Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
 Volume 25, Issues 1, pp. 407–421, DOI 10.12694/scpe.v25i1.2306

RESEARCH ON THE EVALUATION MODEL OF STUDENTS' FOREIGN LANGUAGE LEARNING SITUATION BASED ON ORIENTED ONLINE TEACHING COLLABORATION PLATFORM

FEI HE*

Abstract. "Internet + Education" makes online teaching gradually penetrate the education industry, and makes the industry enter a great revolution based on information technology. The traditional student learning evaluation system cannot satisfy the actual demand of current learning evaluation. This paper constructs an evaluation model for the foreign language learning of online students. Firstly, the DBSCAN algorithm with distance optimization is used to conduct cluster analysis on the description indicators of student behavior, and the student groups with different behavior characteristics are obtained. Then the ANOVA F-test was used to extract the features of different student groups. Finally, a novel N-Adaboost algorithm based on multiple classifiers is proposed and a model is constructed to evaluate students' foreign language learning. The experimental results show that the accuracy of the evaluation model is 74.02% in the pass and fail groups and 73.74% in the excellent and non-excellent groups. Students' listening, speaking, and reading abilities are in a state of upward development overall through the online teaching collaboration platform, but their writing ability is obviously declining. There is a great improvement in foreign language vocabulary. This study provides a new perspective of thinking for the improvement of the quality of school teaching management, the analysis of students' behavior, and the evaluation of learning situations, and provides a new solution for the problem of students' learning situations in modern information teaching.

Key words: Online Education, Foreign Language Learning, Evaluation Model, N-Adaboost, DBSCAN

1. Introduction. With the popularization of Internet technology and the continuous development of cloud computing, big data, artificial intelligence and other technologies, Internet + education is gradually changing the traditional teaching mode. Now, students can study anytime and anywhere through the online learning platform, no longer limited by time and space [1, 2]. The online education platform not only breaks through the time and space restrictions, and provides students with rich learning resources, but also provides students with personalized learning recommendation and evaluation [3] through the intelligent learning system. At present, the online teaching model has attracted the attention of many educators at home and abroad, and many universities have deployed and implemented their own learning platforms. Some online platform courses and open courses have reached tens of millions of users. As an important subject of personal development and social needs, the teaching collaboration platform greatly promotes the efficiency and accessibility of students' foreign language learning. However, online foreign language learning also has the problem of high user dropout rate. Because learners have greater freedom and strong subjectivity, the lack of supervision and intervention when learning risk occurs [4]. Therefore, it is of great significance to establish a model that can evaluate students 'foreign language learning level to improve the teaching quality and students' learning effect. However, the traditional evaluation method often can not adapt to the characteristics of online teaching, so the research combines the density-based clustering algorithm (Density-Based Spatial Clustering of Applications with Noise DBSCAN) algorithm based on distance optimization with the multi-classifier based Adaboost algorithm (N-Adaboost) to build the foreign language learning evaluation model for students on the online teaching collaboration platform. It is hoped that through the establishment of such an evaluation model, students 'learning behavior and learning level can be better understood, provide teachers with more accurate teaching suggestions, and promote students' learning and development.

2. Related Work. Nowadays, the digital construction of colleges and universities continues to advance, the process of education reform continues to accelerate, and student learning evaluation is very important.

*School of Foreign Studies, Henan Polytechnic University, Jiaozuo, 454003, China (feihefhfh@outlook.com)

In the face of the problem of using digital technology to improve the efficiency of English learning and the effectiveness of teaching evaluation, researchers such as Susanty L use search engines to obtain data and use coding, interpretation, and other means to obtain results. Through data analysis and discussion, it is found that the use of digital technology in English teaching can significantly improve students' classroom participation; the introduction of this technology in teaching evaluation has effectively improved evaluation efficiency and accuracy [5].

Liu H's research team found that there are some problems in the current teaching evaluation system. To conduct a more comprehensive quality evaluation, an evaluation system and model that introduced the entropy weight method and gray clustering were proposed. Through example analysis, the results show that the quantitative indicators of the system are practical and innovative; the stability and accuracy of the constructed evaluation model are the best and have great practical application value [6].

Aiming at the problems of fuzzy evaluation index and imperfect system, scholars Li N constructed a fuzzy evaluation model of the analytic hierarchy process. Relevant experimental data show that the model integrates qualitative and quantitative indicators to conduct teaching evaluation from various aspects, which significantly improves the reliability of the evaluation results. At the same time, the model can effectively deal with fuzzy indicators, which greatly promotes the improvement of English teaching quality [7].

Zhang Y's researcher found that when teaching evaluation in online teaching, there are many fuzzy indicators, and the evaluation effect is greatly reduced. The key indicators that restrict the evaluation model are studied and analyzed, and a multi-attribute fuzzy evaluation model is constructed. The simulation experiment results show that the model can accurately evaluate the effect of online teaching, effectively quantify the quality of teaching in all aspects, and has a good development prospect [8].

Facing some of the problems existing in English teaching, Tran TQT and other teams expounded on the factors that affect the effect of English teaching from many aspects. Through questionnaire survey and data analysis, relevant data shows that students are not motivated enough in class, and interesting teaching strategies can improve their learning enthusiasm. At the same time, students' habits and interests will affect the teaching effect. Colleges and universities should formulate corresponding plans to improve students' classroom participation and improve teaching quality [9].

Student learning evaluation is an important means to improve teaching quality and reform teaching modes in colleges and universities. Wang and other scholars found that several factors such as teachers, students, teaching methods and teaching environment can significantly affect the teaching effect. The research starts with teaching methods, reforms the existing teaching strategies, and passes relevant experimental tests. The results show that after implementing the reformed teaching strategy, students are more motivated, the classroom atmosphere is more active, and their professional ability has been significantly improved. The results provide a reference for colleges and universities to formulate teaching effect evaluation indicators [10].

Li and other research teams used the attributes of evaluation indicators to construct an evaluation index system with different attribute dimensions. The system is used in the evaluation of English teaching ability, and the results show that the system has a simple structure, and can flexibly adjust the weight parameters to adapt to the evaluation needs of different grades, and objectively reflect the teaching level of colleges and universities from many aspects [11].

Researchers such as Sun have developed an English system that combines artificial intelligence and teaching assistance, which can mine potential connections between information and use decision tree technology to independently evaluate teaching effects. It can be seen from the actual application data that the system can accurately analyze the students' mastery of knowledge, help teachers provide the basis for improving teaching strategies, and improve the level of teacher education and the efficiency of students and students [12].

Facing the problems of slow speed and low accuracy of the existing English interpreting evaluation models, Lu et al. used principal component analysis to screen indicators, used a radial basis network evaluation model, and used genetic algorithms to optimize parameters. The simulation experiment data shows that the selected indicators are representative, the evaluation efficiency and accuracy of the constructed evaluation model are significantly improved, and it has high real-time performance [13].

Research scholars such as Fang C found that online evaluation can conduct an overall analysis of teaching activities and provide a basis for educational improvement, and proposed an evaluation model based on support

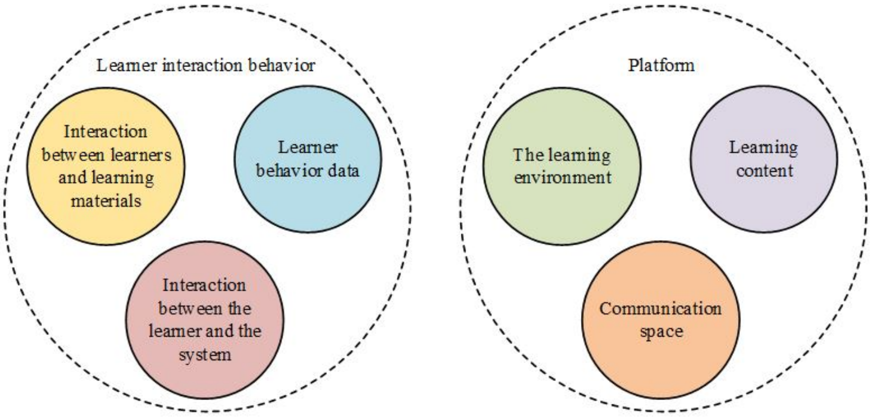


Fig. 3.1: Interaction between platform and learner in online teaching

vector machines. The experimental data show that the improved index dimensionality reduction method can improve the evaluation accuracy and reduce the interference factors. The evaluation effect of this model is better than that of the comparison method, and it has certain value in practice [14].

By expounding the achievements of domestic and foreign researchers, it is found that student learning evaluation plays an important role in education reform. Many scholars have put forward their own improvement plans for different problems, but almost all of them are based on shallow level analysis of simple data and models. Therefore, an improved N-Adaboost model based on multiple classiclassiers is proposed in this study. Through the use of modern intelligent technology to give a number of data evaluation of student learning.

3. Construction of an evaluation model for students' foreign language learning for an online teaching collaboration platform. With the rapid development of science and technology and the popularity of the Internet, more and more people choose to learn foreign languages through online platforms. However, compared with traditional face-to-face teaching, online learning is challenging. To address this problem, the research chose to construct an evaluation model specifically for online students' foreign language learning, providing students with personalized learning advice and feedback, as well as the role of providing guidance and supervision for teachers. Through the evaluation model, online students can better self-learning and improve learning results.

3.1. DBSCAN algorithm based on distance optimization. The rapid development of Internet technology and the innovation and transformation of educational concepts and methods have formed a new teaching method — online teaching [15]. The learners of online teaching break the predicament that traditional teaching is limited by time and space. And backed by big data, data mining and other network technologies, online learning has very rich learning resources, which can meet the diverse and personalized needs of many learners [16, 17]. Online learning is the sum of the teaching activities carried out by online learners and related learning groups through the interaction of online learning platforms in order to complete specific learning tasks, as shown in Figure 3.1.

From Figure 3.1, in online teaching, the platform can recommend a large number and various forms of learning content for learners. It also gives learners an independent space for self-directed learning and a collaborative environment for learning with others. At the same time, it can also meet the fast and efficient communication. At the same time, during online learning, the learner will leave a lot of behavior data, such as the interaction between the learner and the learning materials, the interaction between the learners, and the interaction between the learner and the system. The online learning platform uses cloud computing technology to process the massive data it owns, and then analyzes the characteristics of learners' learning behavior through data mining technology. Through the analysis, the learning rules are obtained, so as to carry out targeted push

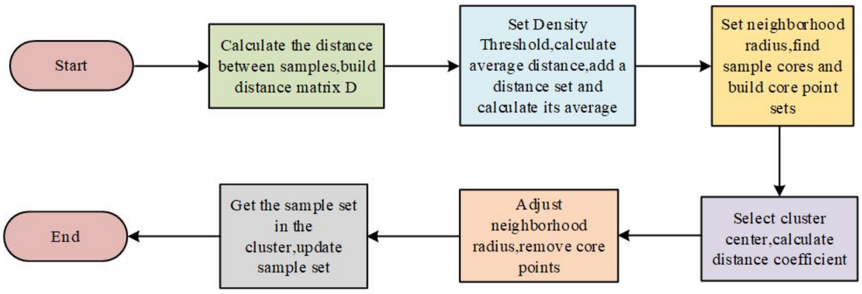


Fig. 3.2: Flow chart of DBSCAN algorithm based on distance optimization

services, and finally achieve the purpose of improving the learning efficiency and effect of learners. Faced with huge and complex online teaching students' academic behavior data, they need to be preprocessed. In 1996, Ester et al. proposed the DBSCAN algorithm. This algorithm differs from other principled clustering algorithms in that it determines the density of the dataset in space by neighborhood [18]. Another way to make the data set to achieve the clustering effect is to describe the density of the data point set in the space. The sample set is set as the neighborhood is for the hypersphere area with the center and the radius. In this area, all sample points constitute a sub-sample set, which satisfies equation 3.1.

$$N_\delta\left(x_j\right)=\left\{x_j \in A \mid \operatorname{distance}\left(x_i, x_j\right) \leq \delta\right\} \tag{3.1}$$

Improves some problems that may exist in the traditional DBSCAN algorithm, and proposes a distance-optimized DBSCAN algorithm. Set the sample set as $A=x_1, x_2, \ldots, x_n$, there is equation 3.2.

$$N\left(x_i\right)=\left\{x_j \in A \mid 0<d\left(x_i, x_j\right)\ <\delta\right\} \tag{3.2}$$

For $x_i \in A, x_i$ the density $N(x_i)$ is the number of data points in x_i the δ neighborhood of. The distance coefficient of the core point is shown in equation 3.3.

$$\theta=\frac{N\left(x_j\right)}{N\left(x_i\right)} \tag{3.3}$$

In Equation 3.3, θ represents the distance coefficient. Therefore, the basic flow of the DBSCAN algorithm based on distance optimization is shown in Figure 3.2.

In Figure 3.2, the algorithm flow can be divided into nine steps. First, the distance between the sample points needs to be calculated to construct a matrix, as shown in equation 3.4.

$$D=\left\{D_{ij} \mid i, j \in R, i \neq j\right\} \tag{3.4}$$

In equation 3.4, D represents a matrix, i and j are two points, and R represents a sample data set. Then given the density threshold, calculate the average distance of the points with the nearest MinPts number, and put them into the distance set, and then find the overall average of the set. Then set the neighborhood radius to half the average distance. Find all core points among all sample points and put them in the core object set. Then select a core point for the cluster center to form a new cluster. Then calculate the distance coefficient between the sample point and the core point to adjust the neighborhood radius. Starting from the sample points in each neighborhood, recursively, using the same method, all the sample points that are density-reachable, and put them into clusters. Then the clusters that have been found are removed and the sample set is updated. Repeat the search and adjustment until all core points are traversed or removed. Finally, the result is obtained, as shown in equation 3.5.

$$C=C_1, C_2, \ldots, C_n \tag{3.5}$$

In equation 3.5, it C represents the whole cluster, C_1 , C_2, and C_n represent the sample points that meet the conditions. In 1987, the silhouette coefficient was proposed. It can evaluate different algorithms with the same original data. And for different parameter indicators of an algorithm, the clustering results can also be evaluated. Assuming that the clustering result has K clusters, and each cluster has I samples, then a sample x is shown in equation 3.6.

$$\begin{cases} a(x) = \frac{1}{i-1} \sum_i \text{distance}(x, x_i), \quad x \neq x_i \\ b(x) = \frac{1}{j} \sum_j \text{distance}(x, x_j), \end{cases} \tag{3.6}$$

In equation 3.6, within the same cluster, the average distance between the sample x and the rest of the samples x_i is represented by, and the average distance $a(x)$ from the nearest point is represented by. Then the silhouette coefficient of the sample x is shown in equation 3.7.

$$A(X) = \frac{b(x) - a(x)}{\max(a(x), b(x))} \tag{3.7}$$

In order to obtain the silhouette coefficient of the overall clustering results, $A(X)$ the average value, $A(X) \in [-1, 1]$.

3.2. Student behavior feature extraction and N-Adaboost model construction. This paper collects students 'school behavior data and network behavior data, and preprocesses the behavioral data, and builds a database of students' network behavior, network viscosity and life rules. The data in the database include students' network traffic data, network authentication data, canteen consumption data, library entry and exit data, dormitory access control data, the names of the courses, course attributes, credits, final examination scores, course semester and academic year of the course. At the same time, the study also desensistudent behavior logs, including electronic account desensitization, domain name desensitization, IP address desensitization, name desensitization, and student number desensitization. Desensitizing the five types of fields of student log information data can obtain an encrypted string with source data rules. This kind of string hides the user's privacy information, which ensures the data security, but also has certain recognition and readability. After desensitizing the school behavioral data, data preprocessing is required. Preprocessing includes data cleaning, data protocol, and data transformation. Study on the preprocessed data to construct a student "portraits" database. The study summarized the data into three aspects, namely, network behavior, network viscosity, and life regularity. Network behavior includes network behavior index, network viscosity includes network viscosity index, and life regularity includes canteen consumption data index and library learning index. Among them, the canteen consumption data and the library learning index jointly describe the offline behavior of students, and reflect the regularity of students' life in life. Therefore, this paper summarizes the two into the same item, namely the "regularity of life". According to the students "portrait" description index, the study using DBSCAN algorithm based on distance optimization of the student cluster, will be divided into different groups with different performance differences, to explore different academic performance, students daily network behavior and school behavior by distance optimization of DBSCAN algorithm after clustering students, can be intuitive analysis learning factors. From a macro perspective, the characteristic indicators with significant influence were found out, and the ANOVA F-test significance test was conducted on them. The F-test test value was used to characterize the influence degree of the influencing factors on the learning situation, and the characteristics of behavioral data were extracted to provide a data basis for the learning situation evaluation model. The influencing factors of students' learning are shown in Figure 3.3.

From Figure 3.3, the influencing factors of students' academic level can be mainly divided into three aspects: network behavior, network viscosity, and life regularity. Among them, network behaviors are divided into five types according to the types of access resources: video, knowledge, game, shopping, and social. In terms of network viscosity, after analysis, students' academic level is more affected by online time and online time. Considering the campus and social environment, the number of online days is not significantly different among different student groups, and the impact on the academic level is small. In terms of the regularity of life, it is believed that the consumption of breakfast can indicate whether students get up early, so it can show the regularity of students' life to a certain extent. ANOVA is a method to test the significance of the

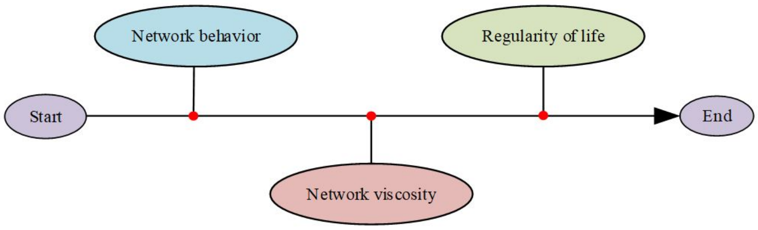


Fig. 3.3: Factors affecting students' academic level

difference between the means of two or more samples [19]. This method takes the F- distribution as the basis of the probability distribution and estimates the F value by the component and within-group mean square value calculated by the sum of squares and degrees of freedom. Then analyze the contribution of variation from different sources to the total variation, to determine the influence of different variables on the research results. Applied ANOVA F test to extract the characteristics of different student groups can understand the extent of different factors affect the evaluation of foreign language learning. In the F test, the study used the foreign language learning evaluation between different groups of students as the dependent variable, and the network behavior factors, network viscosity and life pattern as independent variables. Comparing the differences in variance between different student groups to determine which factors significantly influenced students' evaluation of foreign language learning. In ANOVA F-test, the sum of squares of the total deviation (Sum of Squares Total, SST) is the sum of squares of the error of the variable to the mean of the total sample, as shown in equation 3.8.

$$SST=\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}\right)^{2} \tag{3.8}$$

In equation 3.8, y_{ij} denotes the variable and $\overline{y}$ denotes the mean of the total sample. The Sum of Squares Regression (SSA) between groups represents the error sum of squares of the mean of each group to the total mean, as shown in equation 3.9.

$$SSA=\sum_{i=1}^{k}n_i\left(\bar{y}_l-\bar{y}\right)^{2} \tag{3.9}$$

In equation 3.9, $\overline{y_i}$ represents the mean of each group. The Sum of Squares Error (SSE) within the group represents the sum of squares of the error between the sample data of each group and the mean of the group, as shown in equation 3.10.

$$SSE=\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_l\right)^{2} \tag{3.10}$$

According to the content of the appeal, it can be known that the mean variance between groups MSA, the mean variance MSE within the group, and the F value, as shown in equation 3.11).

$$\begin{cases} MSA=\frac{SSA}{k-1} \\ MSE=\frac{SSE}{k(n-1)} \\ F=\frac{MSA}{MSE} \end{cases} \tag{3.11}$$

In 1996, Adaptive Boosting was proposed. The algorithm can automatically adjust the weight value according to the resulting feedback of the learner to adapt [20, 21]. It ensures that in the continuous iterative process, the classifier can gradually focus on those samples that are difficult to classify, which improves the classification accuracy. However, because the classifier types are the same, there are certain limitations. Therefore, the research proposes a multi-classifier-based N-Adaboost model. Instead of using a single classifier as the base learner, the model integrates multiple classifier models to avoid the problem that homomorphic classifiers only perform well in one aspect, and underperformance in multiple problems. During the training process, the base learner is composed of multiple classifier models, each classifier model classifies the training samples, and the training results of the base learner are determined by the multiple classifier models. In each iteration, the training sample dataset successively passes through multiple classifier models to fit the model with the same weight. By integrating different classifier models, the model overcomes the classification limitations brought about by a single learner and makes the performance of the classifier complementary. The addition of N-Adaboost model can improve the universality of the online course evaluation model and solve the problem of insufficient performance in analyzing a large number of students and multiple online courses. Let the number of iterations be N, the training data set be D_i , and the weight distribution of the initialized training samples is shown in equation 3.12.

$$\begin{cases} W_1 = (w_{1,1}, w_{1,2}, \ldots, w_{1,i}) \\ w_{1,i} = \frac{1}{N}, \quad i = 1, 2, \ldots, N \end{cases} \tag{3.12}$$

Build a learning algorithm consisting of Γi classifiers $H(x)$, and then make classification evaluations on the training data. And count the classification results, and return the final classification results to the algorithm after processing. The learning algorithm is a model trained through the training data set, and input into the ensemble classifier model to obtain a weak classifier $G_n(x)$, as shown in equation 3.13.

$$G_n(x) = \Gamma(D_i, W_n, H(x)) \tag{3.13}$$

In equation 3.13, W_n represents the weight, where $n = 1, 2, , \ldots, n$. According to the classification error rate at this time, the weight in the strong classifier is calculated, as shown in equation 3.14.

$$\alpha_n = \frac{1}{2} \log \frac{1 - e_n}{e_n} \tag{3.14}$$

In equation 3.14, e_n represents the classification error rate. Then update the weight distribution of the training sample set, as shown in equation 3.15.

$$\begin{cases} w_{n+1,i} = \frac{w_{n,i}}{z_n} \exp\left(-\alpha_n y_i G_n\left(x_i\right)\right), \quad i = 1, 2, \ldots, N \\ z_n = \sum_{i=1}^{N} w_{n,i} \exp\left(-\alpha_n y_i G_n\left(x_i\right)\right) \end{cases} \tag{3.15}$$

In equation 3.15, Z_n represents the normalization factor that can make the probability distribution sum of the samples equal to 1. Repeat the previous steps N times to obtain the final classifier as shown in equation 3.16.

$$F(x) = \operatorname{sign}\left(\sum_{i=1}^{N} \alpha_n G_n(x)\right) \tag{3.16}$$

When a model construction is completed, there needs to be a parameter to evaluate the relevant performance of the classification model, and the relevant parameters are called the model evaluation index. Different models have different tasks, and the corresponding ones can be evaluated with different indicators. The study predicted students' academic level mainly based on student behavior and belongs to a classification task. The performance evaluation indexes used in the study were accuracy, recall rate and F1 value. Accuracy refers to the percentage of the number of samples correctly classified by the classifier to the total number of samples, reflecting the situation that the classifier correctly identifies each sample. The calculation formula is shown in equation 3.17.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.17}$$

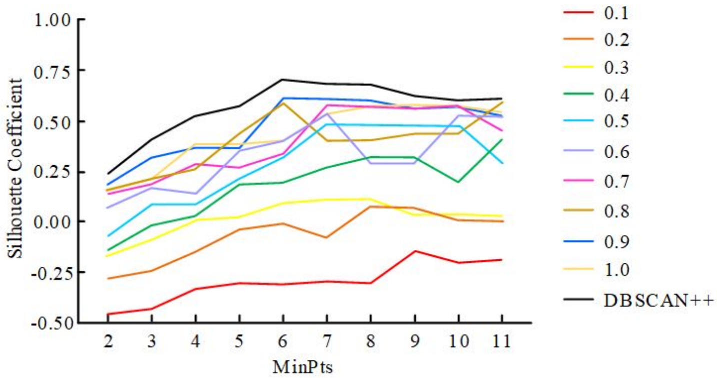


Fig. 4.1: Contour coefficient comparison diagram of traditional and optimized DBSCAN algorithm

In equation 3.17, TP represents true positive; TN represents true negative, FP and FN represent false positive and false negative. Accuracy refers to the number of true cases in the sample where the prediction result is positive, and the calculation formula is shown in equation 3.18.

$$Precision = \frac{TP}{TP + FP} \tag{3.18}$$

Recall refers to the percentage of positive predictions, with the formula shown in equation 3.19.

$$Recall = \frac{TP}{TP + FN} \tag{3.19}$$

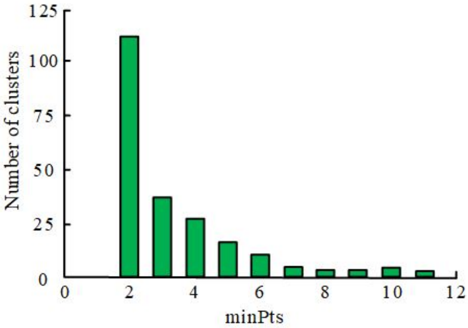
In order to better evaluate the performance of the classifier, the precision and the recall rate are called the measure, and the calculation formula is shown in equation 3.20.

$$F_\alpha = \frac{(1+\alpha^2) \cdot \text{Precision} \cdot \text{Recall}}{\alpha^2 \cdot \text{Precision} \cdot \text{Recall}} \tag{3.20}$$

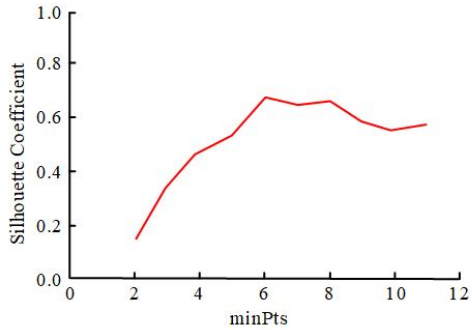
In equation 3.20, α is a non-negative real number. When α is 1, it is F1, which is the harmonic mean of precision and recall.

4. Analysis of the evaluation results of the evaluation model of students' foreign language learning. The data set of the study is the behavior data of students in a university in Henan province, which includes students 'network traffic data, network authentication data, campus "one-card" data and students' foreign language course test scores. To evaluate the clustering effect and determine the optimal number of clusters, the traditional DBSCAN algorithm and DBSCAN algorithm based on distance optimization were respectively used to cluster the data sets, and the corresponding contour coefficients of a different number of clusters were obtained, as shown in Figure 4.1.

According to the data in Figure 4.1, it can be seen that when δ is 0.1 and minPts is 2-11, the value of the silhouette coefficient is always below 0, indicating that the clustering effect is the worst at this time. When δ is 0.2 and 0.3, and minPts is 2-11, the value of the silhouette coefficient is basically below 0, but it can be around 0 when it is very few. When δ is in the range of 0.4-1.0, the value of the silhouette coefficient is constantly changing and can approach or even exceed 0.5 when the individual minPts is taken. Among them, when and when δ is 0.9 and minPts is 6, the contour coefficient exceeds 0.6, and the optimal parameter solution 0.661 is obtained at this time. In Figure 4.1, "DBSCAN + +" represents the distance optimized DBSCAN algorithm, the contour coefficient of the algorithm is always positive, the maximum value was 0.69, and the contour coefficient curve of the algorithm is always higher than that of the traditional DBSCAN algorithm. In conclusion, the results show that the clustering effect of the proposed modified DBSCAN algorithm is better.



(a) The number of clusters under different values of minPts



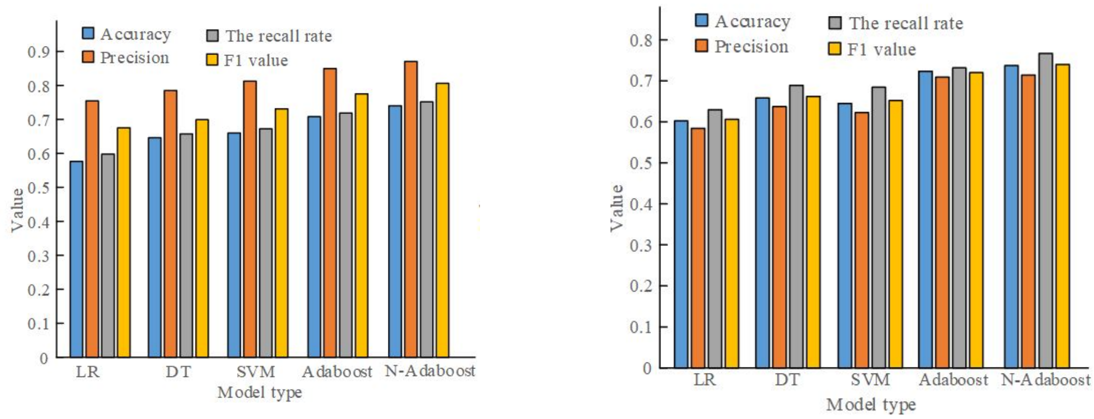
(b) Contour coefficients at different values of minPts

Fig. 4.2: Value of DBSCAN algorithm after distance optimization under different minPts

Therefore, the distance-optimized DBSCAN algorithm is used to cluster the students' network behavior. The changes in the number of clusters under different minPts values are shown in Figure 4.2.

In Figure 4.2, the contour coefficient increases with minPts, and after reaching the highest value, it begins to decrease. At the minPts value of 6, the modified DBSCAN algorithm has the highest contour coefficient of 0.72, which is 0.06 higher than the highest round contour coefficient of the traditional DBSCAN algorithm, when the number of clusters is 4. In conclusion, the results show that the modified DBSCAN algorithm clusters the results best when the minPts value is 6 and the number of clusters is 4. In order to predict the risk of academic level, the study divided students into four groups: pass, fail, excellent and non-excellent. For each group of students, they use the decision tree model, SVM, Adaboost model, and the multi-classifier-based heteromorphic N-Adaboost model. SVM and decision tree model were used as a control benchmark model to verify the effectiveness of N-Adaboost model. Each model was trained and evaluated by ten-fold cross-validation. Both the training set and the test set of the model were randomly divided, and the prediction operation was repeated ten times. The average value of the accuracy and F1 measure were taken as the prediction accuracy of the final model and the F1 measure of the results. The classification measurement results of each model in the passing group, failing group, excellent group and nonexcellent group are shown in Figure 4.3.

Figure 4.3a shows the classification measurement results of each passing group and failure group of the comparison model. As shown in Figure 4.3a, the accuracy of SVM and decision tree benchmark model is 57.63%, while the evaluation accuracy of the N-Adaboost model proposed by the study is 74.02%, which is higher than the benchmark model and also higher than other comparison models, and its accuracy performance is optimal. The accuracy rate of SVM and decision tree benchmark model is 75.46%, while the evaluation accuracy rate of N-Adaboost model is 87.05%, which is higher than other comparison models and has the optimal accuracy performance. Meanwhile, the recall rate and F1 value of the N-Adaboost model were 75.16% and 0.806, respectively, which are higher than the benchmark model and the comparison model, and the proposed N-Adaboost model has the best performance. Figure 4.3b shows the classification measurement results of each comparison model for the excellent group and the non-excellent group. As shown in Figure 4.3b, the accuracy of SVM and decision tree benchmark model is 64.3%, while the evaluation accuracy of the N-Adaboost model proposed by the study is 73.74%, which is higher than the benchmark model and also higher than other comparison models, and its accuracy performance is optimal. The accuracy rate of SVM and decision tree benchmark model is 60.23%, while the accuracy rate of the proposed N-Adaboost model is 68.13%, which is higher than the best accuracy performance. Meanwhile, the recall rate and F1 value of N-Adaboost model were 74.32% and 0.716, respectively, which are higher than the benchmark model and the comparison model, and the proposed N-Adaboost model has the best performance. In conclusion, these results show that the proposed N-Adaboost model performs better than the other contrast models."N" represents the species of individual



(a) Pass and fail

(b) Merit and Non-Merit categories

Fig. 4.3: Evaluation of the learning outcomes of the pass and fail groups and the excellent and non-excellent groups under different models

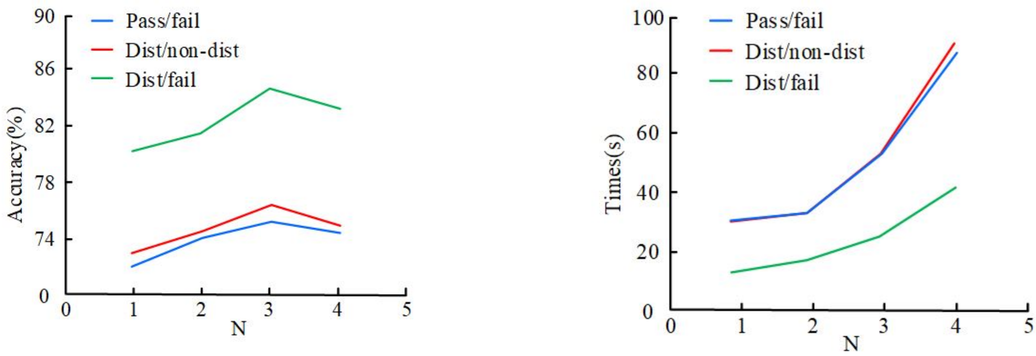
Table 4.1: Evaluation Results of the N-Adaboost Model for Different Values of N

	N=1	N=2	N=3	N=4
Accuracy	0.7138	0.7329	0.7441	0.7362
Precision	0.8535	0.8634	0.8759	0.8674
The recall rate	0.7231	0.7434	0.7474	0.7441
F1 value	0.783	0.799	0.807	0.801

learners in the modified multi-classifier-based N-Adaboost model. However, the variety of individual learners is not the more the better. According to the sample data of the " pass and fail group ", the research explores "N" in the N-Adaboost model and compares the model evaluation performance of N=1, N=2, N=3, and N=4. Among them, N=1 is the traditional Adaboost model, and its base classifier is composed of a decision tree model; N=2 is the N-Adaboost model used in the evaluation experiment, and its base classifier is composed of a decision tree model and an SVM model; N=3 Set the base classifier in N-Adaboost model to decision tree model, SVM model and Naive Bayes model (NB); N=4 set the base classifier in N-Adaboost model to decision tree model, SVM model, Naive Bayes model Yess model and logistic regression model. The results are shown in Table 4.1.

From Table 4.1, when N changes from 1 to 3, the evaluation accuracy is improved to a certain extent. The accuracy of N=2 is 1.91% higher than that of N=1, and the accuracy of N=3 It is 1.12% higher than N=2. However, when N=4, it is found that the accuracy rate has declined, and the accuracy is like that when N=2 and F1 is also reduced accordingly. Similarly, the same method is used to predict the samples of the "excellent and non-excellent group ""Pass and fail group"and "excellent and failing group ", and the results are shown in Figure 4.4.

Figure Figure 4.4a is the accuracy change curve of the three groups of samples under different N conditions. For excellent group and not excellent group, qualified group and unqualified group, excellent group and failing group, when N=3, the evaluation accuracy reached the highest, indicating that the more obvious the difference between the sample data, the higher the prediction ability of the model.(B) For the time-consuming change curve of three groups of samples under different N conditions. As the individual classifier types increase, the time consumption also gradually increases. When N=4, the highest time reached each group. Considering the balance



(a) Accuracy of Prediction

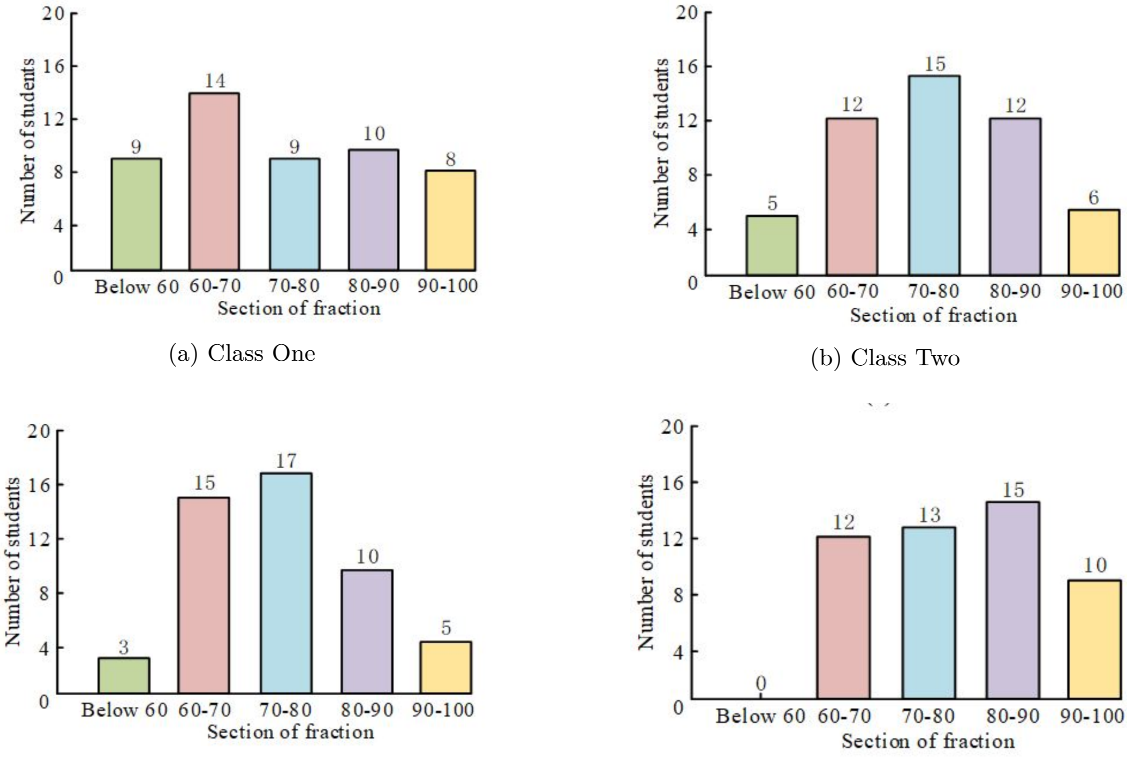
(b) Elapsed Time

Fig. 4.4: Evaluation results of different models for excellent and non-excellent students

between the accuracy and the evaluation time, when N is 2 or 3, the operation time is also controlled within a reasonable range while ensuring the evaluation accuracy of the N-Adaboost model. Four classes of 50 students were selected to evaluate the online foreign language learning of a total of 200 students. Firstly, students' overall learning of the semester was evaluated based on their classroom performance, homework completion, and exam results, with score ranging from 0 to 100 points. The evaluation results consider the students 'understanding and logic, students' classroom performance, students' classroom understanding and summary ability, and the mastery of knowledge and skills. The comprehensive evaluation results of the students are shown in Figure 4.5.

Figure Figure 4.5a shows the study situation of Class 1. According to figure, 4.5a, in Class 1, students with 60-70 account for 28%, students 90-100 account for 16%, and students below 60 account for 18%. Figure 4.5b shows the study situation of Class 2. From Figure 4.5b, in Class 2, students with points of 70-80 account for 30%, and students with points below 60 and 90-100 account for 10% and 12%. Figure 4.5c shows the study situation of Class 2. From Figure 4.5c, in Class 3, students with points of 60-80 account for 64% of the class, and students with points below 60 and 90-100 account for 6% and 10%. Figure 4.5d shows the study of Class 2, as shown from Figure 4.5d. There are no students with scores below 60 points in class 4, and the other four grades are evenly distributed and have the best grades. To grasp students' learning status in more detail and adjust the teaching plan in time, the researchers conducted a detailed evaluation of students' listening, speaking, reading, and writing abilities through the model, as shown in Figure 4.6.

Figure 4.6a shows the listening learning situation of students in each class. From Figure 4.6a, it can be seen that in terms of listening, the listening ability of students of the four classes is on the rise along with the learning progress, indicating that online teaching can improve students' foreign language listening level. Figure 4.6b shows the oral learning of students in each class. As can be seen from Figure 4.6b, we can see that the oral ability of students in Class 1 is constantly improving, the oral ability of Class 2 and Class 4 remains unchanged, and the oral ability of Class 3 is declining. Figure 4.6c shows the learning situation of students in each class in reading. As can be seen from Figure 4.6c, the reading ability of each class increases with the learning progress. Figure 4.6d shows the learning situation of students in each class, According to Figure 4.6d, it can be seen that the writing ability of the four classes has decreased, and the study analyzes the causes of this phenomenon. It is found that the teaching plan has some deficiencies and the lack of relevant writing practice. When teachers assign classroom tasks, they are more inclined to cultivate the "listening", "speaking" and "reading" modules, and the writing tasks are assigned less. At the same time, the lack of students' vocabulary also leads to the decline of the writing ability of each class. In view of this problem, the study put forward corresponding solutions, such as students making personal learning plans, teachers regularly arrange thematic writing tasks and timely feedback and guidance, the school arranged writing competitions, learning communication activities, etc. Through the continuous practice, improve the students' writing ability. Vocabulary is an important part



(a) Class One

(b) Class Two

(c) Class Three

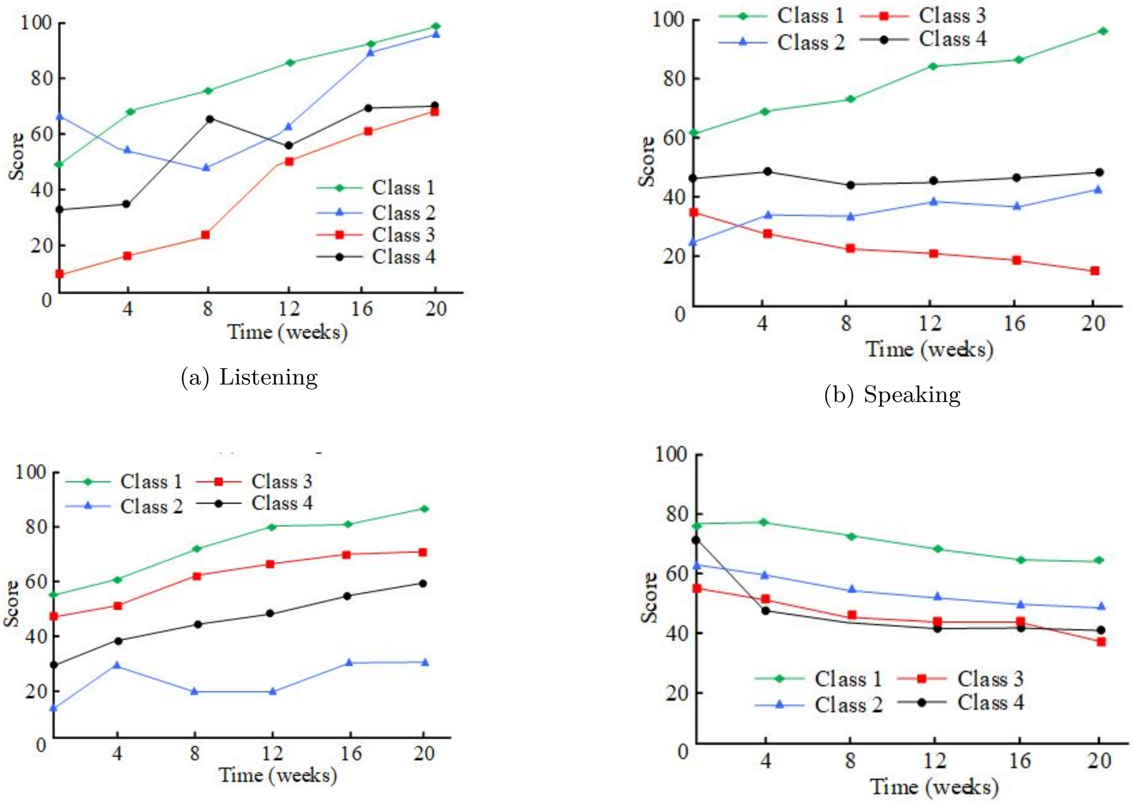
(d) Class Four

Fig. 4.5: Student learning overall evaluation results

of students' foreign language learning. Through the online teaching collaboration platform, teachers not only help students to accumulate vocabulary in class but also require students to independently accumulate at least 30 words a day through the online platform. After one semester, the evaluation model is used to evaluate group A and group B. The vocabulary of a total of 1000 students was investigated, as shown in Figure 4.7.

As can be seen from Figure 4.7, the foreign vocabulary learning of students in each group increased with the learning progress, and the vocabulary of students in Group A increased from less than 1000 to more than 6000. The vocabulary of students in group B increased from 500 to above 6000. In summary, the results show that the evaluation model of students' foreign language learning status for the online teaching collaborative platform can help students to improve their learning situation.

5. Conclusion. The rapid development of Internet technology has brought a new revolution to the traditional offline teaching mode. Aiming at the problem that the traditional teaching quality evaluation system is no longer suitable for modern information-based teaching evaluation, this paper constructs an evaluation model of students' foreign language learning based on the online teaching cooperation platform. DBSCAN algorithm based on distance optimization is used to cluster students' learning behaviors. Based on the traditional Adaboost model, the classification of the base classifier is improved and the N-Adaboost algorithm is proposed. Experiments show that the clustering effect and accuracy are improved by 8.7%, and the comprehensive performance is better than the traditional DBSCAN algorithm. The evaluation accuracy of the N-Adaboost model based on multiple classifiers in the two groups of the pass and fail, excellent and non-excellent is 74.02% and 73.74%, respectively, and the overall performance has been improved. In addition, when N is 2 or 3, the balance between accuracy and evaluation time can be considered. Based on the performance of the class, the completion



(a) Listening

(b) Speaking

(c) Reading

(d) Writing

Fig. 4.6: Student vocabulary changes over time

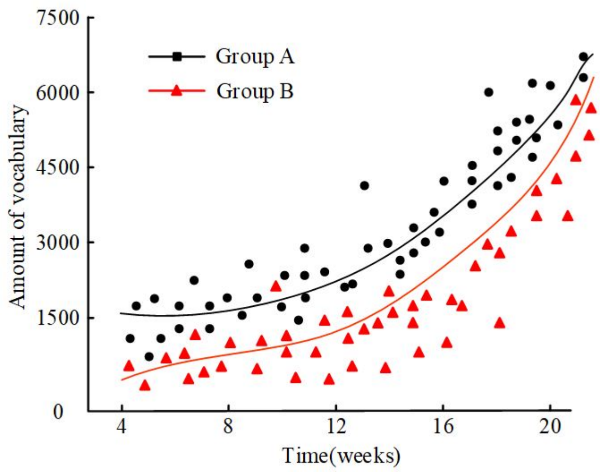


Fig. 4.7: Student vocabulary changes over time

of homework, and the examination results, the learning situation of the four classes was evaluated. There were no failed students in the four classes, and the number of students with 90-100 scores was the largest. Class 1 had the worst overall situation, with 9 students below 60 points and 14 students between 60 and 70 points. The students' listening, speaking, and reading abilities showed an overall improvement, but their writing ability declined seriously, dropping to below 40 points at 20 weeks. In view of this problem, the study put forward corresponding solutions, such as students making personal learning plans, teachers regularly arrange thematic writing tasks and timely feedback and guidance, the school arranged writing competitions, learning communication activities, etc. Through the continuous practice, improve the students' writing ability. The proposed evaluation model of students' foreign language learning for online teaching collaboration platform has certain potential in practical application. This model can accurately evaluate students 'semester situation, and then provide guidance and feedback for teachers, promote students' independent learning, provide students with personalized learning guidance and support for students, optimize teaching design and resource allocation, promote educational reform and innovation, and then improve the quality of education. However, the application of the evaluation model still faces some limitations, the reality is often more complex and changeable, and the evaluation model needs to solve the problems of subjective evaluation factors and the difficulty of comprehensive evaluation. At the same time, the samples selected in this study also have some limitations, such as the under representation of the sample, which still needs to further explore the students who do not have the universal education information sharing platform. The future research direction is to establish a distributed cluster environment using high-performance platform and realize the parallelization of student behavior data processing and computing. The effectiveness verification of the proposed intervention based on the evaluation model proves that the evaluation model improves the teaching quality. At the same time, the research can also add more factors to the evaluation mode, and constantly optimize the online teaching mode, such as students 'participation in class, and evaluate the degree of students' participation and interaction in class, including ask questions, answer questions, discussion and cooperation. And students 'learning experience, such as assessing students' satisfaction and experience of online teaching mode, such as students' willingness to participate, learning motivation and learning interest. By comprehensively considering the factors of students 'classroom participation and students' learning experience, the research can more comprehensively evaluate the advantages and disadvantages of online teaching mode, and continuously optimize the teaching process to improve students' learning effect and learning experience.

Funding. The research is supported by the Henan Provincial Teaching Reform Research and Practice Project "Research and Practice on the Teaching Ecosystem of General Academic English in Science and Engineering Universities under the Double First Class Background" (No. 2019SJGLX060).

REFERENCES

[1] Wang, Y. On College English Teachers' Basic Skills of Classroom Teaching and Evaluation and the Promotion Strategies. *International Journal Of Social Science And Education Research.* **5**, 93-96 (2022)

[2] Gao, H. Application of Iwrite English Writing Teaching and Appraising System in College English Teaching. *International Journal Of Social Sciences In Universities.* **5**, 255-259 (2022)

[3] Jiao, F., Song, J., Zhao, X., Zhao, P. & Wang, R. A spoken English teaching system based on speech recognition and machine learning. *International Journal Of Emerging Technologies In Learning (iJET).* **16**, 68-82 (2021)

[4] Yang, Z. & Feng, B. Design of key data integration system for interactive English teaching based on internet of things. *International Journal Of Continuing Engineering Education And Life Long Learning.* **31**, 53-68 (2021)

[5] Susanty, L., Hartati, Z., Sholihin, R., Syahid, A. & Liriwati, F. Why English teaching truth on digital trends as an effort for effective learning and evaluation: opportunities and challenges: analysis of teaching English. *Linguistics And Culture Review.* **5** pp. 303-316 (2021)

[6] Liu, H., Chen, R., Cao, S. & Lv, H. Evaluation of college English teaching quality based on grey clustering analysis. *International Journal Of Emerging Technologies In Learning (iJET).* **16**, 173-187 (2021)

[7] Li, N. A fuzzy evaluation model of college English teaching quality based on analytic hierarchy process. *International Journal Of Emerging Technologies In Learning (iJET).* **16**, 17-30 (2021)

[8] Zhang, Y. The development of an evaluation model to assess the effect of online English teaching based on fuzzy mathematics. *International Journal Of Emerging Technologies In Learning (iJET).* **16**, 186-200 (2021)

[9] Tqt, T., Tmn, N., Luu, T. & Pham, T. An evaluation of English non-majored freshmen's attitude towards EFL learning at the Can Tho University of Technology. *International Journal Of TESOL & Education.* **1**, 72-98 (2021)

[10] Wang, H. & Cui, J. Evaluation of teaching effect in higher educational institutions and identification of its influencing factors. *International Journal Of Emerging Technologies In Learning (iJET)*. **16**, 226-239 (2021)
[11] Li, Y. & Wang, X. The algorithm and implementation of college English teaching comprehensive ability evaluation system//EAI International Conference, BigIoT-EDU. (216-224,2021)
[12] Sun, Z., Anbarasan, M. & Praveen Kumar, D. Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence*. **37**, 1166-1180 (2021)
[13] Lu, C., He, B. & Zhang, R. Evaluation of English interpretation teaching quality based on GA optimized RBF neural network. *Journal Of Intelligent & Fuzzy Systems*. **40**, 3185-3192 (2021)
[14] Fang, C. Intelligent online English teaching system based on SVM algorithm and complex network. *Journal Of Intelligent & Fuzzy Systems*. **40**, 2709-2719 (2021)
[15] Noor, A., Shahid, A., Ahmed, S. & Ahmad, M. An Evaluation of Communicative Language Teaching in Pakistan: A Study of Undergraduate English Learners of Pakistan. *Pakistan Journal Of Humanities And Social Sciences*. **9**, 259-264 (2021)
[16] Zhang, Y. & Yang, Y. The evaluation method for distance learning engagement of college English under the mixed teaching mode. *International Journal Of Continuing Engineering Education And Life Long Learning*. **32**, 159-175 (2022)
[17] Clayson, D. The student evaluation of teaching and likability: what the evaluations actually measure. *Assessment & Evaluation In Higher Education*. **47**, 313-326 (2022)
[18] Cook, C., Jones, J. & Al-Twal, A. Validity and fairness of utilising student evaluation of teaching (SET) as a primary performance measure. *Journal Of Further And Higher Education*. **46**, 172-184 (2022)
[19] Okoye, K. & Arrona-Palacios, A. Camacho-Zu niga C, Joaquín Alejandro Guerra. (2022) Towards teaching analytics: a contextual model for analysis of students' evaluation of teaching through text mining and machine learning classification. *Education And Information Technologies*. **27**, 3891-3933 (0)
[20] Marshall, P. Contribution of open-ended questions in student evaluation of teaching. *Higher Education Research & Development*. **41**, 1992-2005 (2022)
[21] Marks, B. & Thomas, J. Adoption of virtual reality technology in higher education: An evaluation of five teaching semesters in a purpose-designed laboratory. *Education And Information Technologies*. **27**, 1287-1305 (2022)

Edited by: Mudasir Mohd
Special issue on: Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions
Received: May 17, 2023
Accepted: Nov 1, 2023