Advances in Educational Technology and Psychology (2021) 5: 29-36
Clausius Scientific Press, Canada
DOI: 10.23977/aetp.2021.55006
ISSN 2371-9400

Research on English Teaching Ability Evaluation Algorithm Based on Computer Corpus

Li Zhang

School of Foreign Languages, Zhongyuan University of Technology, Zhengzhou, Henan, 45000, China

Keywords: Computer corpus, English teaching ability assessment, K-means clustering

Abstract: Corpus refers to a large-scale electronic library with a certain capacity, which is built by collecting naturally occurring continuous language application texts or discourse fragments according to certain linguistic principles and using random sampling methods. The application of corpus in English teaching can provide abundant context for learners, which is beneficial to the construction of learners' language knowledge. In view of the low accuracy of English teaching ability evaluation at present, a teaching quality evaluation method based on computer corpus is proposed. Combining K-means clustering algorithm, clustering and integrating the index parameters of English teaching ability, the teaching resource allocation plan is completed, thus realizing the scientific evaluation of English teaching ability. The research shows that the algorithm can combine the characteristics of schools, compare the overall teaching quality of universities of different grades or the same grade, and objectively reflect the overall teaching level of schools.

1. Introduction

Teaching quality evaluation of university teachers is an important way to improve teaching management level and teachers' teaching ability [1]. Using teaching quality evaluation, students can feedback teachers' teaching situation, teachers can reflect on teaching effect, schools can effectively implement teaching management improvement, and provide targeted training for teachers. English teaching is an important part of higher education. At present, all universities have their own teaching evaluation regulations and methods, which mainly include students' learning evaluation and teachers' English teaching ability evaluation [2]. Classroom teaching is the basic link of higher education, the center for teachers and students to work and study in colleges and universities, and the main object of internal control in colleges and universities. In-depth study on English teaching ability evaluation is of great significance to both higher education theory and higher education practice [3].

According to certain linguistic principles, corpus is a large-scale electronic library with a certain capacity, which collects naturally occurring continuous languages and uses text or discourse fragments. In essence, corpus is actually a random sampling of natural language use, with a certain size of language samples representing the overall language use determined in a certain research [4]. With the gradual refinement and deepening of corpus processing, our language knowledge becomes more accurate and profound. There are researches on English majors' learning English, non-English

majors' learning English, and postgraduates' learning English [5]. In view of the evaluation of teaching quality in colleges and universities, researchers put forward a variety of evaluation methods, such as grey relational analysis, analytic hierarchy process, fuzzy comprehensive evaluation method and so on. However, these methods are suitable for linear models, which are difficult to adapt to the non-linear evaluation of teaching quality. They are subjective and random, and can not realize the evaluation of teaching quality well. In this paper, an English teaching ability estimation method based on big data fuzzy K-means clustering and information fusion is proposed, which can cluster and integrate the index parameters of English teaching ability, compile the corresponding teaching resource allocation plan, realize the quantitative planning of English teaching ability evaluation, and realize the accurate evaluation of English teaching ability.

2. Research Method

2.1 Evaluation Index System of English Teaching Ability

(1)Analytic hierarchy process

Analytic Hierarchy Process (AHP) quantifies qualitative factors and semi-qualitative factors, levels various factors to be considered in the problem, and quantifies and compares various related factors layer by layer, providing a quantitative method for analysis, decision-making, prediction or control; It absorbs the characteristics of behavioral science, quantifies the empirical judgment of decision makers, and is especially suitable for problems that are difficult to be completely quantitatively analyzed [5].

The first step: to establish a hierarchical structure model. When applying AHP to analyze decision-making problems, we must first organize and level the problems and construct a hierarchical structure model, which decomposes the relevant factors into several levels from top to bottom according to their attributes. Each factor in the same level belongs to the upper level factor, or has influence on the upper level factor, while dominating the next level factor or being influenced by the lower level factor.

Step 2: Construct a comparative judgment matrix. After the hierarchical structure model is established, the subordinate relationship of factors between the upper and lower levels is determined. It is necessary to compare the influence of N factors $x_1, x_2, \cdots, x_n$ of a certain level on the target Z of the upper level, and quantify the “importance” of each factor by the method of pairwise comparison.

Take two positive numbers $x_i x_j$ at a time, and use positive numbers $b_{ij} = \frac{x_1}{x_j}$ to express the ratio of importance of x_i and x_j relative to the target Z of the previous layer. The matrix obtained from all comparison results is called comparative judgment matrix $A = (b_{ij})_{m \times n}$. There's obviously a $b_{ij} > 0, b_{ij} = \frac{1}{b_{ij}}, b_{ij} = 1$.

Step 3: Hierarchical sorting and consistency check. Calculate the maximum eigenvalue $\lambda_{\max}$ and the corresponding eigenvector w of each comparative judgment matrix. After the consistency test, the feature vector w has satisfactory consistency, which can be used as the ranking weight vector of the importance of each factor in the next level relative to the factor in the previous level. This process is called hierarchical single ranking.

Step 4: Calculate the combination weight vector and make a combination consistency test, that is, hierarchical total sorting and consistency test. What is obtained above is the weight vector of a

group of factors to a factor in its upper layer. Finally, what is need is that sorting weight vector of each scheme in the low layer for the overall goal. The total sorting weight vector should be weighted from top to bottom. At that time, CR was less than 0.1, and the results of hierarchical ranking were considered to be satisfactory. Otherwise, the comparison judgment matrix should be readjusted.

(2)Constructing the evaluation system of English teaching ability

On the basis of a large number of investigations and data collection, according to the main contents of college English teaching ability evaluation, the hierarchical structure system of English teaching ability evaluation is established, as shown in Table 1.

Table 1 Evaluation System Of English Teaching Ability

Target layer	Criterion layer	Method layer
Teaching ability	Teaching attitude B_1	Teaching and educating people, strict management, facing all students C_{11}
		Prepare lessons adequately and skillfully C_{12}
		Explain carefully, be full of energy, and attend classes on time C_{13}
	Content of courses B_2	The purpose is clear, and the difficulties are outstanding C_{21}
		The concept is accurate and the argument is reliable C_{22}
		Combining theory with practice, the teaching content is advanced C_{23}
	Method art B_3	Clear thinking and strong logic C_{31}
		The language is fluent and infectious C_{32}
		Give full play to the leading role of teachers and the main role of students C_{33}
		Teaching students in accordance with their aptitude and flexible methods C_{34}
	Teaching effect B_4	Classroom order is good and students' thinking is active C_{41}
		Complete teaching tasks and achieve teaching objectives C_{42}
		Students' ability and quality have been improved C_{43}

2.2 Text Migration Based on Clustering

Although the existing auxiliary data has expired. However, among these existing data, there should still be some data that are very similar to the test data, which can be used to help the study of the target task [6]. Therefore, it is considered to use clustering technology to find out the data which is very similar to the test data from the existing data.

The purpose of text clustering is to group large-scale text data sets into multiple classes, and to make the texts in the same class have high similarity, while the texts in different classes are quite different. As a function of data mining, clustering can be used as an independent tool to obtain the distribution of data, observe the characteristics of each cluster, and focus on some specific clusters for further analysis [7]. At the same time, clustering technology can also be used as a preprocessing step of other algorithms, which can effectively improve the classification performance of other algorithms.

From a psychological point of view, transfer learning is the basic skill of human beings. When the existing training data is out of date, and there are few new data, or the cost of labeling new data is very high, it is the task of migration learning to use the existing training data under different distributions to help the learning of new data (Figure 1).

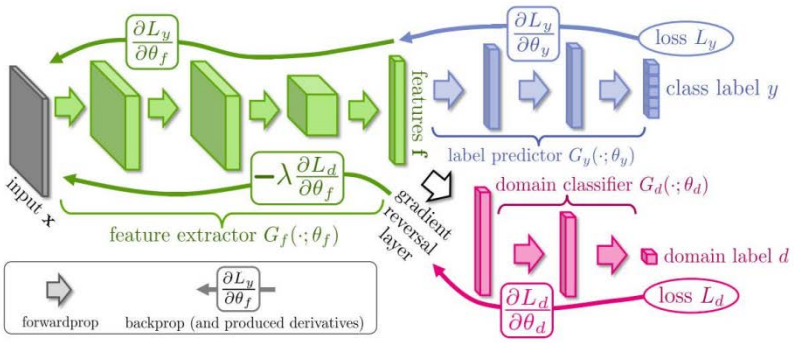

Fig.1 Transfer Learning Process

According to the traditional representation of vector space model, the text content can be expressed as a weighted feature vector.

If D is a text set, d_i represents a text in the set, t_i represents the i-th feature word, and w_i represents the weight of the i th feature word, then a document is expressed as a vector: $d_i\left(t_1, w_1; t_2, w_2, \cdots, t_n, w_n\right)$. The weight w_i can be expressed by $tf-idf$ weight [8] of each feature. $tf-idf$ formula is as follows:

$$tf-idf=\sum_{d\in D} tf\left(d,t\right)\cdot\log\frac{\left|D\right|}{df\left(t\right)} \quad (1)$$

In which $tf\left(d,t\right)$ is the word frequency of word t in text d, $df\left(t\right)$ is the number of texts containing word t in text set D, and $\left|D\right|$ represents the number of texts contained in text set D.

The similarity between two texts [9] can be calculated by the cosine of the angle α between two vectors. Assuming that the two texts are $d_1=\left(t_1, w_1; t_2, w_2, \cdots, t_n, w_n\right)$ and $d_2=\left(t_1, \sigma_1; t_2, \sigma_2, \cdots, t_n, \sigma_n\right)$ respectively, the similarity between d_1 and d_2 is expressed as:

$$sim\left(d_1,d_2\right)=\cos\alpha=\frac{\sum_{i=I}^{n}\omega^i\times\sigma^i}{\left(\sum_{i=I}^{n}\omega_I^2\times\sum_{i=I}^{n}\sigma_I^2\right)^{I/2}} \quad (2)$$

The larger the value of $sim\left(d_1,d_2\right)$, the more similar the two texts are.

2.3 Evaluation Algorithm of English Teaching Ability

Constrained parameter index analysis model is constructed, and big data information model analysis of ability evaluation is completed. In order to further improve the quantitative evaluation ability of English teaching ability, this paper proposes an English teaching ability evaluation algorithm based on fuzzy K-means clustering.

The evaluation problem is transformed into solving the objective function β (K-means clustering) as the least square estimation problem, that is, finding the consistency estimation value of English teaching ability (evaluation resource constraint vector) to minimize $\left\|Y-X\beta\right\|$, where $\left\|\cdot\right\|$ represents Euclidean norm and obtains the entropy feature of ability constraint feature information [10], and its extracted value is expressed as formula (7):

$$P_{loss}=1-\frac{1-\rho_0}{\rho}=\frac{\rho_0+\rho-1}{\rho}=\sum_{n=1}^{N}p_{kn}\ (3)$$

The estimation formula of English teaching ability is transformed into the least square solution, as shown in formula (8).

$$z(t)=x(t)+iy(t)=a(t)e^{i\theta(t)}+n(t)(8)$$

The real part of the evaluation big data distribution time series is represented by $x(t)$, and the imaginary part of the evaluation constraint index series is represented by $y(t)$. The amplitude randomization of English teaching ability is realized by adopting the alternative data method. The empirical distribution data (the evaluation of teaching ability in the k-th class) is perturbed by functional, and the k-th subclass set is obtained. Based on this, the expression of English teaching resource utilization is shown in Formula (9).

$$U_{util}=\gamma\overline{X}\ (4)$$

Constructing hierarchical tree, the establishment of principal component feature quantity of English teaching ability evaluation is realized by using big data analysis method, and the similarity of teaching resource distribution is solved by using fuzzy closeness filling method, and the specific expression is shown in Formula (10).

$$sim_1(d_i,d_{1j})=\frac{\sum_{k=1}^{M}W_{ik}\times W_{1jk}}{\sqrt{\sum_{k=1}^{M}W_{ik}^2}\cdot\sqrt{\sum_{k=1}^{M}W_{1jk}^2}}\ (5)$$

d_i represents the prior distribution feature vector of English teaching ability evaluation; The K-means clustering center vector of tier 1 big data is represented by d_{1j}. Combining with the linear feature fusion method, the ability evaluation index parameters are clustered and integrated, and the output information fusion expression is obtained as shown in Formula (11).

$$P\left(\frac{w}{x}\right)=P\frac{(xw)}{p(x)}(6)$$

On the basis of clustering and integration of index parameters, the allocation plan of teaching resources and teaching ability is completed, so as to optimize the evaluation of English teaching ability and improve the accuracy and practicability of the evaluation.

3. Discussion and Analysis of Results

In order to test the effectiveness of the English teaching ability evaluation algorithm based on fuzzy K-means clustering proposed in this paper, this section uses the collected teaching quality evaluation data of college English teachers to test its performance.

3.1 Selection of Evaluation Index by Ahp

Scientific evaluation of English teaching ability is a necessary measure to evaluate teaching level, sum up teaching experience and promote teaching reform. However, English teaching ability is a quality system composed of many factors, and the feeling each factor gives to the appraisers is a kind of information lacking clear boundaries in many cases. Moreover, it is necessary to process and express this information with the help of reasonable mathematical methods, so that the evaluation conclusion can be scientific, reasonable and credible [11]. In order to make the evaluation results more scientific, reasonable and democratic, and to eliminate the influence of

human factors on the evaluation results, this paper uses analytic hierarchy process to establish a mathematical model, construct an index system for the evaluation of English teaching ability, and make the evaluation quantitative.

The analytic hierarchy process (AHP) is used to analyze the evaluation indexes of English teaching ability, and the indexes that contribute the most to the evaluation of teaching quality are selected. The analytic hierarchy process results are shown in Figure 2.

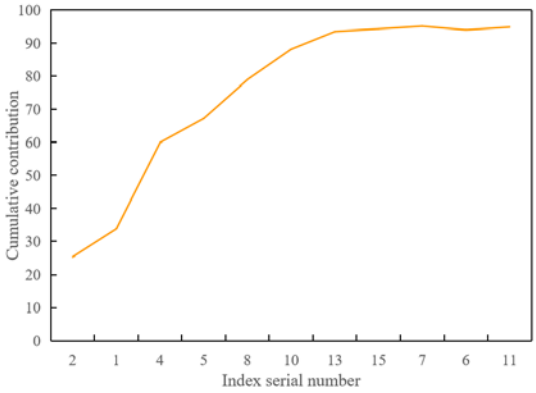

Fig.2 The Result of Analytic Hierarchy Process on Evaluation Index

Figure 3 shows that the cumulative contribution rate of the first eight principal component indicators to English teaching ability evaluation has reached 93.5%, exceeding the set threshold of cumulative contribution rate, which indicates that the first eight indicators can effectively represent the important information contained in all indicators, and these eight indicators will be used to evaluate English teaching ability in the future.

3.2 Analysis of Teaching Ability Evaluation Results

Corpus linguistics is one of the best and effective ways to implement autonomous learning, which is conducive to cultivating students' autonomous and inquiry learning models. Students can rely on a large number of authentic and authentic language inputs in the corpus to awaken their language awareness. Corpus linguistics not only cultivates students' abilities of active learning, independent learning, research learning, exploratory learning, experiential learning, participation in learning and mutual learning, but also cultivates students' abilities of discovering, thinking, analyzing and solving problems independently. It is also a good model of “quality” and “quantity”, “learning” and “getting”.

Corpus is one of the learning activities that can stimulate students' critical thinking. The learning mode of critical thinking is closely coordinated with autonomous learning, which is different from the traditional teaching and passive receiving learning modes. After the students bring up a large number of language examples through the corpus, they should classify and summarize them at multiple levels according to the similarities and differences of context features, and gradually dig out the regular things in them. This kind of autonomous learning, which belongs to the type of analysis, can stimulate students' curiosity.

The fuzzy K-means clustering constructed in this paper is used to evaluate English teaching ability, and the curve of mean square error in the process of model training is shown in Figure 3.

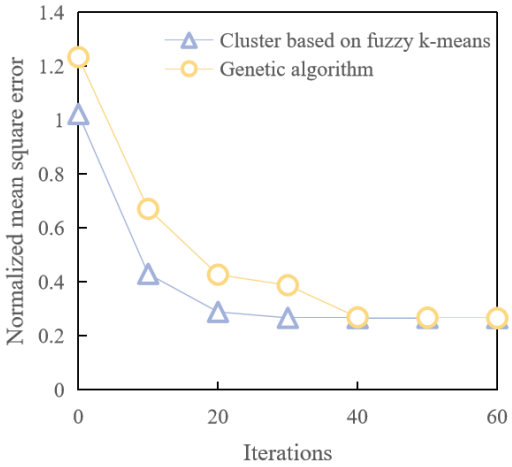

Fig.3 Model Mean Square Error Curve

Fig. 3 shows that genetic algorithm converges after about 50 iterations, while fuzzy K-means clustering only needs about 35 iterations, which shows that fuzzy K-means clustering can effectively improve the convergence speed of the algorithm and reduce the training time of the model. Moreover, the mean square error of fuzzy K-means clustering is always lower than that of genetic algorithm.

The accuracy of teaching ability evaluation of test group data is shown in Figure 4.

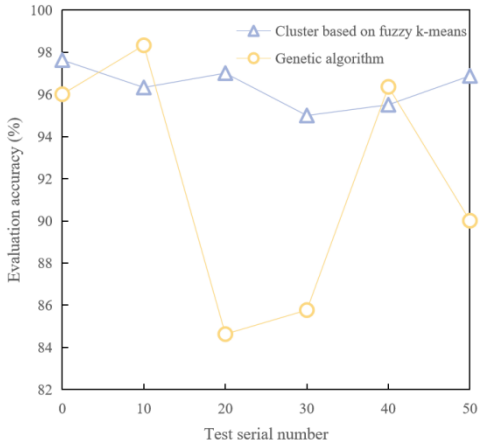

Fig.4 Accuracy Curve of Model Evaluation

Fig. 4 shows that the accuracy of the evaluation of English teaching ability by genetic algorithm is higher than 80%, but the evaluation accuracy of fuzzy K-means clustering is better than that of genetic algorithm. The statistical results show that the evaluation accuracy of genetic algorithm is higher than 91% in 35 groups of 50 test samples, and the evaluation accuracy of 30 groups is higher than 94%, which shows that the model has high approximation accuracy.

4. Conclusion

Practice has proved that the potential of corpus in foreign language teaching can be greatly explored. Corpus-driven learning is not only helpful to improve English teaching ability, but also can fully mobilize students' subjective initiative, stimulate learning motivation and actively participate in foreign language learning, which is helpful to cultivate students' autonomous learning awareness and improve their learning ability. Therefore, this paper mainly designs the evaluation

system of college English teaching ability, and proposes an evaluation method based on big data fuzzy K-means clustering algorithm based on computer corpus. Based on the establishment of constraint parameter index analysis model, it evaluates the ability of big data information model (using quantitative recursive analysis method), extracts the entropy feature of ability constraint feature information, fuses K-means clustering algorithm, clusters and integrates the index parameters of English teaching ability, and completes the preparation of teaching resource allocation plan, thus realizing the scientific evaluation of English teaching ability. The fuzzy evaluation method proposed in this paper effectively reduces the interference caused by the fuzziness of indicators, and makes the evaluation results have higher reliability and validity. For teachers, teachers can compare the gap with other teachers vertically, and can also compare their progress in all aspects of teaching in a year or even longer by students' scores horizontally.

References

[1] Zhaowei. Evaluation of Teaching Ability of Young Teachers in Local Colleges and Universities Based on Students' Evaluation of Teaching Data[J]. Journal of Huaihai Institute of Technology (Social Science Edition), 2018, 016(009):129-131.

[2] WK Zhao, ZQ Wang. Construction of Evaluation Indicator System of Teaching Ability of Physical Education Teachers in Colleges and Universities[J]. Sports research and education, 2018, 033(006):69-74.

[3] Au J, Palmer E, Johnson I, et al. Evaluation of the utility of teaching joint relocations using cadaveric specimens[J]. Bmc Medical Education, 2018, 18(1):41.

[4] Wu, Xiuyun. The Evaluation of the Innovation Teaching Ability of Physical Education in Colleges and Universities under the Regional Difference[J]. IEEE Computer Society, 2017:346-349.

[5] Ying W, Fei Y. A study on the influence mechanism of PBL teaching model on non-English major students' critical thinking ability[J]. International technology management, 2017(6):39-41.

[6] Liu Z. A Multi-Index Measurement Model of English Classroom Teaching Level in Colleges and Universities Based on Fuzzy System Theory[J]. Journal of Computational and Theoretical Nanoscience, vol. 14, issue 1, pp. 46-50, 2017, 14.

[7] Nejad H H, Bagherabadi M, Sistani A, et al. Effectiveness of resident as teacher curriculum in preparing emergency medicine residents for their teaching role[J]. J Adv Med Educ Prof, 2017, 5(1):21-25.

[8] A S A V, B W K. Instructional model development to enhance critical thinking and critical thinking teaching ability of trainee students at regional teaching training center in Takeo province, Cambodia[J]. Kasetsart Journal of Social Sciences, 2017, 38(1):88-95.

[9] Liang G, Yang C, Zhu S, et al. Application and Practice of Remote Measurement and Control Platform in the Experiment Teaching Course of "Data Acquisition and Interface Technology"[J]. Creative Education, 2018, 9(5):779-790.

[10] Lu L S, Guan S S. A STUSY ON DIFERENCE BETWEEN CASE-BASED INSTRUCTION AND DIDACTIC INSTRUCTION ABOUT STUDENT'S WORKS PERFORMANCE OF THE EXHIBITION DESIGN[J]. Bulletin of Japanese Society for Science of Design, 2017, 58:75-84.

[11] Duan Aiyuan, Liu Yang. Evaluation on the Implementation Effect of University Teaching System to Cultivate Students'Innovative Practice Ability[J]. higher education of sciences, 2018, 000(005):87-93.