EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Design of a Big Data-based Evaluation System for Teacher-Student Cooperation in College English Writing

Jia Li[1,*] and Jinpeng Zhang[1]

[1] English Department, School of English Language & Literature, Xi'an Fanyi University, Xi'an 710105, Shaanxi, China
Author's address
[2] The Security Affairs Office, Xi'an Fanyi University, Xi'an 710105, Shaanxi, China

Abstract

INTRODUCTION: With the constant change in information technology, people have entered the era of big data, and the reforms brought about by the era of big data have profoundly affected people's way of life and learning. English writing is one of the four basic skills for mastering English, and the skilful mastery of writing is an essential form of accurate expression in English. Different from the traditional teaching modes of "separate assessment of teaching" and "separate assessment of learning", "cooperative assessment between teachers and students" is a new type of assessment mode proposed by professors at the National University in recent years. This cooperative evaluation model advocates the combination of the teaching and evaluation processes and the evaluation and learning processes. In the cooperative evaluation model, evaluation is considered an extension of the teaching process, and the evaluation process itself is also a school process. The arrival of the significant data era also brings new opportunities and challenges for university English teacher-student cooperative evaluation.
OBJECTIVES: We design a university English writing teacher-student cooperative evaluation system based on big data by combining big data with the teacher-student cooperative evaluation.
METHODS: All college English writing classes in one university were selected as the background for ample data research, and with the help of theoretical knowledge and technology related to big data, a model was constructed to analyze the factors affecting the cooperative evaluation system of college English writing teachers and students.
RESULTS: The university English writing teacher-student cooperation evaluation system is analyzed and summarized using big data.
CONCLUSION: By using big data to analyze the evaluation of college English writing teacher-student cooperation, big data can better help teachers understand the weaknesses of students' knowledge points and better help students improve their English proficiency.

Keywords: big data, BP neural network, college physical test, performance prediction

Received on 19 January 2023, accepted on 21 August 2023, published on 4 September 2023

doi: 10.4108/eetsis.3830

*Corresponding Author. Email: 20347305@qq.com

1. Introduction

Since McKinsey first proposed the concept of big data, people have been paying more and more attention to the concept of big data, and people are gradually recognizing the value embedded in big data. Since the value of big data can be compared to the value of natural gas and oil, big data is being increasingly emphasized both domestically and internationally. Research on big data has had a significant impact on scientific advancement and commercial use, bringing about a drastic change in people's lives. Big data is often used for commercial purposes, and more research needs to be done on its

application to students' learning. The existence of big data can analyze the problems arising from students' learning habits and methods and provide new ideas for solving them(Giffen et al., 2022). Through the construction of an extensive data system, there can be a systematic assessment of student learning, using the most scientific data as a basis for the development of student learning. Higher education institutions have the responsibility to provide higher talents for society, and the application of big data technology in higher education to explore its potential resources and application value will play an essential role for higher education institutions to manage students better, cultivate students, and carry out a brand new intelligent change in higher education. In the current higher education system, college English exists in the learning process of every major as a compulsory course. Among them, English writing can reflect students' mastery of English to a large extent(Liang et al., 2021). However, in the current situation, college English writing classes face many problems, such as the fact that students are not able to produce high-quality essays, the content learned in class cannot be flexibly applied in writing, and some students will plagiarize to complete the composition(Hilali & Mckinley, 2021). The standard evaluation methods in college English writing courses often consist of four evaluation modes: teacher's evaluation, student's self-evaluation, peer evaluation, and machine scoring. Although the teacher evaluation mode has a better effect among the above four evaluation modes, the teacher evaluation will consume too much energy from the teacher, who cannot follow up on the writing situation of each student in real time. Due to the students' different attitudes towards the teacher's evaluation, they may need help rectifying the problems the teacher pointed out. Hence, the teaching results fail to achieve the expected results. The other two evaluation methods, student self-assessment and peer assessment, can only find some simple grammatical or spelling errors and cannot analyze the reasons for students' errors, and students' trust in these two evaluation methods needs to be improved (Meziane, 2021). The last type of evaluation, machine evaluation, can help students improve their writing level. However, the problems found by machine evaluation are more focused on superficial language errors, and there is no way to analyze the reasons why students produce problems in depth. Therefore, traditional teaching and evaluation methods often need to be more effective. Teacher-student cooperative evaluation is an excellent solution to the problems encountered in traditional evaluation. Unlike traditional evaluation, the teacher plays a leading role in teacher-student cooperative evaluation, and students are learning-centred during the whole learning process. Students can realize diversified evaluations of tasks with the help of the teacher, which can achieve the effect of learning while evaluating(Meziane, 2021). The use of big data technology in the design of a cooperative evaluation system for college students' English writing Teachers and students can process a variety of data, saving time and providing precise results. In addition, the use of big data technology in the exploration of college students' English teacher-student cooperative evaluation system, through the use of big data technology in the statistical analysis of the problems encountered by the student's writing, makes the data have a high degree of credibility, which can make the evaluation results of the teacher-student cooperative evaluation system have a higher degree of credibility and can be improved in time for the existing deficiencies.

2. Research Background

Although there are high requirements for listening, speaking, reading, and writing in university English teaching, English writing has recently been included as one of the subjects in the English examination in recent decades in the actual teaching process. Unlike listening, speaking, and reading, English writing has not attracted the attention of many teachers since its inception. However, with the acceleration of reform, opening up, and globalization, the practical use of English has been broadened. English writing has gradually come into people's view, attracting attention and becoming more and more critical in teaching English. The separate subject of English writing is mainly offered in the teaching process of English undergraduates and specialists. In contrast, English writing is not taught separately in the public English course of the university. Therefore, teachers cannot systematically prepare particular textbooks for English writing, so many students' English writing ability is low, and the English writing system could be better. Therefore, it takes work to improve the English writing of these students. Moreover, the traditional teaching methods can no longer meet the needs of teachers and students, and the problems in English writing teaching are becoming more and more prominent. For example, the content of the textbooks needs to be updated, the information in the textbooks is different from modern information, and it needs to be in line with modern society and compatible with modern teaching reform(Giesler, 2021). At the same time, in traditional education and teaching methods, the teacher is often used to lead the teaching, and the student's enthusiasm for learning is poor. Learning motivation in the traditional teaching mode cannot be stimulated. In daily teaching, teachers dominate and concentrate on students' theoretical learning after explaining the knowledge points, omitting students' role in learning. Writing is a process that requires long-term accumulation, and teachers need to provide timely guidance for students' essays and make timely corrections to the problems that arise in the writing process. According to the traditional teaching method, more than students' self-control is required. Teachers' energy is insufficient to grasp the level of all students' compositions, so students' English writing in the class and the classroom does not mean getting timely feedback on their problems which cannot be corrected in time. The level of writing cannot be improved, which affects the students' motivation to learn

and the results of the teaching and learning of this aspect of the English language. This affects the students' motivation to learn, affecting the results of English writing education.

With the development of the information age, the global popularity of online education is getting bigger and bigger. While generating a large amount of data, it promotes the rapid development of online education research. How to make these vast amounts of data better for research and get the law to find out the problem is an important issue that connects the direction of big data and education. The participation of new computer technology effectively improves the lack of relevant information in the field of education, providing a new program for a more convenient college education that is conducive to using objective data to help teachers teach students according to their abilities. Big data is essential to promote the development of education reform. Big data not only directly impacts education but also reshapes the education system as a whole, which has an essential impact on the decision-making and planning of the entire education system. Before the era of big data, the structure and organization of the education system relied mainly on the subjective judgment of educators and the subjective decisions of relevant leaders, and the education data generated in the education process was not fully utilized, making it challenging to identify the real problems encountered in the education process. Applying big data technology to promote education can provide more empirical evidence, a theoretical basis, and a decision-making basis for educational change(Yurkofsky, 2022). Through a large amount of data analysis and argumentation to improve the quality of education and the accuracy of educational teaching methods and educational technology, we can better help solve the problems in education. Educational data is relatively easy to collect, but the data analysis is not. People often use big data to answer an educational question that already has an answer and hope that big data can help put forward new, correct questions. The acquisition of big data is a long-term and complex process, so there is a need for a correct way to analyze educational data. Improving the quality of educational data analysis and refining the intrinsic value and laws from the collected data is a new way of thinking about considerable data-driven education. Analyzing learning in the context of big data is one of the most critical research areas in applying big data to the field of education, and it is an essential pillar of significant data-driven educational change. In the whole research process, the data generated in actual education should be used as the basis, and the technology's visibility, accuracy, and universality should be taken as the goal to build a technical framework applicable to the analysis of educational data. The better application of big data in daily teaching is a trend in the development of the whole education field, and the development of the education paradigm driven by big data is of great significance to education reform and innovation(Conn et al., 2022). Integrating learning technology and education in the context of big data, based on obtaining valuable data and relying on information technology and other technical means to apply the information to the actual educational situation, is critical in realizing the future of education and intelligent education.

Combining big data technology with the teacher-student cooperative evaluation system used more often nowadays and designing a new teacher-student cooperative system based on big data for college English writing will help teachers understand the problems in students' English writing more accurately and quickly and help students solve the problems in time. Students will gain more enthusiasm and initiative in learning, and their English writing ability will improve.

3 Methods and Materials

3.1 Clustering algorithm

Clustering algorithms are very versatile and are used in a variety of fields. In business, clustering analysis is used to help mall operators and salespeople analyze consumer demand and focus on future sales. In healthcare, clustering analysis can also effectively help medical professionals and hospital administrators, helping doctors to understand patients' episodes of illness, medication intake, and post-surgical recovery over time, making it easier to assign medical staff, understand the focus of purchasing medications, and make a scientific prognosis for patients' post-surgical recovery. Although the various types of cluster analysis algorithms are not the same, their features have a certain degree of overlap. Therefore, it is difficult to use a precise classification of sentence pairs to analyze the division(Andreis et al., 2021). The main classification methods used so far are as follows: division method, hierarchical method, density method, graph theory-based method, grid-based method, and model-based method.

3.1.1 Division method

The division method divides data collection containing N samples into K classification families (K < N). In the beginning, there are initial K clustering centres, and according to the principle that the closer the distance between similar records, the more distant the distance between dissimilar records, K is the centre of clustering using the distance priority principle. K clustering centres are constantly updated and improved until they finally get the optimal clustering division.

3.1.2 Hierarchical Approach

The hierarchical method refers to the continuous bottom-up or top-down decomposition of a given sample according to the desired level of the sample data until the conditions are met. The basic idea is to divide the data by

distance, density, and connectivity until a given condition is met.

3.1.3 Density Method

The density method refers to setting a threshold value for a given data interval before dividing the previous data and subsequently dividing the data according to the density. Whenever there is a region where the density of data points exceeds the threshold, it is divided into similar clusters.

3.1.4 Based on the graph theory approach

Based on the graph theory method, which refers to the nodes of the graph to characterize the data being classified and the edges of the graph to characterize the degree of similarity of the data being classified, build a graph corresponding to the sample data of the problem and to present the sample data as a graph, i.e., to be classified, objects are viewed as nodes in a fully connected undirected graph, G = [X, E], where X is the set of nodes and E is the set of edges; the distance between two points is the weight of the edge to generate the minimum tree; and a threshold is set to analyze the objects for clustering.

3.1.5 Grid-based method

The grid-based method is based on the space-driven idea of dividing the entire data space into several cells. Processing data in cells has a faster processing speed.

3.1.6 A model-based approach

The main idea of the model-based approach is to optimize the fitness of a given data object to the data model to achieve the most clustering. Usually, clustering assumes a numerical model to find similar sample data so that the sample and mathematical models match. Clusters are determined by constructing a density function that reflects the spatial distribution of data points or by statistical results.

3.2 *K*-means clustering algorithm

The *k*-means clustering algorithm is a division method and is an algorithm used for signal processing. Its most important function is to mine information. The core of his algorithm is to divide the objects of N datasets into K classes according to the principle of closest distance so that each object can be attributed to the nearest cluster, and each such algorithm should satisfy the following two conditions:
Clustering contains at least one sample of data.
The sample data belongs to only one cluster.

Given in advance, a set of objects ($x_1, x_2 \cdots, x_n$), each of which is an m-dimensional vector, is divided by an algorithm into k sets (k ≤ n), where each object in the set minimizes the sum of the squares of the differences between the objects in the centre of the set.

3.2.1 Basic flow of calculations

In the classification dataset, the initial step involves selecting K points as the initial centroids for the *K*-means clustering algorithm. Based on the principle of dividing data points into clusters based on their distances from the centre of K, it is observed that the smallest distance is allocated to the nearest K centres. In the presence of multiple data points with the same smallest distance, the new class points can be arbitrarily assigned to one of these centres. Subsequently, the average value of these points is calculated, which is considered the new clustering centre based on the average value of the nearest point (Kumar et al., 2021). If there are multiple nearest clustering centres, choose one as its new data centre, calculate the distance to its K clustering centre for each point, and then continue to divide each point according to the principle of the nearest to find the K clustering centre of all the data to divide the point of whether the change if there is a duplication, repeat step 2.3; if there is no change, the end of the division, as shown in Figure 1.

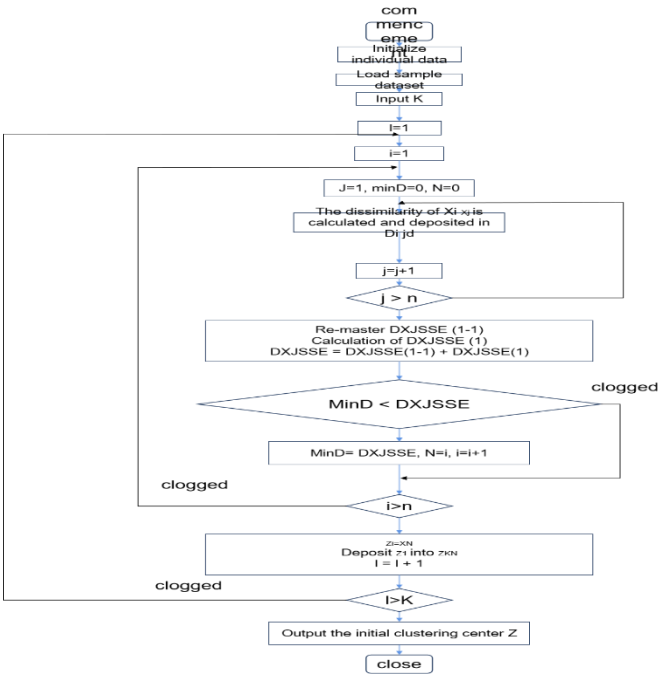

Figure 1. Flowchart of Clustering Center Implementation

3.2.2 *K*-means Clustering Algorithm Distance Calculation Equations

Let $X = \{x_1, x_2, \cdots, x_n\}$ be the set of sample $X = \{x_1, x_2, \cdots, x_n\}$ points and the sample of attributes$\{A_1, A_2, \cdots, A_m\}$. For the attribute that takes the

value of $Dom(A_i) = \{a_1^i, a_2^i, \cdots a_l^i\}(l \geq 2)$ any two sample points, the distance of the K-modes is given by

$$D(x_i, x_l) = \sum_{l=1}^{m} d(x_{il}, x_{jl}) \quad (2)$$

included among these $d(x_{il}, x_{jl}) = \begin{cases} 0, & x_{il} = x_{jl} \\ 1, & x_{il} \neq x_{jl} \end{cases}$

Determine the number of classifications K and randomly generate K clustering centres; respectively, calculate the distance from each sample point to each clustering centre, according to the principle of minimum distance, the sample point is classified as the closest class, if there is more than one closest centre, arbitrarily select one to be classified in it; respectively, calculate the average value of the distance of the sample data points to their corresponding centres in each clustering, based on the results of the calculation of the corresponding clustering centre; determine whether each clustering centre and samples are consistent with the last clustering results; otherwise, repeat the steps and steps to continue iterating. Clustering centre: Determine whether each clustering centre and samples and the last clustering results are consistent; if consistent, the clustering is complete, the end of the algorithm; otherwise, repeat the steps and steps to continue the iteration.

3.3 Anomalous Data Detection Methods

Anomalous data refers to some data that is different from other data in the data concentration area or is far away from the data concentration area(Su & Zhang, 2022). The existence of anomalous data may affect the accuracy of the work to some extent, cause some mistakes, and impact the work as a whole.

Anomalous data detection, also known as outlier data detection, is called de-drying in this whole process, i.e., detecting data that is distinguishable from other data, generally to eliminate noise and discover knowledge of potential value significance.

The basic idea is to calculate the sampling frequency by calculating the sample data based on the number of times each attribute appears in the attribute's value. Based on the frequency of the sample attribute, the smaller the frequency indicates, the more abnormal. The equation is as follows:

$$AVF(x_i) = \frac{1}{m}\sum_{j=1}^{m} f(x_{ij}) \quad (3)$$

Among them are the sample data, the sample data dimension, and $f(x_{ij})$ is the sample data in the attribute frequency in the attribute.

4. Results and Discussion

As shown in Figure 2, unfolding teacher-student cooperative evaluation in college English writing teaching is often implemented by the following three steps: the teacher chooses the evaluation focus and samples before the class and evaluates the samples; the teacher guides the students to carry out cooperative evaluation during the class; and the students carry out self-assessment according to the teacher's critique at the end of the class and then subsequently carry out mutual evaluation or machine evaluation, as shown in Figure 2.

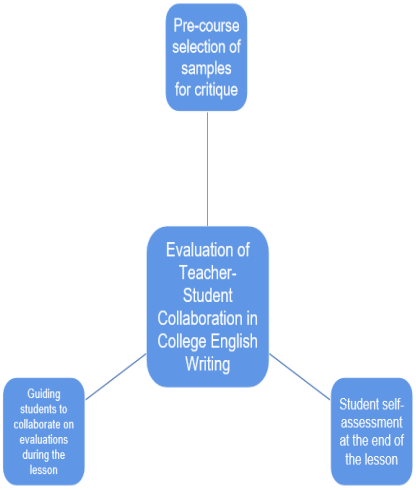

Figure 2. Evaluation of Teacher-Student Cooperation in College English Writing Teaching

In the whole process of teacher-student cooperative evaluation, teachers play various roles(Agasisti et al., 2022). On the one hand, the teacher, as the decision-making object, needs to decide the content of classroom evaluation and the task based on understanding the student's learning situation and needs to mobilize the students' learning enthusiasm in the classroom to ensure the efficiency of the classroom. On the other hand, the teacher should also play a supporting role in the post-class student self-assessment. The teacher should also play a supporting role, provide necessary support in the process of peer evaluation and machine evaluation, guide students to revise their essays, and help students improve their English writing ability. Teacher-student cooperative evaluation is not just a simple evaluation method; it is, more importantly, an effective means to promote student learning(Hamlin, 2021). Applying teacher-student cooperative evaluation in the college English writing classroom has more significance. It can not only improve the quality of students' output but also realize the teaching goal of writing, make students more deeply aware of their original writing problems, make students' classroom participation higher, improve their writing ability, and change from passive acceptance of knowledge to active exploration and learning.

4.1 English Vocabulary Mastery of College Students

Using the constructed big data model to analyze students' college English vocabulary using the *K*-means clustering algorithm and collating college English essays from the same academic year, a large amount of English writing material was accumulated. As shown in Figure 3, the

analysis was first carried out from the vocabulary dimension. Many college students' English writing materials were accumulated by using the traditional evaluation model to evaluate students' lectures in the last semester and by collecting college students' English writing essays enrolled in the same academic year and sorting them out. As shown in Figure 2, analyzing the word dimension, it can be seen that the word length of the student's English writing articles is relatively average, mainly focusing on the number of students between 4 and 5, which is 673, accounting for 67.3% of the total number of students; some of the students' higher vocabulary level reaches 8, which is 126, accounting for 12.6% of the total number of students; some of the students' lower level of vocabulary is less than 3, which is 201, accounting for 21.6% of the total number of students; and the ratio of the national college students' English level to the national college students' English level(Felix, 2021). The number of students with a lower vocabulary level is 201, accounting for 21.6%, which is similar to the proportion of national college students' English proficiency, proving that the English writing word length of college students in this grade is representative. When the college students' English vocabulary richness dimension is explored, it can be found that the students' vocabulary richness is mainly concentrated in the 5.25.7 between the number of 743, accounting for about 74.3%; some students' composition level higher vocabulary reached 7, the number of 98, and its proportion is 9.8%; some students' composition level lower vocabulary is less than 3, the number of 159, accounting for 15.9%, similar with The gap between the vocabulary richness used in the English writing of the national college students is relatively small, proving that the relationship between the use of traditional teaching and evaluation so that the diversity of students' vocabulary is consistent with the diversity of vocabulary used in the English writing of the national college students is representative of the situation, as shown in Figure 3.

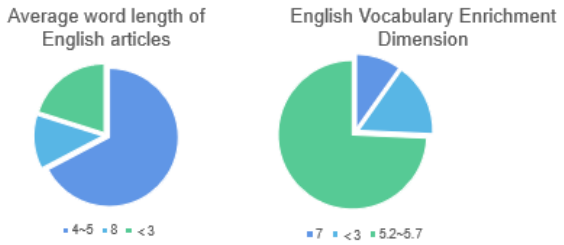

Figure 3. English Vocabulary Mastery of College Students Under Traditional Evaluation Methods

In the next semester, using the teacher-student cooperative evaluation system to teach and evaluate the students and collecting the English writing articles of college students enrolled in the same grade, it can be found that, from the vocabulary dimension, with the use of the teacher-student cooperative evaluation system for teaching, the college students' English level of English words of 56 is about 713 students, accounting for 71.3%, and some students' higher vocabulary level of composition reached 8, with 198 students, accounting for 19.8%; some students' lower vocabulary level of composition reached 45, with 89 students, accounting for 8.9%. The number of students is 198, and its proportion is 19.8%; some students' composition level is lower than the vocabulary length of 45, and the number of students is 89, and its proportion is 8.9%. In exploring the dimension of college students' English vocabulary richness, it can be found that students' vocabulary richness has a significant improvement, which mainly focuses on the period between 5.7 and 6.3, and the number of students is 698, accounting for about 69.8%, and the level of some students' composition is more Vocabulary richness of 7.3, the number of people for 143, its proportion is 14.3%; some students composition level lower vocabulary richness of 4.5–5.5, the number of people for 159, its proportion is 15.9%; its richness has exceeded the level of the national college students English dimensional richness. It is proven that using the evaluation system of cooperation between university teachers and students can significantly improve the mastery of English vocabulary of college students, and there is a significant improvement in the word length of students' English essays as well as in their vocabulary richness, as shown in Figure 4.

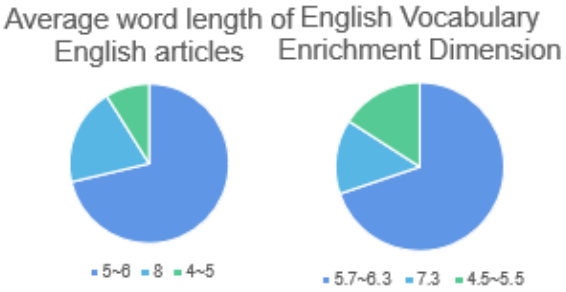

Figure 4. Mastery of English Vocabulary of College Students Under Teacher-Student Collaborative Evaluation

4.2 Analysis of college students' English utterances

Using the constructed big data model, the student's college English utterances were analyzed using the *K*-means clustering algorithm(Patsy et al., 2021). The syntactic structure of college students' English utterances was studied regarding average sentence length and density. As shown in Figure 5, the average sentence length of students' compositions mainly focuses on the distribution of students between 20 and 23, accounting for 583, accounting for about 58.3%. A small number of students have a lower level of composition, with the number of students with a sentence length of 12 times or less being 231, accounting for 23.1%. Some students have

a higher level of English writing, with a sentence length of more than 25, accounting for 186, accounting for 18.6%, which is within normal limits and the normal range of the English writing level of college students. The difference is not significant and is within the normal range, which is representative to a certain extent. Analyzing from the perspective of sentence density, it can be found that the sentence density of this grade of college students' English writing is mainly distributed in the range of 0.7–1.0; the number of students is 745, accounting for 74.5%; the number of students concentrating in the range of 0.54–0.8 is 132, accounting for 13.2%; and the sentence density of a minimal number of students with poorer English is lower than 0.3; the number of students is 123, accounting for 12.3%. It can be seen that students can use certain variations of sentences in their writing to enrich their essays with a certain degree of readability. However, it is not yet able to reach the ideal level of college English writing skills, as shown in 5.

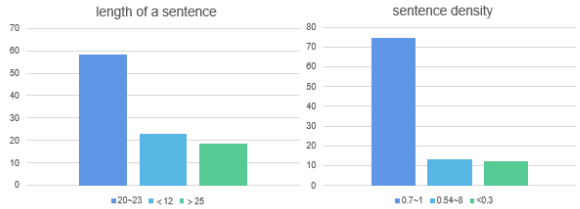

Figure 5. Mastery of English utterances by university students under traditional evaluation methods

As shown in Figure 6, through the use of teacher-student cooperation after the evaluation, it can be seen that the students' English writing sentence length has been significantly improved, mainly distributed in the number of 23–28 between 634, accounting for 63.4%, the level of English poorer students' compositions up to 15, the number of 184 accounted for 18.4% of the ratio of about 18.4%, which is the more excellent students' utterance length of the most extended distribution of more than 30 numbers of 182, accounting for about 18.2%, from the overall point of view, the length of their sentence length has been significantly improved. From the perspective of sentence density, it can be found that the main distribution of the sentence density of college students has been improved, from 0.71 to 0.91.2; the number of students is 683, accounting for 68.3%; the number of students with sentence density between 0.60.9 is 213, accounting for 21.3%; and the number of students with poorer English proficiency is 104, accounting for 10.4%; with the number of students with sentence density between 0.20.4, accounting for 10.4%. The number of students with sentence density between 0.2 and 0.4 is 104, accounting for 10.4%. Using the cooperative evaluation system figure 6 shows that the student's English writing level seems to have improved dramatically from the point of view of utterance.

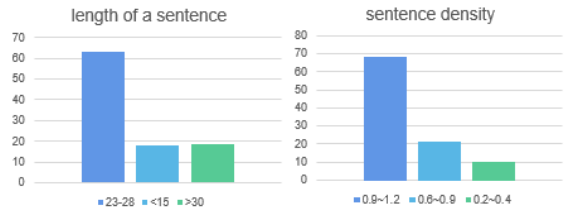

Figure 6. Mastery of English utterances by university students under the cooperative evaluation of teachers and students

4.3 Analysis of the length of college students' English articles

The *K*-means clustering algorithm uses a big data model to analyze college students' English proficiency. This analysis focuses specifically on the dimension of article length. As depicted in Figure 7, the essay length encompasses various components, such as the word count, conjunction usage, and clause incorporation inside the essay. From the dimension of students' essay length, the number of words used by students mainly focuses on the number of words between 150 and 230, accounting for 582, accounting for 58.2%. Although there is no clear requirement for word count in English writing classes for college students, some students believe that a word count of 150 words is more reasonable, similar to the word count in CET-4 and CET-6 Chinese and English writing. A small number of students have a word count of over 230, accounting for 35.2%. A small number of students with poor English proficiency have a word count of less than 100, accounting for 66%, which is similar to the national average and within the normal range(Shivakumar et al., 2021).

Among them, the length of the analysis from the point of view of the number of conjunctions used, the length of the number of conjunctions in the essay to play the role of the beginning and end, as well as the function of the reference, can reflect the logic of the essay, so it is a measure of the organization of the essay and an indicator of the organization of the ability of the key(Krasulina, 2023). The number of conjunctions in the length of the students' essays is more scattered, mainly distributed in the number of 810 659 people, accounting for 65.9%; a small number of students with the number of conjunctions below 3 253 people, accounting for 25.3%; some of the students who have a better level of English use the number of conjunctions more than 16 88 people, accounting for 8.8%; but the overall number of conjunctions used in the English language falls short of the ideal value.

Finally, the number of subordinate clauses used in the English articles of college students was counted. The number of subordinate clauses used in the English articles represents the students' ability to control English. The number of subordinate clauses used by the students is concentrated in the range of 5–6, with the number of

students being 542, accounting for 54.2%; the number of students using subordinate clauses is lower than 4, with the number of students being 243, accounting for 24.3%; and the number of students using subordinate clauses in the English articles is more than eight sentences with the number of students being 215, accounting for 21.5%. The number of students using subordinate clauses in English texts is less than ideal, as shown in Figure 7.

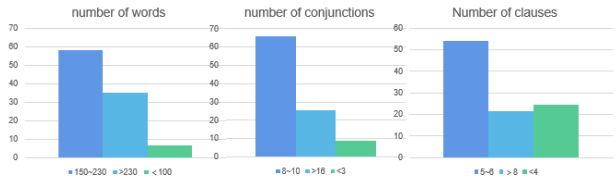

Figure 7. College students' English article writing level under traditional evaluation methods

As shown in Figure 8, after using teacher-student cooperative evaluation to guide students' English writing in the next semester, it was found that the number of students' word count increased from 150-230 to 645 between 180-260, accounting for 64.5%. A small number of students were able to exceed 300 words in English, with 294 students accounting for 29.4%. Some students had poor English, with 61 students between 100-130 words, accounting for 6.1%, compared to using traditional evaluation methods for teaching. The number of English words has significantly increased.

Analyzed from the perspective of hyphenation use, the number of students' compositions with hyphenation mainly distributed in the number of 911 is 639, accounting for 63.9%; some students' hyphenation can be more than 18; 213 account for 21.3%; and a small number of students' hyphenation use is between 57 and 148, accounting for 14.8%; and the overall number of English hyphenation use has increased compared with the traditional mode of discussion.

Finally, from the comparison of the number of clauses, it can be seen that the number of students using English clauses has increased from the original 5-6 to 678, accounting for 67.8%. The number of students using clauses less than 5 is 98, accounting for 19.8%, and the number of students using clauses more than 9 is 224, accounting for 12.4%. It can be seen that the number of students using clauses has significantly increased.

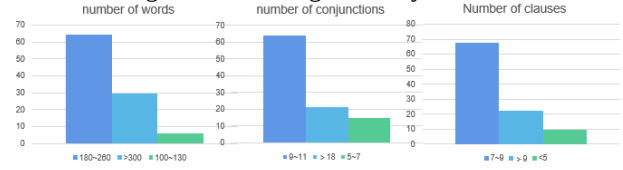

Figure 8. College students' English article writing level under the evaluation of teacher-student cooperation

With the rapid development of modern society and the increasing amount of information, people have entered the information society, and big data has entered every aspect of people's lives in an unstoppable mode. The application of big data in education realizes data-driven process description, diagnosis, prediction, and decision-making feedback. It promotes the formation of an orderly and interrelated cyclic structure of participants, educational resources, educational strategies, and educational tools related to education. Such a structure needs to be supported by appropriate and practical learning technologies, and the data generated by such a cyclic structure is the object of the application of learning analytics(Carter & Egliston, 2023). Therefore, under the premise of being supported by a large amount of data, big data is a powerful impetus for educational change. In teaching foreign languages, teachers gradually realize that foreign language learning can be a unilateral output and, more importantly, let the students receive the input information. Although college English writing has been paid more attention than before, teacher is still the main body of the evaluation, with the absolute right to evaluate the students. The students can only be passive and listen to the requirements and instructions, and they are in the passive position of writing learning without the internal drive to learn. Students can only passively follow the requirements and instructions and are passive in their learning without the internal drive to learn. Teachers have a more significant burden as the main body of evaluation, and there is no way to evaluate students' evaluations individually. In addition, students need to understand the teacher's comments thoroughly, and the mistakes they make are repeated in the writing process, so the teacher's evaluation does not get positive feedback and cannot achieve the expected results. Therefore, in recent years, the emerging teacher-student cooperative evaluation has attracted teachers' attention(Lin et al., 2021). Teacher-student cooperative evaluation is not only a simple one-way evaluation method but also a means to promote students' learning, and in the college English writing classroom, the use of teacher-student cooperative evaluation mode to evaluate students' essays has a more significant meaning. The teacher-student cooperative evaluation model focuses on more than just the quality of the students' English compositions. More importantly, the teacher-student cooperative evaluation mode can realize the teaching goal of writing, which makes students realize their deficiencies more clearly. The students' participation in class is further improved, and their initiative in learning is enhanced. The teacher-student cooperative evaluation mode emphasizes students more. Under the leadership of the teacher, the primary mode of learning involves the simultaneous process of evaluation and self-assessment, transcending the traditional demarcations between learning and evaluation. According to Yiran (2021), the implementation of the teacher-student cooperative assessment approach has the potential to enhance students' writing skills more efficiently. After processing a large number of students' English compositions through big data, the English level of students using the traditional evaluation method could be higher. After half a semester's study, their level must reach ideal. However, after learning through the teacher-student cooperative assessment, the student's English writing level has

significantly improved in terms of vocabulary, sentence length, and sentence structure.

5 Conclusion

In this paper, a large amount of student information is processed using big data. A reasonable algorithm is constructed to calculate the teacher-student cooperative evaluation system for college English writing, through which it is found that using the new teacher-student cooperative evaluation system can help students better integrate into the classroom so that they can become the main body of learning and be the master of learning. Moreover, the teacher plays more of a guiding role. Taking the same group of students as the main body of learning and using different evaluation modes in the two semesters, it can be seen from the vocabulary mastery of college students in English writing that the length of the words and the density of the words of the students have been improved; from the analysis of college students' sentences, it can be seen that the length of the sentences used has been increased significantly, and the density of the sentences has been increased as well. Finally, the analysis from the length perspective shows a significant increase in the number of words, conjunctions, and clauses. Teaching and learning are the two main aspects of educational activities. In today's massive amount of educational information, constantly analyzing and exploring the potential value of data, promoting the integration of teaching and data to cultivate the quality of talent is of great practical significance. Integrating a new evaluation model and big data can help teachers better understand the problems in students' learning and help students improve their learning ability.

References

[1] Agasisti, T., Celma, D. O. R., & Montemor, D. S.. The efficiency of Brazilian elementary public schools. International Journal of Educational Development, (2022).93(2), 256–260. https://doi.org/10.1016/j.ijedudev.2022.102627

[2] Andreis, L., Knig, W., & Patterson, R. I. A. A large ↘ eviations principle for all the cluster sizes of a sparse Erds–Rényi graph. Random Structures and Algorithms, (2021). 2(3), 570-. https://doi.org/10.1002/rsa.21007

[3] Carter, M., & Egliston, B. What are the risks of Virtual Reality data? Learning Analytics, Algorithmic Bias and a Fantasy of Perfect Data: New Media & Society, (2023). 25(3), 485–504. https://doi.org/10.1177/14614448211012794

[4] Conn, K. M., Lovison, V. S., & Hyunjung, M. C. How Teaching in Underserved Schools Affects Beliefs about Education Inequality and Reform. Public Opinion Quarterly, (2022). 5(1), 1. https://doi.org/10.1093/poq/nfab072

[5] Felix, K. (2021). Enhancing EFL students' participation through translanguaging. ELT Journal, 9(10), 12-. https://doi.org/10.1093/elt/ccaa058

[6] Giesler, T. Much ado about nothing? The impact of the Reform Movement on (northern) German English language teaching. Language and History, (2021). 2(3), 23-. https://doi.org/10.1080/17597536.2021.1996088

[7] Giffen, B. V., Herhausen, D., Fahse, T., & Woodside, A. G.. Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods. Journal of Business Research, (2022) 144(4), 57-.

[8] Guan, K., Matsushima, K., Noguchi, Y., & Yamada, T. Topology optimization for rarefied gas flow problems using density method and adjoint IP-DSMC. Journal of Computational Physics, (2023). 5(11), 444-. https://doi.org/10.1016/j.jcp.2022.111788

[9] Hamlin, D.. Can a Positive School Climate Promote Student Attendance? Evidence From New York City: American Educational Research Journal, (2021) 58(2), 315–342. https://doi.org/10.3102/0002831220924037

[10] Hilali, T. S. A., & Mckinley, J. Exploring the socio-contextual nature of workplace writing: Towards preparing learners for the complexities of English L2 writing in the workplace. English for Specific Purposes, (2021). 63(3), 86–97. https://doi.org/10.1016/j.esp.2021.03.003

[11] Krasulina, E. G. On implementation of some systems of elementary conjunctions in the class of separating contact circuits. Discrete Mathematics and Applications, (2023). 33(1), 19–29. https://doi.org/10.1515/dma-2023-0003

[12] Kumar, R., Nayak, S., Garbrecht, M., Bhatia, V., & Saha, B. Clustering of oxygen point defects in transition metal nitrides. Journal of Applied Physics, (2021). 129(5), 055305. https://doi.org/10.1063/5.0038459

[13] Liang, Q., Torre, J. De. La., & Law, N. Do Background Characteristics Matter in Children's Mastery of Digital Literacy? A Cognitive Diagnosis Model Analysis. Computers in Human Behavior, (2021). 3(6), 106850. https://doi.org/10.1016/j.chb.2021.106850

[14] Lin, Y., Xu, B., Feng, J., Lin, H., & Xu, K. Knowledge-enhanced recommendation using item embedding and path attention. Knowledge-Based Systems, (2021).233(3), 107484-. https://doi.org/10.1016/j.knosys.2021.107484

[15] Meziane, R., FaridBenhadid, MostefaNiar, AbdelatifMamache, BakirMeziane, Toufik. Comparative evaluation of two methods of pregnancy diagnosis in dairy cattle in the East of Algeria: Proteins associated with pregnancy and ultrasonography. Biological Rhythm Research(2021)., 52(1a2), 34-.

[16] Patsy, D. P., Jonas, B., & Vera, M. 274Early adversity scale for schizophrenia (EAS-SZ) constructed and validated using linked register data. International Journal of Epidemiology, (2021). 9(Supplement_1), Supplement_1. https://doi.org/10.1093/ije/dyab168.158

[17] Shivakumar, N., Chandrashekar, A., Handa, A. I., & Lee, R. Use of deep learning for detecting, characterizing and predicting metastatic disease from computerized tomography: A systematic review. Postgraduate Medical Journal, (2021) postgradmedj-2020-139620. https://doi.org/10.1136/postgradmedj-2020-139620

[18] Su, N., & Zhang, F. Anomalous overland flow on hillslopes: A fractional kinematic wave model, its solutions and verification with data from laboratory observations. Journal of Hydrology, (2022). 604, 127202-. https://doi.org/10.1016/j.jhydrol.2021.127202

[19] Yiran, L. Evaluation of students' IELTS writing ability based on machine learning and neural network algorithm. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, (2021). 4(4), 40.

[20] Yurkofsky, M. Environmental, Technical, and Representational Uncertainty: A Framework for Making Sense of the Hidden Complexity of Educational Change: Educational Researcher, (2022). 51(6), 399–410. https://doi.org/10.3102/0013189X221078590