54(14):143-147.
[18] Luo Juan. Research on dynamic evaluation model of college English classroom teaching [J]. Journal of dongguan university of technology, 2017, 24(04): 107-111.