

Retraction

Retracted: Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data

Journal of Function Spaces

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Journal of Function Spaces. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] M. Zhang, "Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data," *Journal of Function Spaces*, vol. 2022, Article ID 3367256, 11 pages, 2022.

Research Article

Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data

Ming Zhang 

School of Applied Technology, China University of Labor Relations, Beijing 100048, China

Correspondence should be addressed to Ming Zhang; zhangm@culr.edu.cn

Received 27 April 2022; Revised 14 June 2022; Accepted 16 June 2022; Published 12 July 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Ming Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuous development and application of data computing in the whole society in the construction of digital campus and intelligent campus of each higher education institution. In the environment of education, universities use these data well, which not only affects the orderly operation of higher education but also will become an inexhaustible power to help higher education promote the reform and innovation of education and teaching system. In this paper, we focus on the teaching operation and students' independent learning by taking students' evaluation data and students' online learning data of Y school as the research objects. We conducted a preliminary analysis and transformation of students' evaluation data of a university, eliminated the abnormal evaluation data by using the improved cosine phase dissimilarity algorithm, standardized the evaluation data by using the normalization method, and used the traditional *K*-modes algorithm. Based on these three problems, the traditional *K*-modes algorithm was improved in three aspects, including the determination of the number of clustering families, the determination of the measurement of clustering distances, and the experimental results showed that the improved algorithm was more reasonable and effective.

1. Introduction

Since the McKinsey report first introduced the concept of "big data," human society has been paying more and more attention to "big data," and people are widely aware. The potential value of "big data" has been widely recognized. Therefore, in view of the great potential, all countries in the world regard "big data" as important as oil, coal, minerals, natural gas, etc., and the research on "big data" at home and abroad has entered a brand new era. Research on "big data" has a crucial impact on various fields such as scientific progress, national security, and national education [1].

According to the study, the continuous expansion of data application will have a far-reaching impact on all fields of the whole society [2]. Big data is not only affecting people's learning, living, working, and thinking in an unprecedented way but also influencing the way of production, operation, organization, and management in various industries. It can be said that the era of big data is not only affecting the industry revolution but also the big changes in education [3]. Therefore, many experts and scholars have

conducted a series of fruitful researches and discourses based on this, which have reform and innovation of university education and teaching system.

On March 29, 2012, BDRI was launched in the United States, elevating big data to a national strategy [4]. At the same time, some U.S. teaching fields have made all-round reports on intelligent computing, which has accelerated the change of the whole field. [5]. In the report, the basic situation of online education big data application in the United States is analyzed with the case of online education big data, especially the case of adaptive personalized learning system discusses more comprehensively the development situation and a series of challenges and problems that will be faced by online education in the future [6]. Desire2Learn, a Canadian company, develops systems for student-directed learning, academic alert, and intervention services, and so on [7].

Its technology is no longer a traditional information technology, but provides the needs of intelligent transformation, integrating technologies, services, and solutions such as artificial intelligence, 5G, cloud computing, and the Internet of Things. At present, from intelligent manufacturing to

intelligent transportation, remote diagnosis and treatment, and online education, all walks of life are experiencing the process of digital and intelligent transformation [8]. In particular, 5G not only accelerates the intelligent transformation of the industry but also makes the traditional “cloud pipe end” IT architecture including cloud computing center, network pipeline, and terminal ineffective, thus giving birth to a new architecture [9]. “New IT” can also support intelligent manufacturing and promote the high-quality development of manufacturing industry. With digital and intelligent transformation and upgrading, China’s huge manufacturing industry is transforming into high-quality high-end manufacturing and intelligent manufacturing [10].

It is very important to deeply explore the potential resources and application value of big data in higher education to guide higher education institutions to fully implement the fundamental task of moral education and the concept of student-centered education, to promote the deep integration and intelligent transformation of higher education teaching activities, to promote the scientificization of higher education teaching system and decision-making system, to help the transformation and upgrading of scientific research paradigm in higher education institutions, and to build a more scientific management system in higher education institutions [11]. It is very important to promote decision-making system of higher education, help the transformation and upgrading of scientific research paradigm of higher education, and build up a more scientific management system of higher education [11]. A large amount of dynamic data will be generated in the process of school teaching and learning. With the help of big data technology, people can easily collect and mine these data to improve school teaching and learning and promote school quality and efficiency. And decision-making can make the evaluation results with higher reliability and credibility and make the decision-making more scientific and accurate [12].

Taking the student evaluation data and student online learning data of Y school as the research object, this paper focuses on teaching operation and students’ autonomous learning. The traditional *K*-modes algorithm is innovatively improved in three aspects: the determination of cluster family number and the determination of cluster distance measurement. The experimental results show that the improved algorithm is more reasonable and effective. This paper makes a targeted exploration from the two aspects of teaching operation and students’ online learning and draws some preliminary conclusions with certain reference value. Student evaluation, expert evaluation, peer evaluation, supervision, and evaluation institutions, and daily teaching monitoring data can be formed into more comprehensive, accurate, popularized, and influential results. It will provide more targeted decision-making, reform, and innovation basis for education and teaching managers and front-line teachers [13].

This paper focuses on two aspects of new era of big data environment of university education as the research background, takes the real student evaluation data and student online learning behavior data of a university as the research object, and carries out the relevant application research on

the teaching operation and learning situation of a university by using the improved clustering algorithm and the neural network algorithm based on machine learning, and the experimental result method has good scientific, and the thesis is divided into five sections: Section 1 describes the research background of this paper and the main structure of this paper; Section 2 introduces the current situation of domestic and foreign research in related fields and summarizes the research significance of this paper; Section 3 takes the student evaluation data of a university as the research object and adopts to model and analyze the school teaching and learning situation based on the abnormal data elimination and specification of the student evaluation data. The operation is modeled and analyzed, and the teacher teaching status evaluation model is established. In Section 4, the proposed scheme is tested and analyzed. Section 5 summarizes the research content of this paper and provides an outlook on future research directions.

2. The Related Works

Scholars have different interpretations on the definition or connotation of precision teaching. Precision teaching is divided into four stages: precision teaching breeding, precision teaching creation, precision teaching expansion, and precision teaching informatization. Carefully analyzing the statements of scholars at home and abroad on precision teaching, we find that no matter at which stage of precision teaching, scholars’ statements on precision teaching have the following common characteristics: data-based teaching, evaluation based teaching, learner centered teaching, and teaching that emphasizes recording and analyzing students’ learning behavior and performance. Therefore, we believe that the essence of accurate teaching is the spirit of scientific reflection [14, 15].

Chen Shongyeh et al. pointed out that precision teaching is supported by many theories, such as teaching theory, learning theory, curriculum theory, and technology theory. There are also many idealized models in implementation. Its experience is that the theory must be localized, and school-based. Only by integrating theory with practice can it be more widely used and reflect the localization characteristics of precision teaching [16]. Zhang Junchao et al. pointed out that when carrying out precision teaching, many regions or schools will adopt a one size that fits all approach in operation, requiring teachers to fully adopt new technologies and subvert the original teaching methods or teaching habits. According to our practical experience, we believe that teachers’ habits should be changed step by step and students’ habits should not be changed as much as possible [17]. Zhang Yannan et al. believe that the arrival of the era of big data in education will certainly be a revolutionary change due to its many characteristics such as comprehensiveness, real-time, and potential, which make a series of problems that are difficult to solve in the traditional education field are greatly improved, such as the balance, science, and rationality of education, and then the education model, education implementation path, and scientific evaluation of education [18]. Yang Xianmin et al. believe that one of the important

reasons for students' heavy academic burden is that teachers do not grasp the learning situation. Teachers' teaching effect is positively correlated with their understanding of students. The greatest value of big data precision teaching lies in accuracy, which is to deeply and fully understand the learning situation by collecting student data. Therefore, if we want to collect students' data fully and completely, we should collect students' classroom homework, synchronous test, simulated test, and other data horizontally. The data of students from enrollment to graduation are collected vertically [19].

At present, many studies have comprehensively discussed the concept of educational evaluation and the context of big data in education and pointed out specific implementation paths, which have indicated the direction for the comprehensive reform and coordinated development of education teaching [20]. In the environment of big data in education, based on the accelerated evolution of the new round of technological revolution and industrial change, take the national strategy, new industry, new economic development, and the future development direction of industry as the guide, and take the cultivation of innovation and practical as the guide. With the cultivation of innovation spirit and practical ability of college students as the core, the universities use these data well, which not only affects the orderly operation of the whole education teaching system but also will become the inexhaustible power teaching system in higher education.

K -modes algorithm has many advantages, but there are also many shortcomings, so He, San, Ng et al. proposed the method of calculating the distance between sample data by the frequency of occurrence within a class; Hsu et al. calculated the distance between sample data based on hierarchy; Ganti et al. and Ahmad et al. used the degree of cooccurrence data; these studies have made some improvements to K -modes algorithm. These studies have improved the K -modes algorithm, and its clustering ability has been greatly improved in the corresponding applications, but there are still shortcomings in the problems discussed in this study, mainly because they all ignore the different attributes of the connections and differences between different objects, and for this reason, some improvements are made to the K -modes algorithm for the specific problems studied in this paper. On this basis, this paper takes the big data of education and teaching in a university as the background and establishes relevant mathematical models with the help of big data related theories and technologies to guarantee that the evaluation and prediction of school teaching are more scientifically based and convincing.

3. Application of Improved K -Modes Algorithm in the Evaluation of Teaching Status of University Teachers

3.1. Basic Structure of Neural Network. K -modes algorithm is a clustering algorithm used to classify attribute data in data mining. K -modes algorithm is an extension of K -means algorithm. It can only deal with numerical data, but not classified attribute data. However, the traditional K -

-means algorithm is only suitable for data sets with continuous attributes, and neural algorithm is needed to supplement the data sets with discrete attributes. In the process of practical application, a single neuron cannot fit too complex mapping relationship. We need to build more complex networks to approximate those more complex objective functions. Using multilayer networks can sometimes find a good convergence relationship after less iterative training, as shown in Figure 1.

Neurons, as the most basic building blocks of neural networks, are responsible for computational or processing functions, which are usually expressed in terms of computational functions called excitation functions, and their functions vary greatly for different applications, so they should be selected with great care. The biological structure of a neuron is shown in Figure 2.

3.2. Problem Formulation and Data Selection and Data Structure Analysis of Evaluation Data. The evaluation indicators are as follows:

(1) Quality literacy

They have high moral character and serious and responsible teaching, pay attention to the image of teachers, strictly abide by teaching discipline, and do not suspend and transfer classes at will, no late arrival, and early departure, profound professional knowledge, solid practical skills, fluent expression in Mandarin, and standardized board writing.

(2) Teaching attitude

The teaching attitude is correct, the class is well prepared, and the teaching schedule can be reasonably arranged according to the syllabus; the class is organized in an orderly and active atmosphere, and the teaching content is organized rigorously without reading from the text.

(3) Teaching skills

The lectures are clear, accurate, and focused on difficult points; they focus on the systematic, scientific, and advanced knowledge and are able to update the teaching content in a timely manner by combining the frontiers and dynamics of the field.

(4) Extracurricular sessions

He teaches and nurtures students, focuses on the cultivation of students' innovative and practical abilities, pays attention to the process management, reviews homework carefully, provides timely counseling, answers questions, often takes the initiative to communicate with students, pays attention to students' feedback, and continuously improves teaching.

The alternatives are the following: excellent, good, moderate, pass, and fail.

In order to illustrate the problem, improve the efficiency of the experiment, and focus the work on the improvement

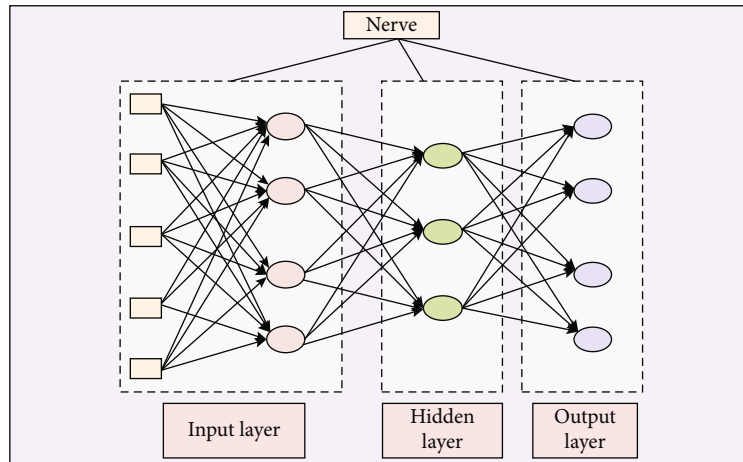


FIGURE 1: Schematic diagram of neural network structure.

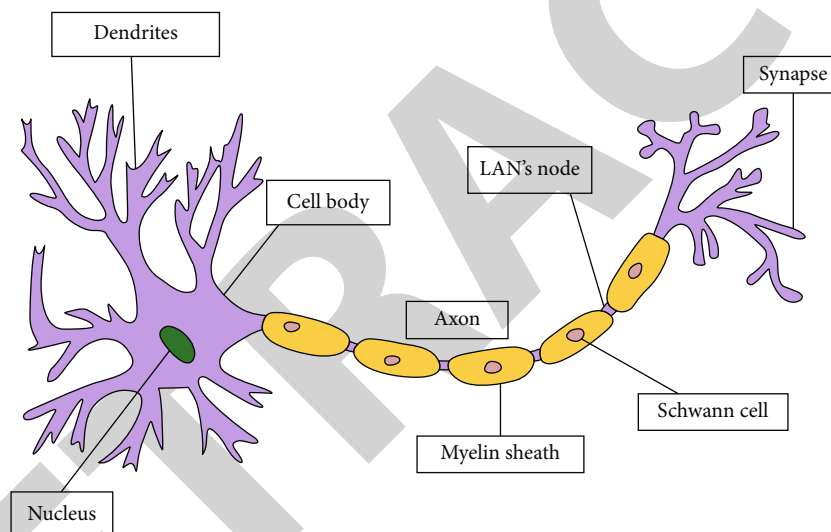


FIGURE 2: Schematic diagram of the biological structure of a neuron.

of the method, all the data of student evaluation in the follow-up of this paper are the data of one college of the university (referred to as College R for the convenience of the subsequent narrative) for the last ten academic years, i.e., from the second half of 2009 to the second half of 2019, with a total of 237,924 student evaluation records.

The data tables related to the evaluation data table in the Academic Affairs Management System of Campus Y are: Student Information Data Table (Student Number, Name,), Teacher Information Data Table (Employee Number, Name,), Course Information Data Table (Course Code, Course Name,), teaching task data sheet (academic year, semester, course code, course selection number, employee number,), course selection data sheet (academic number, course selection number,), and student evaluation. There are 7 tables of indicators (serial number, evaluation item, content, and evaluation level). The student evaluation data table is based on XN, XQ,

XH, XKKH, and ZGH as the main code, and the sex associated with other data is the external code, for example, for the student information table data, its external code is XH, to effectively ensure the integrity and consistency of all relevant data in the system.

3.3. Student Evaluation Data Preprocessing. In this paper, there are four evaluation indicators for student evaluation data, and each indicator has five values, which is a multidimensional subtype problem. Since the subtypes are different from the numerical data, it is difficult to measure the differences between the data in a hierarchical way, and also for the sake of presentation, without changing the interpretation, each evaluation value of student evaluation is transformed accordingly in this paper, i.e., each different subtype is assigned a rank that exactly matches its value. The specific transformation strategy is as follows:

$f(x)$ is the transformation function of the evaluation value;

x : for PJ_i original rating, $x \in \{\text{excellent, good, moderate, failing}\}$, $i = 1, 2, 3, 4$;

$$\begin{aligned} f(x) &= 5, x = \text{excellent}, \\ f(x) &= 4, x = \text{good}, \\ f(x) &= 3, x = \text{medium}, \\ f(x) &= 2, x = \text{pass}, \\ f(x) &= 1, x = \text{failed}. \end{aligned} \quad (1)$$

Anomalous data in student evaluation data are often due to the fact that individual students do not evaluate a course taught by a particular instructor objectively and fairly, but rather evaluate it with a strong personal touch, resulting in a large deviation between their evaluations and those of other students. There is a significant correlation between students' cognitive style and some dimensions of college students' social adaptability. Field-independent cognitive style is significantly correlated with learning adaptability, career choice adaptability, self-care adaptability, and physical and mental adaptability. Field-dependent cognitive style is significantly correlated with interpersonal adaptability and role adaptability. Therefore, these anomalous data need to be removed from the evaluation of teaching data.

For example, for two sample data x, y , their original evaluation values are (1, 1, 1, 1) and (5, 5, 5, 5), and the replacement values are (1-3, 1-3, 1-3, 1-3) and (5-3, 5-3, 5-3, 5-3), respectively.

The corresponding phase dissimilarity is calculated as

$$\cos(x, y) = \frac{(-1 \times 2) + (-1 \times 2) + (-1 \times 2) + (-1 \times 2)}{\sqrt{(-1)^2 + (-1)^2 + (-1)^2 + (-1)^2} \times \sqrt{2^2 + 2^2 + 2^2 + 2^2}} = -1. \quad (2)$$

Applying the method above, 237924 student evaluation records from 1326 categorical sample data files of College R were subjected to anomalous data elimination, and the results are shown in Figure 3.

3.4. Application of the Improved K-Modes Algorithm to the Evaluation of College Teachers' Teaching Ability. After eliminating and merging the abnormal evaluation data in the previous section, the K-modes algorithm will be applied to evaluate the teaching ability with the evaluation of 59 instructors in a semester in College R. As we can see from the discussion of K-modes algorithm in Section 2, K-modes algorithm is a 0-1 matching algorithm based on whether the values of each attribute of the sample data are the same or not. However, the algorithm also has three serious shortcomings in solving some specific problems with certain relationships among attributes.

Firstly, in the k-means algorithm, K is given in advance, and the selection of this k value is very difficult to estimate. In the initial K-means clustering, we need to determine an initial partition algorithm, and then, we need to optimize an initial partition algorithm. The selection of this initial clustering center has a great impact on the clustering results.

Finally, the algorithm needs to constantly adjust the sample classification and constantly calculate the adjusted new cluster center. Therefore, when the amount of data is very large, the time overhead of the algorithm is very large.

In this paper, each uses the method of minimizing sum of squared error (SSE) to determine the number of clustering families K .

(1) Definition of SSE

$$\text{SSE} = \sum_{l=1}^k \sum_{x \in L_l} \text{Dist}(x, Z_l)^2, \quad (3)$$

where k is the number of clustering families, Z_l is the clustering center of the l cluster L_l , and $\text{Dist}(x, Z_l)$ is the similarity between the sample data x and the clustering center Z_l . Different similarity calculations often bring large differences to the clustering results, since the frequency-based similarity calculation is measured by the frequency of each attribute of each sample data in the whole sample data set.

(2) Calculation of the frequency of sample data

In order to obtain the frequency of each attribute of the sample data, it is necessary to count the number of occurrences of each rating value in the whole sample data set. Let x_{ij} be the evaluation value of the attributes of the sample data, $f(x_{ij})$ be the frequency of x_{ij} , C be the set of allowed values of x_{ij} , for this example, C be {5 (excellent), 4 (good), 3 (moderate), 2 (pass), 1 (fail)}, C_p be the number of times the element of p in the set C appears in the sample data set, where $i = 1, 2, \dots, n$, n is the number of sample data, $j = 1, 2, \dots, m$, m is the dimension of the sample data, $p = 1, 2, \dots, P$, $P = |C|$ is the base of the set C , i.e., $f(x_{ij})$ can be calculated by the formula:

$$\begin{aligned} f(x_{ij}) &= C_1, x_{ij} = 5, \\ f(x_{ij}) &= f(x_{ij}) = C_2, x_{ij} = 4, \\ f(x_{ij}) &= f(x_{ij}) = C_3, x_{ij} = 3, \\ f(x_{ij}) &= f(x_{ij}) = C_4, x_{ij} = 2, \\ f(x_{ij}) &= f(x_{ij}) = C_5, x_{ij} = 1. \end{aligned} \quad (4)$$

(3) Frequency-based similarity calculation

The frequency of the sample data is calculated by the following formula:

$$\text{AVF}(x_i) = \frac{1}{m} \sum_{j=1}^m f(x_{ij}), \quad (5)$$

where x_i is the sample data and $f(x_{ij})$ is the frequency of the sample data x_i on the attribute j .

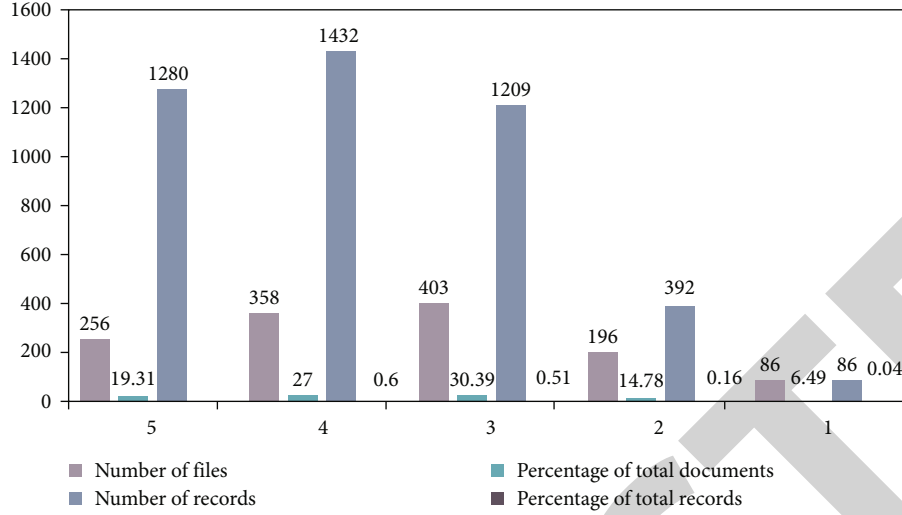


FIGURE 3: Number of records containing anomalous data versus number of files and number of records.

(4) Preclustering to determine the number of clusters K

The frequency-based similarity was used to pre-cluster the classes 1-7 with K -modes, and the minimum error sum of squares versus the number of clusters K was obtained as shown in Figure 4.

From Figure 5, it is obvious that the inflection point of the image appears at $K = 3$, so it is more appropriate to cluster 3 classes for a semester of student evaluation results in College R. In order to prove this point, the same method was used to precluster all teachers' evaluation data of school Y. The preclustering results showed that the inflection point of the image appeared at $K = 5$; when the dimensionality of the sample data was doubled, the preclustering results showed that the inflection point of the image appeared at $K = 7$. Therefore, the value of K varies for different types and scales of problems and should be treated with caution in specific problems.

From the principle of K -modes algorithm to achieve clustering, it is known that for the randomly selected K clustering centers, iterate and update the clustering centers continuously according to the closest distance principle, and the goal is to minimize the SSE. Inspired by this idea, the scheme of this paper is specified as follows.

Each sample data point other than the one already selected as a cluster center is used as the hypothetical initial cluster center at $l + 1$, and the SSE with $l + 1$ is calculated, and the data point with the fastest decreasing SSE value is selected as the true initial cluster center at. The data point with the fastest decreasing SSE value was selected as the true $l + 1$ initial cluster center, repeated, and terminated when K is the known number of classifications.

Based on the above idea, the following model is developed:

$$\begin{aligned}
 Z &= \{z_l | z_l \in X, l \leq K\}, \\
 z_l &= \max_{x \in X} \text{DJJSSE}(x), \\
 \text{DJJSSE}(x) &= \left(\sum_{j=1}^{l-1} \text{Dist}(x, x_j)^2 + \min_{x \in X} \left(\sum_{i=1}^n \text{Dist}(x, x_i)^2 \right) \right), \quad (6)
 \end{aligned}$$

where Z is the set of initial clustering centers, z_l is the sample data point that makes SEE fall fastest x , i.e., the l initial clustering center, x is the sample data, and n is the number of sample data.

The traditional K -modes algorithm of clustering distance calculation uses the similarity (0-1 match, 0 for same and 1 for different) of sample data and cluster center data for the metric; this method is feasible and effective for applications where the sample attributes take values independently of each other, but if there is a certain connection between different attributes of the sample, and different values under the same attribute also have a certain connection; the method's exposed deficiencies are very large. For example, in this case, if the students' evaluation of teacher 1 is (2, 2, 3, 2), and the evaluation of teacher 2 is (5, 4, 5, 4), and the cluster center is (3, 5, 4, 3), the distance of teacher 1 from the cluster center is (1, 1, 1, 1) = 4, and the distance of teacher 2 from the cluster center is (1, 1, 1, 1) = 4, using the traditional K -modes algorithm for similarity measurement. (1, 1, 1, 1) = 4, which is obviously far from the actual situation. Therefore, on the one hand, the results of clustering may vary greatly depending on the initial clustering centers, and the results are unstable, as can be seen from the table comparison; on the other hand, because the distance measure does not consider the connection between different attributes and different values of the same attribute and simply uses 0-1 matching to measure the distance, the clustering results are poor and do not match the actual situation. Therefore, in this paper, the distance measure of co-occurrence proposed by Ahmad et al. is used to improve the similarity of the traditional K -modes algorithm by using the similarity between different attribute values under the same attribute with other attributes.

The application of the improved K -modes algorithm for cluster analysis of student evaluation data to achieve the evaluation of teachers' teaching ability can be summarized in two stages and eight basic steps, as follows.

Phase I: data preprocessing, consisting of four basic steps

Step 1: sample data conversion, i.e., rating values from (excellent, good, moderate, passing, and failing)

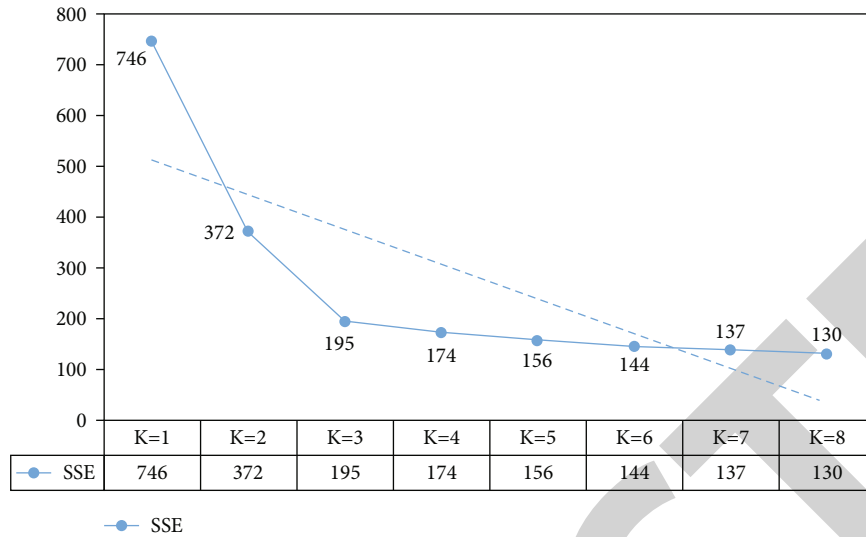


FIGURE 4: Plot of the minimum error sum of squares versus the number of clusters K .

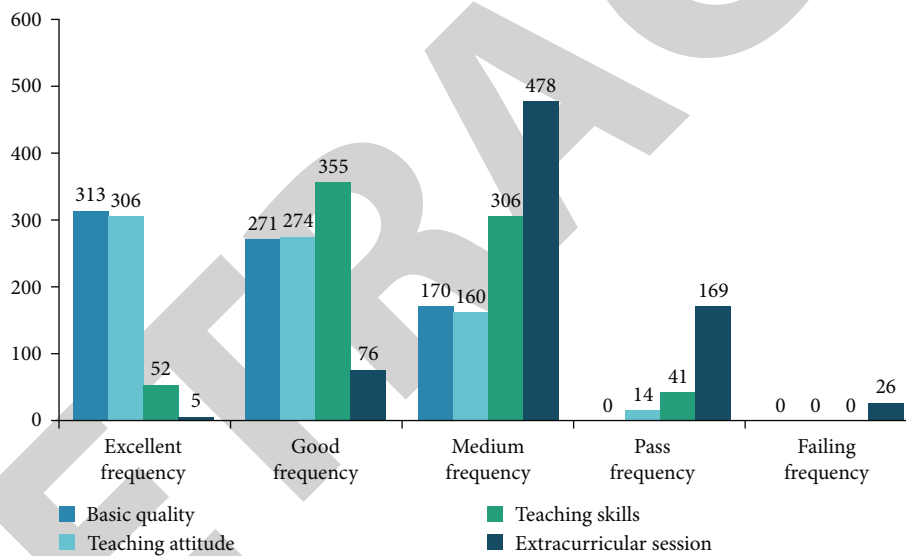


FIGURE 5: Statistics of students' evaluation of teaching in a semester in school Y.

Convert to (5, 4, 3, 2, 1)

Step 2: sample data splitting, i.e., all the evaluation data are split into multiple classification data files by “academic year + semester + course code + employee number” as the splitting code

Step 3: sample anomaly data elimination, i.e., using the cosine dissimilarity calculation model, the sample data whose evaluation values in each categorical data file are significantly different from other evaluation values are eliminated from the corresponding categorical data file

Step 4: merge and restore the sample classification data files, i.e., for the sample classification data files, merge each teacher's evaluation into one record according to the merge code of “academic year + semester + employee number”, and restore it

Phase II: clustering evaluation, consists of four basic steps

Step 1: a reasonable number of clusters K is determined by applying the frequency-based (AVF) similarity calculation algorithm for clusters that have not been given a specific number of clusters K or have been given a number of clusters K but have poor clustering results

Step 2: even poor usability of clustering results due to the selection of outlier data points or proximity points, K initial clustering centers are determined by using the sum of least errors squared (SSE) algorithm to avoid the occurrence of two types of situations

Step 3: traditional K -modes clustering algorithm does not consider the relationship between different attributes and different values of the same attribute in the 0-1 matching method when calculating the clustering distance, the

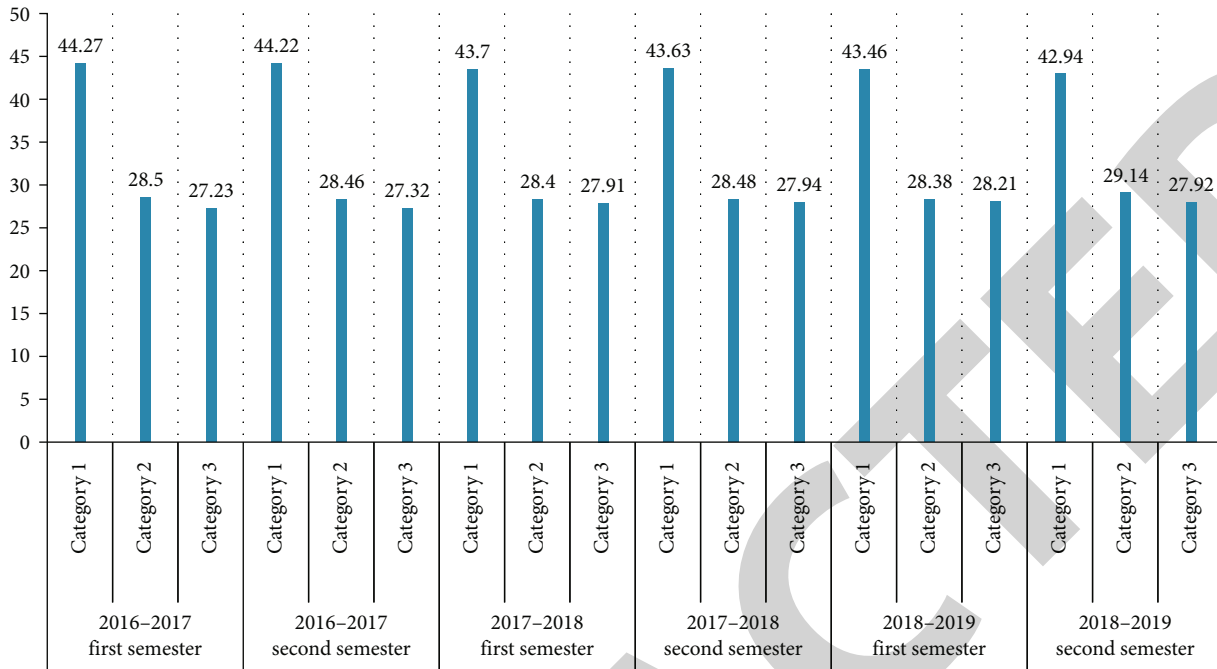


FIGURE 6: Histogram of the percentage of teachers in the three categories at school Y.

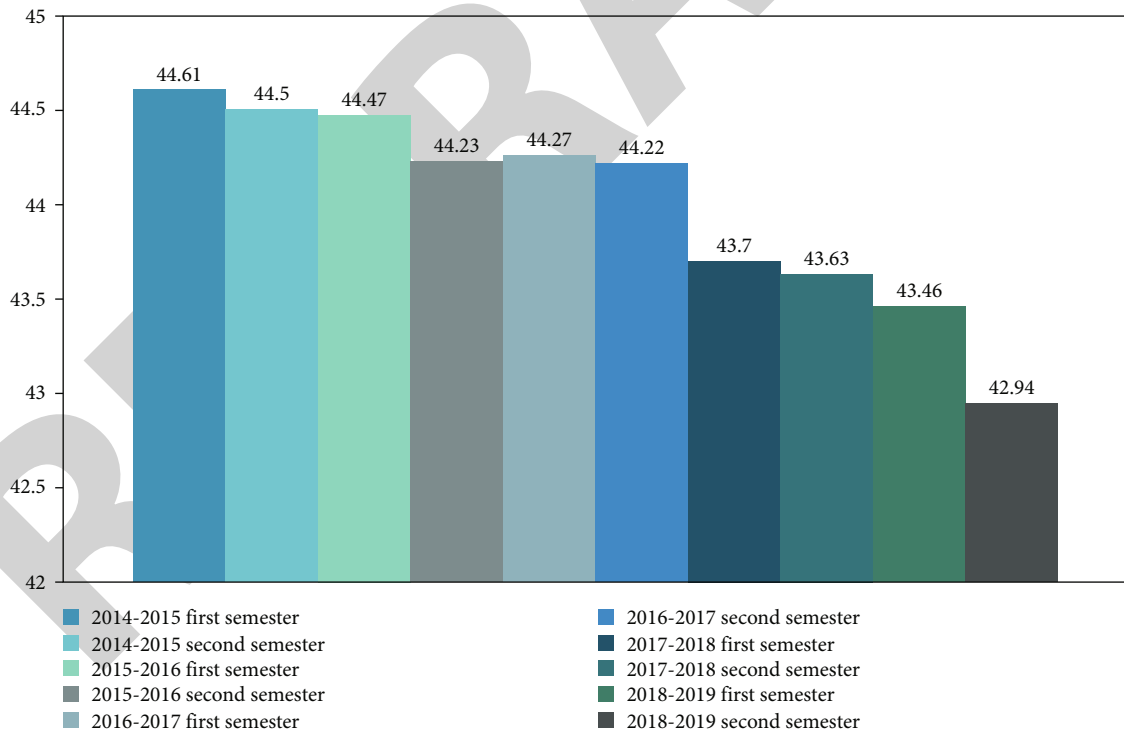


FIGURE 7: The proportion of first-class teachers has changed in the past five years.

traditional *K*-modes clustering algorithm is improved by using the cooccurrence of evaluation value as the distance measure

Step 4: based on the clustering results, we analyze the overall teaching situation of all teachers and the teaching sit-

uation of each teacher in each semester by semester, managers to grasp the overall teaching situation of all teachers, and teachers in each teaching unit in a timely manner and thus provide a scientific basis for correct decision-making and targeted measures; on the other hand, we provide

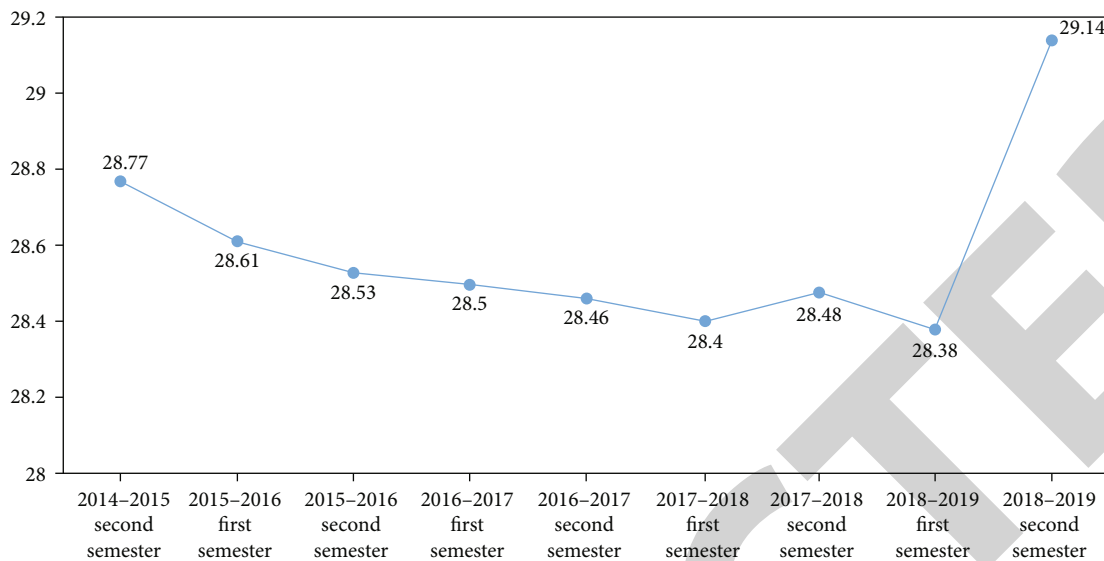


FIGURE 8: The percentage of category 2 teachers in school Y over the past five years.

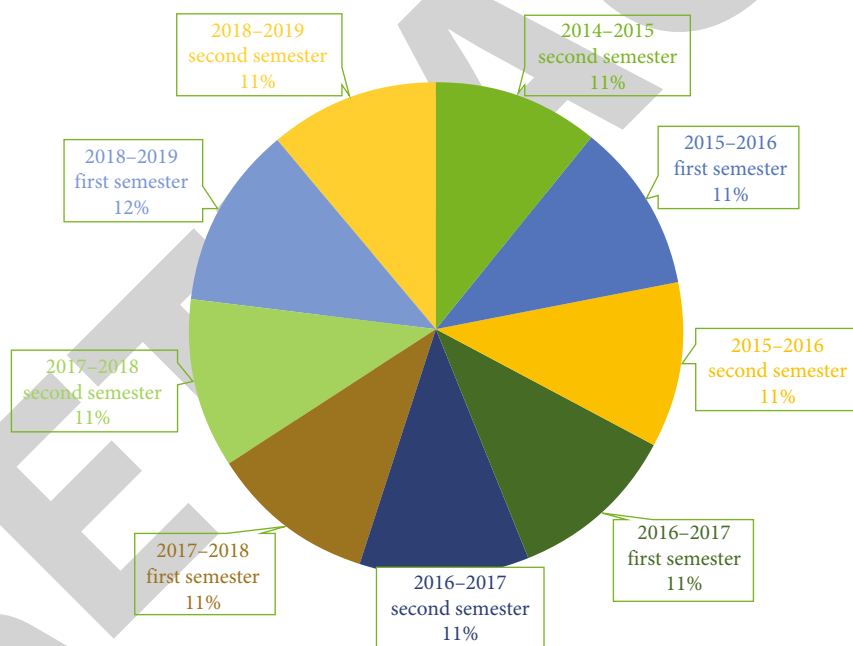


FIGURE 9: Changes in the proportion of three types of teachers in Y school in the past five years.

teachers with a timely understanding of their own teaching situation and that of other teachers, so that they can take targeted measures to improve their teaching. On the other hand, it provides teachers with a timely understanding of their own and other teachers' teaching conditions, so that they can take targeted measures to enhance the internal driving force for continuous improvement of teaching

4. Analysis of Simulation Results

The teaching status can reflect the teaching operation of a university as a whole, but this whole is also composed of

individual teachers, so it can also reflect the overall teaching ability of teachers in a university. Based on the improved *K*-modes clustering algorithm, the teaching situation of the university was analyzed by clustering the teachers of the whole university, the teachers of secondary teaching units, and the individual teachers at three levels based on the time period of academic year semester.

The change in the percentage of teachers in the three categories is shown in Figure 6. The percentages of teachers in the three categories in each semester were basically equal, with the first category (at least 3 out of 4 indicators rated "good" or higher) accounting for at least (43%, 45%), and

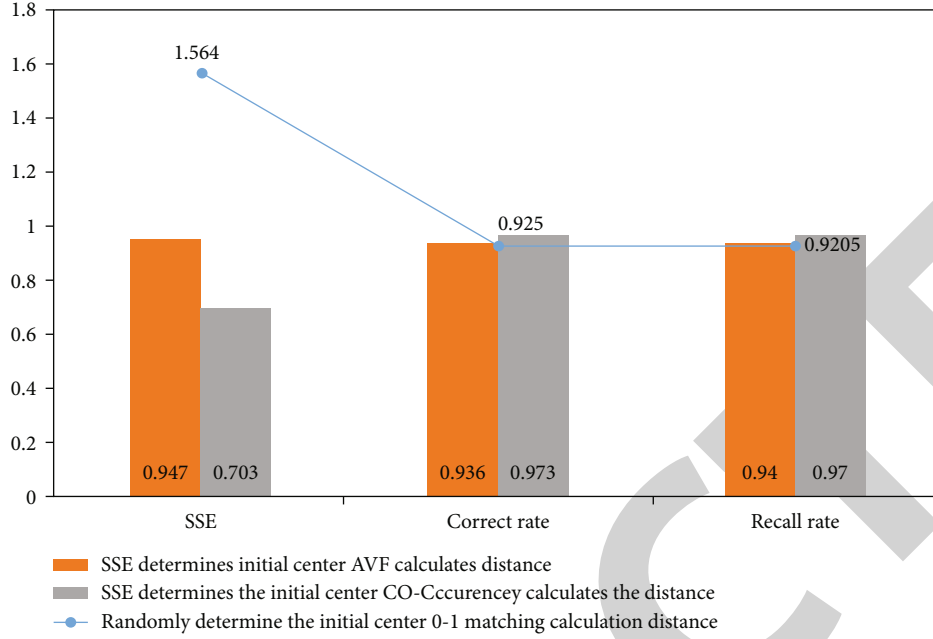


FIGURE 10: Comparison of three different clustering algorithms for each index in teaching evaluation data.

the second category (2-3 out of 4 indicators rated “good”), “good” grade. The second category (at least one of the four indicators is rated as “moderate”) is within the range of (26%, 28.5%); the third category (all four indicators are rated below “moderate”) is within the range of (26%, 28.5%). It can be seen that although students’ evaluation groups are changing, their perceptions of teachers’ teaching are basically the same, indicating that students’ evaluation is fair and just overall. Figure 7 shows the change in the proportion of first-class teachers over the past five years.

Figure 8 shows the proportion of class II teachers in school y in the past five years. Figure 9 shows the changes in the proportion of three types of teachers in school y in the past five years. From the clustering results, it is obvious that a little more than 40% (the first category of clustering) of teachers in the school are relatively qualified, nearly 30% (the second category of clustering) are in the basic qualified state, and nearly 30% (the third category of clustering) are not in the “basic quality” and “teaching.” This indicates that the overall teaching operation of the university is not optimistic and is far from the fundamental goal of establishing moral education and still needs to be reversed and improved with great efforts.

The results of three k -modes were compared in terms of cluster centers, number of clusters, percentages, sum of squares of least errors, correctness, and recall rates, using the evaluation of a semester at YR as an example.

$$\begin{aligned} \text{Correct rate} &= \frac{\left(\sum_{i=1}^k e_i\right)}{n}, \\ \text{Recall rate} &= \frac{\left(\sum_{i=1}^k (e_i/e_i + m_i)\right)}{k}. \end{aligned} \quad (7)$$

Figure 10 compares three different clustering algorithms for each index in the teaching evaluation data. From the experimental results of the three different algorithms, it is obvious that the improved K -modes outperform the first two algorithms in all metrics.

5. Summary and Outlook

Focusing on teaching operation and students’ autonomous learning, this paper uses the improved K -modes algorithm to cluster analyze the classroom teaching operation and preliminarily analyzes and converts the student evaluation data of a university. On this basis, the abnormal evaluation data is eliminated according to the improved cosine algorithm, and the normalization algorithm is used to normalize the data. The traditional K -modes algorithm is used to cluster the data, and the main problems are pointed out. Taking the real data of a university as the research object, this paper makes a targeted exploration from the two aspects of teaching operation and students’ online learning and draws some preliminary conclusions with certain reference value, which can form a more comprehensive, accurate, popularized, and influential result of students’ evaluation, expert evaluation, peer evaluation, supervision, and evaluation institutions, and daily teaching monitoring data. It will provide more targeted decision-making, reform, and innovation basis for education and teaching managers and front-line teachers. However, there are still some limitations in this paper. Research needs to determine the number of cluster family K . Determine the initial K cluster centers and then the distance measurement from the K cluster centers. In future research, K -modes algorithm can be improved in these three aspects to evaluate teachers’ teaching ability.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities of China University of Labor Relations: Study on contingency table analysis and implementation in EXCEL and SPSS (no. 21ZYJS018).

References

- [1] X. Chen, "The construction of teaching evaluation index system of programming courses in primary and secondary schools based on AHP," *Learning and Education*, vol. 10, no. 1, 2021.
- [2] Z. Cailian and G. Yao, "Retraction note to: mountain rainfall estimation and online English teaching evaluation based on RBF neural network," *Arabian Journal of Geosciences*, vol. 14, no. 22, 2021.
- [3] S. O. Kim and S. M. Park, "Analysis of perception gaps regarding a university course and teaching evaluation between students and faculty groups: based on a case study of a course and teaching evaluations at K university," *Korean Association For Learner-Centered Curriculum And Instruction*, vol. 19, no. 18, pp. 501–520, 2019.
- [4] M. Yong-tong and M. Zhou, "Facing big data information fusion and data mining technology to construct college physical education teaching evaluation system," *Journal of Sensors*, vol. 2021, Article ID 7168855, 15 pages, 2021.
- [5] Z. Chenyang and A. Gudamu, "Ranking of teaching evaluation index system of public physical education curriculum based on management by objectives using analytic hierarchy process," *Mobile Information Systems*, vol. 2021, Article ID 2729574, 15 pages, 2021.
- [6] Z. Xiaoying, "Teaching evaluation algorithm based on grey relational analysis," *Complexity*, vol. 2021, Article ID 5596518, 9 pages, 2021.
- [7] Z. Jing, "Research on classroom teaching evaluation and instruction system based on GIS mobile terminal," *Mobile Information Systems*, vol. 2021, Article ID 9790766, 11 pages, 2021.
- [8] X. Liu, X. Cao, R. Yang, and L. Zhang, "Physics practice teaching evaluation system based on organization and end-user calculation," *Advances in Educational Technology and Psychology*, vol. 5, no. 9, 2021.
- [9] X. Xiao, "Research on higher education evaluation based on fuzzy comprehensive evaluation method," *International Journal of Education and Management*, vol. 6, no. 4, 2021.
- [10] L. Yueguang, "The application of fuzzy comprehensive evaluation algorithm in college classroom teaching evaluation," *Education*, vol. 11, no. 4, pp. 1340–1347, 2021.
- [11] F. R. Vicente, A. Z. Llinares, and N. M. Sánchez, "Curriculum analysis and design, implementation, and validation of a STEAM project through educational robotics in primary education," *Computer Applications in Engineering Education*, vol. 29, no. 1, pp. 160–174, 2021.
- [12] L. Cao, "Construction of project-based teaching evaluation index system for computer courses in secondary vocational schools," *International Journal of Social Science and Education Research*, vol. 5, no. 3, 2022.
- [13] I. Aartun, K. Walseth, Ø. F. Standal, and D. Kirk, "Pedagogies of embodiment in physical education – a literature review," *Sport, Education and Society*, vol. 27, no. 1, pp. 1–13, 2022.
- [14] C. Shen, "Higher education evaluation based on principal component analysis and time-series analysis," *Frontiers in Economics and Management*, vol. 2, no. 7, 2021.
- [15] P. Pandey and K. Sharma, "Implementation of rubrics in assessment of teaching from students' and teachers' perspectives," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, no. 4, pp. 83–88, 2020.
- [16] Y. Wang, P. Anand, K. S. Cheung Simon, H. C. Ching, and D. Sadia, "Similarity detection of English text and teaching evaluation based on improved TCUSS clustering algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 7555–7565, 2021.
- [17] Q. Liang, Y. Pang, X. Xiong, and D. Pan, "Research on higher education evaluation system based on AHP-NBM comprehensive evaluation model," *Frontiers in Educational Research*, vol. 4, no. 4, 2021.
- [18] S. J. Curran, "Pitfalls in using small number statistics in teaching evaluations: a case study," *Assessment and Evaluation in Higher Education*, vol. 45, no. 3, 2020.
- [19] C. Pong-inwong and W. Songpan, "Sentiment analysis in teaching evaluations using sentiment phrase pattern matching (SPPM) based on association mining," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 8, pp. 2177–2186, 2019.
- [20] J. Qiu, "Research on the problems and countermeasures of undergraduate teaching fine management based on teaching evaluation," *Frontiers in Educational Research*, vol. 2, no. 6, 2019.