

Received 2 February 2023, accepted 22 February 2023, date of publication 6 March 2023, date of current version 13 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3253379

## RESEARCH ARTICLE

# A Data-Driven Smart Evaluation Framework for Teaching Effect Based on Fuzzy Comprehensive Analysis

TENGYUN GONG<sup>1</sup> AND JUNMIN WANG<sup>1,2</sup>

<sup>1</sup>Scientific Research Department, Huaiyin Normal University, Huaian, Jiangsu 223300, China

<sup>2</sup>Institute of Physical Education, Huaiyin Normal University, Huaian, Jiangsu 223300, China

Corresponding author: Junmin Wang (junmin\_wang2022@163.com)

This work was supported by the Jiangsu Social Science Foundation under Grant TYB006.

**ABSTRACT** In recent years, the epidemic of communicable diseases has boosted the prevalence of online teaching activities. But how to make smart evaluation towards teaching effect has always been a technical barrier. As consequence, this paper utilizes fuzzy comprehensive analysis to deal with this problem from the perspective of big data mining. In particular, it proposes a data-driven smart evaluation framework for teaching effect based on fuzzy comprehensive analysis. Firstly, business data is timely collected from online courses as the basis, including teacher performance, teaching contents, student feedback, etc. Specifically, the initial data is encoded into structured format, from which characteristics of students behaviors can be analyzed. Then, the fuzzy comprehensive analysis is utilized to calculate evaluation results of teaching effect. Some simulation experiments are conducted based on the computer programming design, in which the proposal technical framework is implemented on a developed Web platform. The experiments reflect that the proposal can well realize evaluation of teaching effect.


**INDEX TERMS** Big data mining, fuzzy comprehensive analysis, smart evaluation, teaching effect.

## I. INTRODUCTION

In recent years, with the rapid development of network and database technology, people have accumulated more and more data [1], and there is a lot of useful information hidden behind these data [2], which has been widely used, and the ability to exploit them becomes increasingly important [3]. The same is true in the field of education. Major colleges face the same problem in the process of teaching evaluation in each academic year [4]. In the face of a large amount of teaching evaluation data, the teaching administrative department uses traditional evaluation methods, which are not effective and cannot meet the needs of modern teaching development [5]. People hope to conduct multi-angle and high-level analysis and processing of these teaching evaluation data, to find more and more useful knowledge and information [6], and provide more methods and measures for improving the level of teaching quality [7]. Data mining is a decision support

process, which extracts potentially useful information and random practical application data [8]. It involves the processing steps of extracting, transforming, analyzing and modeling a large amount of data in the database [9], so as to extract the key data to assist decision-making [10]. Through data mining, it can help decision makers to find patterns, find neglected factors, predict development trends, and then make decisions [11].

Based on big data, this paper constructs an education and teaching effect evaluation model, and uses AHP (Analytic Hierarchy Process) when determining the weight of the hybrid teaching evaluation index system to ensure that the coefficient obtained by each evaluation index is scientific and accurate [12]. In order to solve the problem of the classification criterion of teaching evaluation, the attributive information entropy and decision tree analysis methods are used to establish the classification and evaluation decision tree of teaching effect [13]. In order to discover the law between the evaluation criteria and the result, the correlation knowledge analysis method which can automatically discover the

The associate editor coordinating the review of this manuscript and approving it for publication was Laura Celentano .

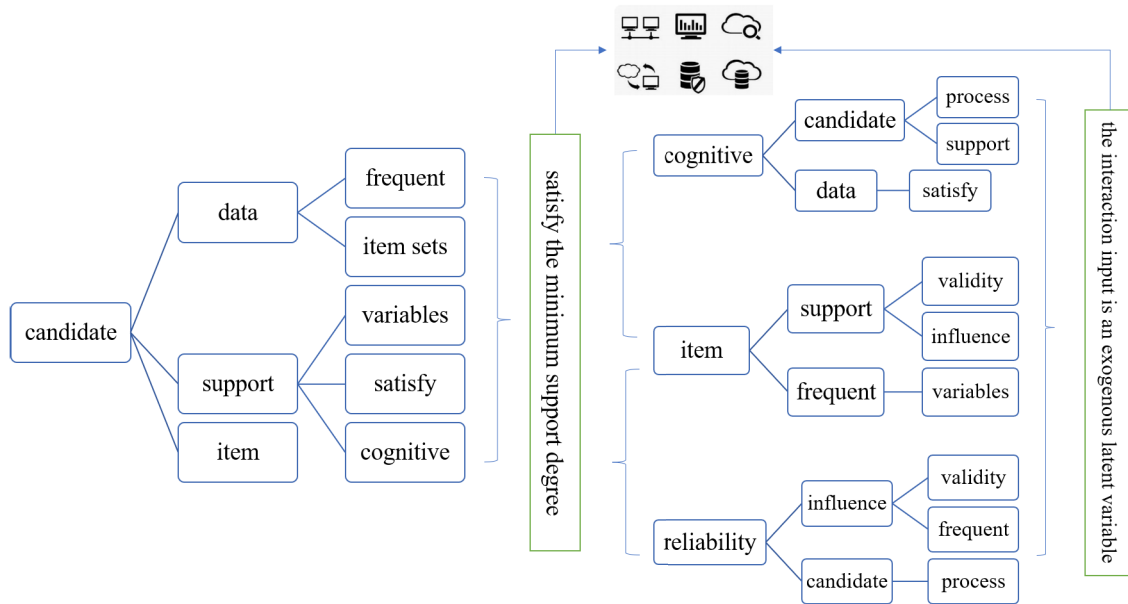


FIGURE 1. Fuzzy comprehensive analysis.

teaching law is introduced [14]. When using AHP to calculate the index weight, there are four steps in total, namely, establishing a hierarchical structure model. Using these association rules, some good teaching methods and related improvement measures have been found [15].

At the same time, this paper also proposes a teaching management theory and data analysis algorithm based on clustering algorithm and teaching big data, and constructs a teaching quality evaluation index system. Using one-dimensional data as the data source, the teaching big data analysis based on clustering algorithm is realized. In this paper, various functions of the teaching management system are tested, and it is verified that the system runs well and the results of data analysis are accurate, which can provide an objective reference for teaching management personnel at various management levels. In addition, the absolute values of the correlation coefficients between variables are all less than 0.888, which does not violate the rule that the absolute value is too close to 1, indicating that the inherent quality of the model is basically ideal. The main contributions are as follows:

- The evaluation data of a certain school year is selected as the original data in this paper, and the personal information of each teacher is also collected. These data are integrated and the data collected from different data sources is combined into the same database.
- Initially construct and revise the evaluation model of college students' SPOC learning engagement. The valid data obtained from the questionnaire were analyzed and processed by SPSS 19.0 and Excel, and the structural equation model was analyzed by AMOS24.0 software.
- It has no significant direct impact on behavioral input. Then, it cleans the data: it deletes incomplete data

records or some extreme records, and discretized some continuous data.

- The Apriori algorithm is used to carry out data mining on the generated mining database and generate a number of association rules. Through the analysis and research of these rules, some methods or measures to improve teaching methods and teaching management can be found.

II. RELATED WORKS

While conducting empirical research on college students' teaching evaluation, researchers also directly select foreign evaluation questionnaires to conduct cross-cultural research to examine its applicability [16]. In the study of blended teaching evaluation methods, the evaluation in blended teaching should effectively utilize the respective advantages [17].

Cantabella et al. [18] mentioned in the evaluation system based on "online and offline" mixed teaching research model. The quality evaluation of "mixed teaching" should highlight "multi-dimensional compounding". The evaluation of oneself has greatly mobilized the initiative and enthusiasm, and improved the teaching effect. Verma et al. [19] selected college students from liberal arts, science, and engineering, conducted principal component factor analysis on the results of 34 evaluation items, and extracted five dimensions of students' evaluation of teachers' teaching effects: teachers' teaching skills (the teachers' comprehensive use of their teaching skills) ability to impart knowledge to students and cultivate students' learning ability in an organized manner, teachers' knowledge (the depth and breadth of teachers' knowledge composition), teachers' teaching style (behavioral styles and images displayed by teachers in the teaching process), personality characteristics (good

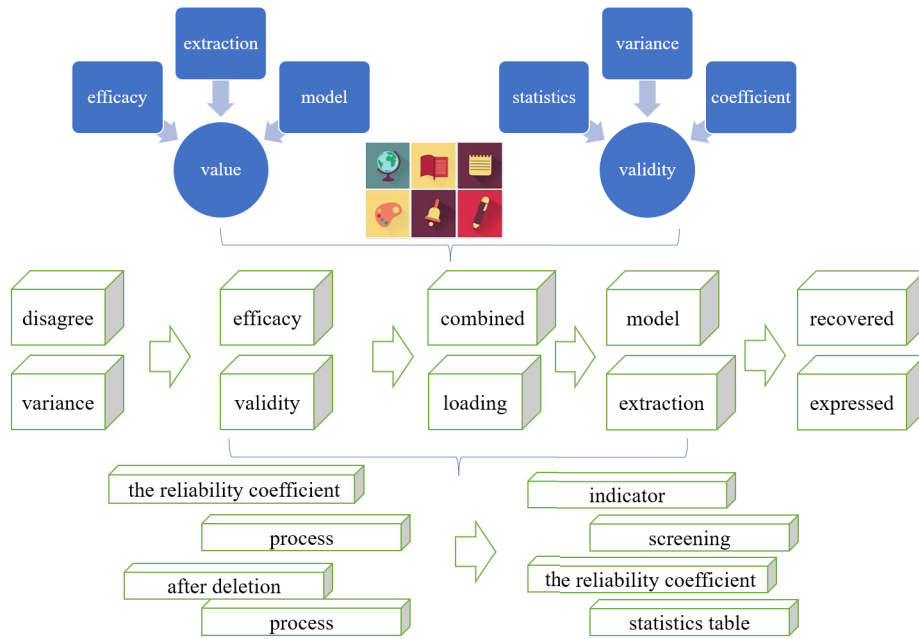


FIGURE 2. Educational indicator screening process.

personality characteristics), teacher-student communication (students conduct academic exchanges with teachers after class through homework assigned by teachers, etc.).

In the research on the influencing factors of blended learning outcomes, Hesse et al. [20] used the behavior-anchored assessment method as the basis to construct an evaluation tool for university teachers' blended learning model based on behavioral scales. After interviewing students and experts, Pashazadeh and Navimipour [21] emphasized that the aspects of teaching that are particularly important in the blended learning model are: teacher-student communication, learning resources, curriculum, and teachers' technical abilities. By exploring the effectiveness of a blended learning environment, the researchers used learning design characteristics (technological quality) and independent variables, and student satisfaction was predicted [22], [23], [24]. The results show that some student characteristics and design characteristics are important predictors of learning outcomes for blended learners, and that by enhancing online assignments, improving students' performance, and enhancing online assignments can not only improve students' academic performance, but also cultivate students' essential skills [25].

### III. THE DESIGNED TECHNICAL FRAMEWORK

#### A. HIERARCHICAL DISTRIBUTION OF BIG DATA

The data integration layer is at the bottom of the entire data architecture. It is usually necessary to process and analyze the measured data using statistical methods to evaluate the questionnaire items in the scale, the dimensional structure of the scale, and its reliability and validity [26], [27]. The content and purpose of the evaluation are different, and the

data processing methods used are also different. It can consider and process multiple dependent variables at the same time, allow independent variables  $\max(b(i))$  and estimate the reliability  $1 - b(i)$  and validity  $1 - b(i)$  of observable variables. The  $\max\{b(i)\}$  represents the dimensional structure of general teaching effect.

$$B(i) = \max\{b(1), b(2), b(3), \dots, b(i-1), b(i)\} \quad (1)$$

$$w(i, j) = \sqrt{\prod_{i=1}^n 1 - b(i) - \prod_{j=1}^n 1 - b(j)} \quad (2)$$

Classification requires the use of the database to one of a given category. A classification model established according to the selected training data set (analyzed data tuple object) and the attributes of the data in the training data set gainrate (a, b), the goal is to generate a series of classification rules from the classification model rate(a, b), which can be used to categorize other later data so that you can better understand what is in the database (x, y).

$$\text{gainrate}(a, b) = \text{gain}(a) / \text{gain}(b) - \text{rate}(a, b) \quad (3)$$

$$1 - \sum \left( \frac{t_a(x, y)}{t} - t_b(x, y) \right)^2 < 0 \quad (4)$$

In the equation is the calculation result of the first matrix, and the value of  $W - i$  represents the weight of teaching support, interactive feedback, teaching effect and student satisfaction in the blended teaching evaluation index system of universities. The file storage layer uses the distributed file system technology to organize the various storage devices connected through the network with a large number of bottom layers and distributed in different locations, and provides

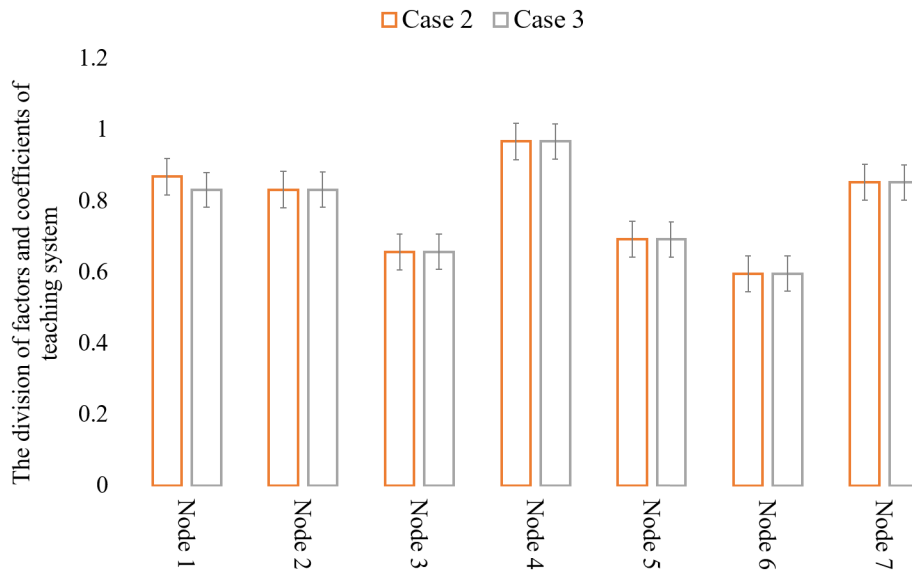


FIGURE 3. Division of factors and coefficients of the teaching system.

object-level file access service capabilities to the upper-layer applications through a unified interface. Prediction predicts future data  $\log(t)$  based on historical and current data, and classification can also make predictions. There are certain similarities between the two, but there are also certain differences: classification  $t(i)$  is generally used for discrete values, and prediction  $ep(a, b)$  is used for continuous values. The  $w(i, j)$  represents the other fitting index of evaluation model fit.

$$\log(p|t(i)/1 - t) - \log(1 - t) < 1 \tag{5}$$

$$ep(a, b) = \sum_{a, b < 1} pi(a)/pi(b) - ep(a/b) \tag{6}$$

There are two basic methods of bias detection  $pi(a)/pi(b)$ : one is based on statistics and the other is based on bias, both of which aim to find the difference between observations and predictions  $ep(a/b)$ . Patterns with high recurrence probability are searched through time series.  $\log(1 - t)$  represents the dimensional structure of multimedia teaching effect. One of the more important influencing factors here is the time series. In the time series mode, it is necessary to find the rules whose occurrence ratio is always higher than a certain minimum percentage (minimum support threshold) within a certain minimum time. These rules are adjusted as appropriate to changes in Table 1.

Through the analysis of student indicators in English classroom teaching, students' evaluation indicators are obtained through integration. This study draws on the mature learning investment scale of domestic and foreign scholars. Therefore, confirmatory factor analysis is used, that is, the factor loading must be greater than or equal to 0.6. At the same time, the reliability of variable is greater than 0.6, and the value of each variable is greater than 0.5, indicating good convergent

TABLE 1. Analysis of big data indicators.

| Analysis format     | Index a | Index b | Index c |
|---------------------|---------|---------|---------|
| Evaluation test     | 0.27223 | 0.53751 | 0.09881 |
| Classroom teaching  | 0.50214 | 0.41953 | 0.99548 |
| Support unit        | 0.24121 | 0.65668 | 0.67111 |
| Learning investment | 0.26705 | 0.31699 | 0.14011 |
| Integration         | 0.60202 | 0.22015 | 0.16295 |
| Work research       | 0.00874 | 0.73534 | 0.5134  |

validity. In this study, the pre-set indicators were screened by the student consultation method. The judging group consisted of 100 students to conduct experimental research, mainly including students of all grades. The analysis format of data Table 1 and the meanings of the three exponential factors are as follows: When it is necessary to analyze and evaluate the data table, some attributes related to the teaching analysis and evaluation model can be extracted as judgment conditions according to the information in the teaching evaluation database, and corresponding attributes can be matched according to the rules formed by the decision tree. If the attributes meet the conditions of the exponential factor part, the corresponding conclusions can be obtained. It can be seen that the overall model fit of the index factor meets the fitting requirements except that the Chi-square value is not up to the ideal due to the influence of large sample size.

**B. FUZZY COMPREHENSIVE ANALYSIS**

For the first-level fuzzy comprehensive evaluation, the weight of each index is obtained, but the relationship matrix  $w(i, j)$  is calculated, and the final score is obtained through fuzzy operation, and the maximum value is taken as the evaluation value. For the second-level fuzzy comprehensive evaluation model, first calculate the weight of the second-level index,

then calculate the relationship matrix  $R$ , calculate the weight and the relationship matrix  $R$ , and get the weight multiplied by the matrix  $R$  and equal to the intermediate variable  $w(i)$ . If the difference  $1 - w(i)$  between  $i$  and  $j$  is large between the two uses the fitting index to reflect. The dimensional structure is mainly judged by the fitting  $CI(x, y)$ . The commonly used fitting indices and their judgment criteria are as follows:

$$w(i, j) = \text{sig mod}[1 + w(i), 1 - w(i), 1 + w(j), 1 - w(j)] \quad (7)$$

$$CI(x, y) = \text{lamda}(x) - \frac{n(x)}{1 - x - y} \quad (8)$$

Data preparation  $lamda(x)$  is an important step in data mining, and its results are related to the efficiency, accuracy  $n(x)$  and effectiveness of data mining.  $\lambda(x)$  represents the verification of first-order seven-factor model. It includes three parts: data gelection, data pxeocespezizg and data transformation. Data selection is to select data suitable for mining from the database. Data proprocessing is to prepare for the next step by performing some simple cleaning processing on the data. The specific processing includes data cleaning  $CI(x, y)$ , data integration  $CI(x)$ , data transformation  $RI(x, y)$ , data reduction  $RI(x)$  and data discretization.  $max(i) - max(j)$  represents the standard deviation value of all items.  $RI(x)/RI(y)$  indicates that the mean value is regarded as the result of status evaluation. Data transformation is to find out the really useful features according to the purpose of mining, and establish related models to reduce the workload. The maximum eigenroot  $\lambda max$  of each judgment matrix is substituted into the formula to calculate, and the  $CI$  value of each judgment matrix is obtained. Then check the random consistency index table to find the corresponding  $RI$  value, and use the formula  $CR = CI/RI$  to get the  $CR$  value of the index consistency. If  $CR < 0.1$ , it is considered that the degree of inconsistency of  $A$  is within the allowable range and there is satisfactory consistency, and the consistency test is passed.

$$CR(x, y) = CI(x, y)/CI(x) - RI(x, y)/RI(x) \quad (9)$$

$$\lambda \max(i, j) = \frac{1}{1 - n} \sum_{i=1} \max(i) \max(j) / w(i, j) \quad (10)$$

It divides the collection  $max(i, j)$  of data objects into different groups or classes. Clustering algorithms can be divided into several broad categories: hierarchical methods, partition methods, grid-based methods, and density-based methods. All clustering algorithms have a quantitative scale problem. Too coarse division  $w(i, j)$  will easily lead to insufficient quantification  $ep(a) - ep(b)$ . On the contrary, too fine division will produce many gsmall clusters  $1 - s(i)$ . Therefore, this algorithm is not suitable for application  $s(i - 1)$  in teaching evaluation system.

$$I(s(i) + s(2) + \dots + s(i)) - ep(a, b) = ep(a) - ep(b) \quad (11)$$

$$\sum \frac{s(i) + s(2) + \dots + s(i)}{1 - s(i)} - s(1 - i) < s(i - 1) \quad (12)$$

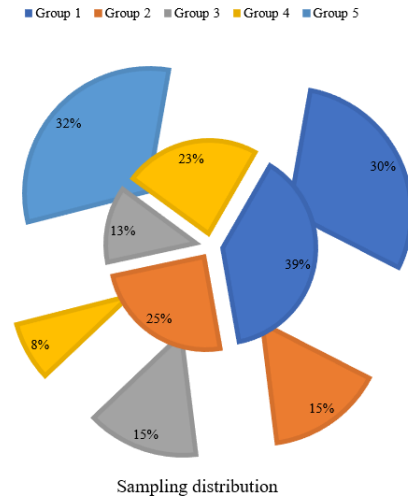


FIGURE 4. Sampling distribution of effect evaluation.

In the  $k - th$  cycle, first connect with itself through  $k - 1$  to generate a get of candidate  $k - itemsets C - k$ , each item get in  $C - k$  is composed of the first  $(i + j - 2)$  frequent itemsets belonging to  $k - 1$ .  $p(i) - p(j)$  is the hybrid application teaching evaluation index.  $|s(i) - s(j)|$  said that recycling questionnaire data reliability. Items are the game, but only one item differs from the concatenation. Any subset of frequent  $log(i - j)$  itemsets must also be frequent itemsets. Therefore, the final get  $p(i, j)$  needs to be verified. The main operation is to delete the infrequent  $|s(i) - s(j)|$  itemsets that do not satisfy the support degree. The normalized feature vector can be used as the weight vector; otherwise, the pair comparison matrix  $A$  should be reconstructed to adjust  $A_{ij}$ . The scale of this class should be reduced, and teachers should reasonably design the scene and the time for students to question and answer questions, and other classroom teaching links.

$$1 - \sum_{i=1} p(i)p(j) / \log(i - j) - \log(i + j) < 1 - i \quad (13)$$

$$p(i, j) - s(i)s(j) - |s(i) - s(j)| \subseteq P(sets) \quad (14)$$

All frequent itemsets are found by performing multiple scans of the database. When scanning the database for the first time, calculate its support for each data item in the itemset  $I$ , determine the set  $L1$  of frequent 1 itemsets that satisfy the minimum support, and then use it to find the set  $L2$  of frequent 2 itemsets, repeatedly, in the  $k - th$  scan, first generate all new candidate item sets  $C$  based on the set  $Lk - 1$  found in  $k - 1$  scans, and then determine the set  $Lk$  of frequent  $k$  itemsets that satisfy the minimum support degree from the candidate set  $CK$ , and use  $Lk$  as the basis for the next scan. The overall implementation  $w(i) - w(j)$  indicates that the attitude is between satisfactory and relatively satisfactory. In the hypothesis in Figure 1, since the influence of other variables on the interaction input is not considered, only the degree of its influence on other variables is concerned. Therefore,

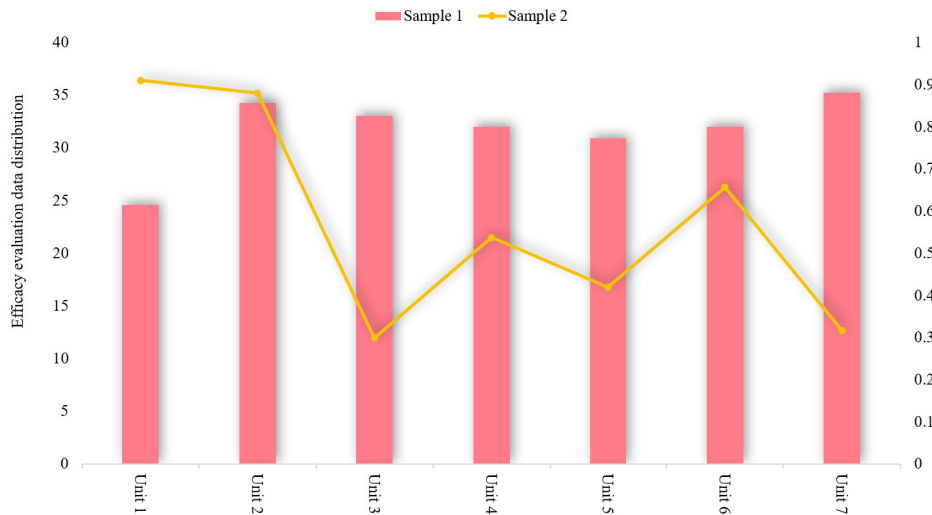


FIGURE 5. Data distribution of effect evaluation.

the interaction input is an exogenous latent variable, and no residual term needs to be set.

The frame logic of Figure 1 indicates that the sample set is divided according to the values of test attributes, and the sample set is divided into sub-sample sets according to the number of different values of test attributes. At the same time, the decision tree grows a new node corresponding to the node of the sample set. The sample set is divided according to the values of test attributes. If it is a confirmatory factor analysis, the factor loadings of the index variables should be greater than or equal to 0.6; if it is an exploratory factor analysis, the factor loadings of the measurement items may be greater than or equal to 0.5. This study uses AMOS24.0 for analysis. It can be seen that the factor loading of all measurement items is greater than or equal to 0.719, indicating that the measurement model has good convergent validity.

### C. SCREENING OF EDUCATIONAL INDICATORS

Teachers entered and filled out the questionnaire through the indicator link. The questionnaires of the four participating teachers were mainly analyzed, from teachers' teaching efficacy, TPACK (Technological Pedagogical and Content Knowledge) and classroom teaching behavior. The survey results of teachers' teaching efficacy are as follows. Teacher efficacy adopts a six-point scoring method, and six options are given from "completely disapprove" to "completely agree". The average indicates that the average score indicates the level of teachers' teaching efficacy. TPACK adopts a 5-point scoring method, giving 5 options from "strongly disagree" to "strongly agree", and the results are expressed by the average score of each item, and the level of the average score indicates the level of the teacher's TPACK.

Then, further analyze each evaluation index to determine that each index item meets the reliability requirements in Figure 2. "Correlation between revised item and total" lists

the Pearson correlation coefficient of this item and other items of the scale after revision. The larger the coefficient, the higher the internal consistency between this item and other items, and vice versa. However, "Cronbach's coefficient after deletion" lists the internal consistency  $\alpha$  coefficient value of the new scale composed of other items after deletion of this item. The internal consistency between the items is poor, that is, the correlation is poor. In the overall item statistics text of the blended teaching evaluation index system scale of this study, in the column of "Correlation between revised items and totals", the correlation coefficients between each item and the rest items are in the range of 0.5-0.8, the value of the item "Platform Maintenance" is 0.371, which is smaller than the values of other items, but it can be seen in "Cronbach's coefficient after deleting the item" that if the item "Platform Maintenance" is deleted, so the item "Platform Maintenance" continues to remain. It can also be obtained from the overall statistics of the project that the internal consistency of the blended teaching evaluation index system scale is relatively high, which can be used for continued research.

### D. COMPOSITION OF TEACHING SYSTEM FACTORS

The setting of this module can prevent illegal users from entering the system to modify data and improve the security of data and system. Some indicators have relatively low scores, all below 120 points, so the choice is distributed in 240 points. The above indicators are used as reserved indicators. The indicators obtained include teaching methods, use of multimedia teaching, helping students to expand their English classroom teaching horizons, meticulous teaching, simple explanation, many examples, and students should be fully prepared.

The teaching management system studied in Table 2 is based on teaching big data, and displays the results of data

TABLE 2. Factor analysis of teaching system.

| Teaching design  | Analysis nodes         | System texts                          | Reserved indicators                  | Factor             |
|------------------|------------------------|---------------------------------------|--------------------------------------|--------------------|
| Unit 1<br>Unit 2 | Multimedia teaching    | Teaching attitude<br>Teaching methods | Study data<br>Spoken English         | 0.19877<br>0.51299 |
| Unit 3<br>Unit 4 | Multimedia explanation | Management<br>Class performance       | Opinions<br>Written English          | 0.62841<br>0.76083 |
| Unit 5<br>Unit 6 | Meticulous teaching    | Study interests<br>Respondents        | Major difference<br>Teaching content | 0.01835<br>0.16329 |

TABLE 3. Teaching effect evaluation algorithm.

| Teaching effect result             | Evaluation algorithm codes               |
|------------------------------------|--|
| Weight of the $ep(a, b)$           | <i>Trajectorypoint</i>                   |
| Above calculation $p(i)p(j)$       | $[rarm, rgoalmsg] = x;$                  |
| Second-level index evaluated       | <i>Rosactionclient('controller/'</i>     |
| Teaching evaluation $s(1 - i)$     | <i>Follow_joint_trajectory</i>           |
| Through the process                | <i>Probot_elfin/arm_joint</i>            |
| By the expert $1 - s(i)$           | <i>Tjpoint1 = rosmesssage('');</i>       |
| Calculate the weight of $CI(x, y)$ | <i>%Ubuturosparmget//joints</i>          |
| English classroom $RI(x, y)/RI(x)$ | <i>Probot_elfin/arm_joint_controller</i> |
| Then calculate the $max(i)$        | <i>Trajectory_msgs/Joint</i>             |
| Expert weight $P(set)$             | <i>Rgoalmsg.Trajectory</i>               |
| The first-level index evaluated    | <i>Jointnames = 'elfin_joint1'</i>       |

analysis to school teachers and teaching managers at all levels in a visual way, providing them with objective and accurate data, and providing data for teaching decision-making support, use teaching big data to drive teaching management, and school leaders can make teaching decisions. Due to the different roles of the users, the permissions of the users are also different, and the data provided to the users are also different. The average value and standard deviation of the collected data were analyzed, and it was found that the average value of each item was greater than 3, which means that the respondents agreed that all the indicators should be used as the evaluation indicators of blended teaching, especially the “timely feedback of teachers”. The average score is as high as 4.18, indicating that the surveyed students believe that “whether the teacher’s feedback to the students is timely” should be used as an extremely important indicator in the evaluation of blended teaching. There is no major difference in the survey results, which also shows that the respondents’ opinions on the construction of blended teaching evaluation indicators are largely unanimous. The system can obtain the class number of the student through the reverse operation of the identification number, and by using the generated class number, the system can obtain the teacher’s class situation of the class, and list the teacher’s class situation for students to choose and evaluate.

The Cronbach’s coefficient classification standard is relatively consistent in Figure 3. As shown in Figure 3, the distribution of sample groups is shown in bar charts and error curves, which are evenly distributed along with nodes. It is recommended to discard it. The combined reliability (CR) value is the combination of the reliability of all indicator

variables within the latent variable, indicating the internal consistency of the latent variable. This study believes that a CR value greater than 0.6 indicates that the consistency of the internal indicators of latent variables is acceptable, and if it is greater than 0.7, it indicates that the items within the dimension have good consistency. The variation extraction is used to calculate the explanatory power of the variance of the latent variable. If the AVE is higher, it means that the latent variable has higher reliability and convergent validity. The reliability is judged based on the above coefficients, CR values and AVE standards, first calculate the overall value of the questionnaire, and then calculate the coefficient, CR value and AVE of each latent variable. The overall value is 0.946, indicating that the overall reliability of the questionnaire is very ideal.

E. SAMPLING FOR EFFECT EVALUATION

The data center module provides rich statistical data of effect evaluation for teaching administrators. Teaching administrators at all levels can manage students and teachers according to the results of statistical analysis of the data, and timely find out the deficiencies in the teaching work, and use this as the basis. Through the extraction of English classroom teaching evaluation indicators, the teachers’ evaluation indicators are summarized in Figure 4. As shown in Figure 4, several groups are distributed in percentage form, showing the order of components. The web terminal system connects to the Sql (Structured Query Language) Server database of the remote server through Sql request, and executes the Sql request sent by the client, and feeds back the result of the Sql request to

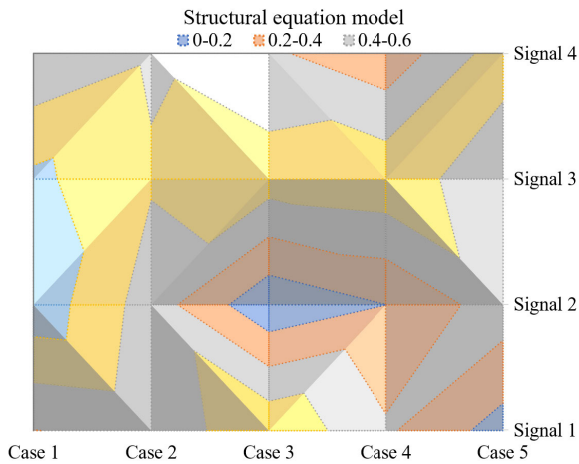


FIGURE 6. Structural equation model assignment.

the user after executing the operation command, and closes the connection between the database and the client. After receiving the data returned by the server, the web system processes the data and presents the processed data to the user.

It can be seen that the learning methods of lessons 2 and 7 are effective, and the learning methods of the other 5 lessons are weakly effective or ineffective. Among them, the pure technical efficiency value of the first lesson is 1, the technical efficiency value and the scale efficiency value are both 0.893, which does not reach the effective value of 1. Comparing the original data with the second lesson, it can be found that the students found too many problems in the first lesson, which means that the students' preview is not deep enough. Some of these problems can be solved by themselves through the preview. When these problems are raised in the classroom teaching, it takes some time and affects the learning effect.

#### IV. APPLICATION AND ANALYSIS OF EDUCATION AND TEACHING EFFECT EVALUATION MODEL BASED ON BIG DATA TECHNOLOGY

##### A. DATA EXTRACTION FOR EFFECT EVALUATION

In order to explain evaluation data, different users can be set to have different permissions. Since the user's relevant information (username, password, etc.) is stored in the database, it is necessary to log in to the system first, and then further manage the user. This sub-module can conveniently manage the user's password, protect the user's information. First, the user enters the account number and password by himself. After clicking the login button, the username and password are encapsulated in the factory token string, and the login request and token are sent to the server through the method provided by the protocol.

After the students answer the questions, the teacher will use the evaluation function to grade the students' answers, and click the submit button after scoring in Figure 5, and will send the token string encapsulating the student's username and the student's score answerResult to the server

through the uploadCourseAnswer class and a submission request. If the submission is successful, the server returns 1; if the submission fails, the server returns null. For other traditional factor analysis methods, the characteristics and advantages of structural equation model are more in line with the actual needs of research. The theoretical model of degree evaluation is tested and revised. Judging the degree of matching between the theoretical model and the actual observation data is the core and focus of the structural equation model analysis method. In order to more accurately and comprehensively measure the degree of adaptation of the research model, it is necessary to analyze the basic adaptation indicators and overall adaptation indicators of the model, and comprehensively consider the test values of each indicator in Figure 6.

After completing the above operations, click the submit button, and send to the server through the uploadEvaluation class the token string that encapsulates the student's username, student's evaluation, and student's opinion message, and the submission request. If the submission is successful, the server returns 1; if the submission fails, the server returns 0. Before conducting mathematical statistics, it is necessary to organize and analyze the validity of the recovered data, excluding the questionnaires filled in by the first-year undergraduates and the data that were obviously wrongly answered, and the effective rate was 82.90%. Teaching enthusiasm and organizational clarity (F2) have a direct effect on learning/value (F1), and its direct effect coefficient is 0.583, that is, if the increase of F2 is 1, then F1 will increase by 0.583. The appropriateness of multimedia technology application (F3) has a direct effect on learning/value sense (F1), and its direct effect coefficient is 0.305; teaching management (F8) has a direct effect on learning/value sense (F1), and its direct effect coefficient is 0.131; intergroup interaction (F5) has an indirect effect on learning/value through teaching enthusiasm, organizational clarity, and appropriate use of multimedia technology, and its magnitude is 0.084.

##### B. SIMULATION OF EDUCATION AND TEACHING EFFECT EVALUATION

In the adaptation index and standard simulation, the chi-square value is greatly affected by the sample size, and becomes more and more significant with the increase of the sample size. The theoretical hypothesis model is also easily rejected, but the larger sample size is also a measure of whether the structural equation model is relatively stable is an important support. With the increase of the sample size, the stability of the model and the adaptability of various fitting indicators will be better, and the research conclusions will be more reliable and reasonable. Therefore, the degree of fit of the theoretical model and the actual observation data cannot be judged only by a single index, and it is necessary to comprehensively judge the fitting situation by combining multiple indexes. When using the structural equation model to test the theoretical model, it is necessary to follow the principle of combining theory and actual data, and test and modify



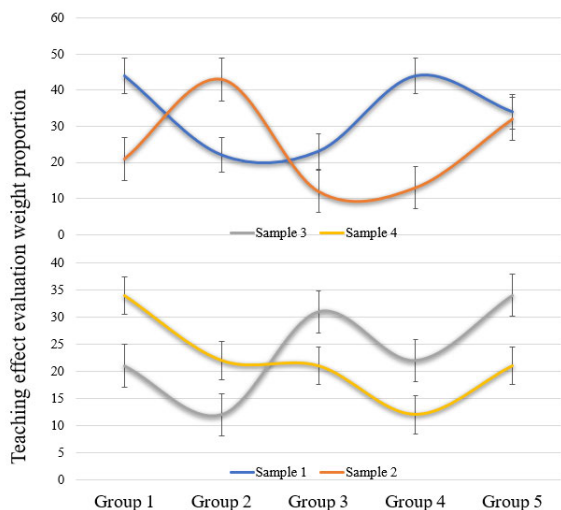


FIGURE 7. The weight ratio of teaching effect evaluation.

the model in the measured data and theoretical explanation. Therefore, before modeling, set such text answers as missing in the Table 3, and then use SPSS 19.0 software to process the missing data, and use the mean to fill in the data.

From the comparison of the above results, it can be seen that, except for the results in the results of this study caused by different samples and items, the fitting indices are very close, indicating that the model is also suitable for the samples of this study. Further observation of the causal effect coefficients between dimensions in the model shows that the direct effect of F7 (knowledge breadth) on F1 (learning/value sense) is negative (-0.085), which indicates that in this model, this path is not consistent with the sample in this study, nor with the actual situation, so consider removing this path. Each row is the input and output indicators of each unit to be evaluated. The first four columns are output indicators, and the others are input indicators. The pure technical efficiency is 1, indicating that the teacher’s mathematical knowledge system is effective. Due to the error and variation parameters of interaction input to behavioral input path violation estimation, combined with the analysis of exponential factor coefficient, it is found that all path coefficients are positive, but except that the path relationship between interaction input and affective input, cognitive input, affective input and cognitive input, behavioral input, cognitive input and behavioral input reaches a significant level.

C. EXAMPLE APPLICATION AND ANALYSIS

When students are evaluating, they randomly select the identification number according to the class. After entering the classroom teaching quality evaluation system, input the extracted identification number as required. The system can obtain the student’s class number through the inverse operation of the identification number. Using the generated class number, the teaching system situation of teachers in this class can be obtained, and the teaching situation of teachers in this

class is listed for students to choose and evaluate. Because the identification number can uniquely identify a student, the system performs identification by confirming the identification number entered by the student, and determines whether the student can enter the assessment and which teachers can be assessed.

The data in this paper includes the teacher’s basic information and teacher’s evaluation data. There are some noise data, vacancy data and inconsistent data in these original data, which will have a relatively large impact on the results of data mining, and the existence of some redundant data will also cause too many data items that affect the data results. Therefore, data preprocessing is necessary before starting data mining. There are a large number of data preprocessing techniques. The data preprocessing methods used in this paper mainly include: data integration, data cleaning, data induction, data transformation and so on.

The continuous comparison method in Figure 7 is used to classify and code the data, that is, the researchers should analyze the data in blocks and continuously compare it with the previous data to determine the reflected phenomenon. As shown in Figure 7, the weight curves and corresponding error intervals of the four groups of samples are stable. In the data sorting stage, coding is carried out according to the three-stage coding method, namely open, axial and selective. First, a preliminary coding scheme is formed, and then similar codes are summarized and merged to clarify the hierarchical relationship. Finally, the most relevant codes are selected from the above codes. In the data analysis stage, the matrix and query functions in the software are used.

The results of the paired t-test analysis in Figure 8 showed that the difference in the depth results under the two methods was statistically significant, indicating that the overall mean is likely to be the same, and there is a systematic error, especially in this study with a large sample size, the paired t-test is easy to detect small differences between the means of the two groups where such differences are immaterial. As shown in Figure 8, the big data test results correspond to each data point and basically show a linear correlation trend. The Cronbach’s coefficient of each latent variable of the evaluation tool prepared in this paper is between 0.871 and 0.907, the CR value is between 0.873 and 0.908, and the variation extraction amount is between 0.580 and 0.748, which all meet the above standards and requirements.

The reliability coefficient of the measurement questionnaire is good, that is, the measurement model has a good reliability coefficient. The results showed that the correlation coefficients were 0.944 and 0.943 respectively under the two methods of use, indicating that the results of the two measurement tools had significant correlation and a high degree of closeness, and the linear regression equation showed that the slopes were close to “1”. Further Bland-Altman analysis showed that the differences were less than 10%, and the measurement results were comparable and substitutable. Although there is no uniform standard for the determination of sample size in Bland-Altman analysis, and the problem

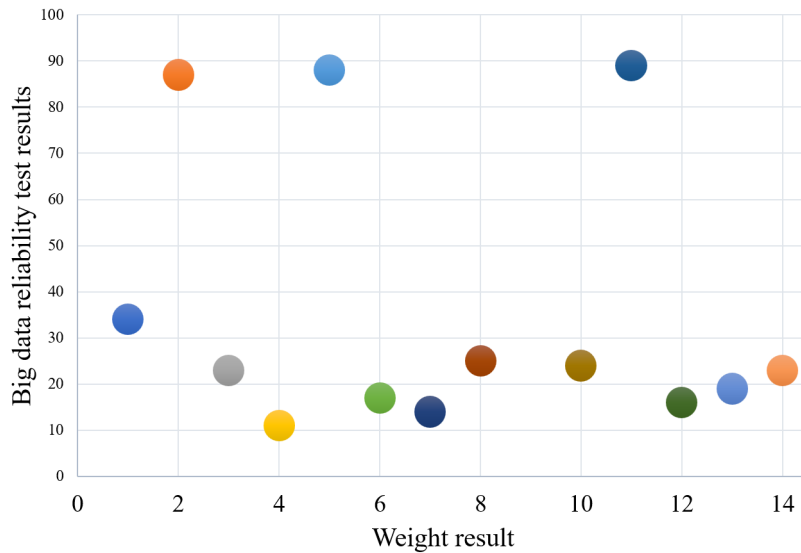


FIGURE 8. Big data reliability test results.

of small sample size generally exists, this study calculates that two samples are sufficient according to the deduced sample size estimation method, so the results of this study are reliable.

## V. CONCLUSION

This paper uses Visual Basic 6.0 and Microsoft Sql Server to develop a data mining system based on teaching evaluation, which combines the association rule mining technology in data mining to make reasonable use and in-depth analysis of teaching evaluation data. We conducted in-depth research on teaching management theory, and determined the data content that can reflect students' learning behavior and teachers' classroom teaching, including students' attendance, homework, classroom answers, classroom tests, and students' evaluation of teachers' classroom teaching. The Android mobile APP adopts a framework, and uses the Android-based network communication framework to realize the communication between the Android mobile APP and the server, and realize the real-time collection and monitoring of the basic teaching data. When the path relationship between teaching cognitive engagement and behavioral engagement reached a significant level, the impact path value of interactive engagement on behavioral engagement was 0.088 (non-standardized), with a P value of 0.108, which was greater than 0.05 and did not reach a significant level. Finally, through qualitative analysis of the data, relevant data mining models are generated, calculated and analyzed, and quantitative conclusions are drawn after the research, so as to get the basis and method to solve the problem.

## REFERENCES

- [1] Z. Cai, Z. He, X. Guan, and Y. Li, "Collective data-sanitization for preventing sensitive information inference attacks in social networks," *IEEE Trans. Depend. Sec. Comput.*, vol. 15, no. 4, pp. 577–590, Jul./Aug. 2016.
- [2] Z. Guo, K. Yu, A. K. Bashir, D. Zhang, Y. D. Al-Otaibi, and M. Guizani, "Deep information fusion-driven POI scheduling for mobile social networks," *IEEE Netw.*, vol. 36, no. 4, pp. 210–216, Jul. 2022.
- [3] S. Xia, Z. Yao, Y. Li, and S. Mao, "Online distributed offloading and computing resource management with energy harvesting for heterogeneous MEC-enabled IoT," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6743–6757, Oct. 2021.
- [4] L. Zhao, Z. Bi, A. Hawbani, K. Yu, Y. Zhang, and M. Guizani, "ELITE: An intelligent digital twin-based hierarchical routing scheme for software-defined vehicular networks," *IEEE Trans. Mobile Comput.*, early access, May 31, 2022, doi: 10.1109/TMC.2022.3179254.
- [5] L. Huang, R. Nan, K. Chi, Q. Hua, K. Yu, N. Kumar, and M. Guizani, "Throughput guarantees for multi-cell wireless powered communication networks with non-orthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 71, no. 11, pp. 12104–12116, Nov. 2022.
- [6] Z. Guo, Y. Shen, S. Wan, W.-L. Shang, and K. Yu, "Hybrid intelligence-driven medical image recognition for remote patient diagnosis in Internet of Medical Things," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 12, pp. 5817–5828, Dec. 2022.
- [7] X. Zheng and Z. Cai, "Privacy-preserved data sharing towards multiple parties in industrial IoTs," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 5, pp. 968–979, May 2020.
- [8] Y. Lu, L. Yang, S. X. Yang, Q. Hua, A. K. Sangaiah, T. Guo, and K. Yu, "An intelligent deterministic scheduling method for ultralow latency communication in edge enabled industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 1756–1767, Feb. 2023.
- [9] D. Xu, K. Yu, and J. A. Ritcey, "Cross-layer device authentication with quantum encryption for 5G enabled IIoT in industry 4.0," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 6368–6378, Sep. 2022.
- [10] K. Rabah, "Convergence of AI, IoT, big data and blockchain: A review," *lake Inst. J.*, vol. 1, no. 1, pp. 1–18, 2018.
- [11] Z. Guo, K. Yu, A. Jolfaei, F. Ding, and N. Zhang, "Fuz-spam: Label smoothing-based fuzzy detection of spammers in Internet of Things," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 11, pp. 4543–4554, Nov. 2022.
- [12] D. Peng, D. He, Y. Li, and Z. Wang, "Integrating terrestrial and satellite multibeam systems toward 6G: Techniques and challenges for interference mitigation," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 24–31, Feb. 2022.
- [13] Q. Zhang, K. Yu, Z. Guo, S. Garg, J. J. P. C. Rodrigues, M. M. Hassan, and M. Guizani, "Graph neural network-driven traffic forecasting for the connected Internet of Vehicles," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 5, pp. 3015–3027, Sep. 2022.
- [14] Y. Li, H. Ma, L. Wang, S. Mao, and G. Wang, "Optimized content caching and user association for edge computing in densely deployed heterogeneous networks," *IEEE Trans. Mobile Comput.*, vol. 21, no. 6, pp. 2130–2142, Jun. 2022.

- [15] Z. Shen, F. Ding, Y. Yao, A. Bhardwaj, Z. Guo, and K. Yu, "A privacy-preserving social computing framework for health management using federated learning," *IEEE Trans. Computat. Social Syst.*, early access, Dec. 22, 2022, doi: 10.1109/TCSS.2022.3222682.
- [16] A. Joshi, M. Vinay, and P. Bhaskar, "Impact of coronavirus pandemic on the Indian education sector: Perspectives of teachers on online teaching and assessments," *Interact. Technol. Smart Educ.*, vol. 18, no. 2, pp. 205–226, Sep. 2021.
- [17] D. Dessì, G. Fenu, M. Marras, and D. R. Recupero, "Bridging learning analytics and cognitive computing for big data classification in micro-learning video collections," *Comput. Hum. Behav.*, vol. 92, pp. 468–477, Mar. 2019.
- [18] M. Cantabella, R. Martínez-España, B. Ayuso, J. A. Yáñez, and A. Muñoz, "Analysis of Student behavior in learning management systems through a big data framework," *Future Gener. Comput. Syst.*, vol. 90, pp. 262–272, Jan. 2019.
- [19] S. Verma, S. S. Bhattacharyya, and S. Kumar, "An extension of the technology acceptance model in the big data analytics system implementation environment," *Inf. Process. Manage.*, vol. 54, no. 5, pp. 791–806, Sep. 2018.
- [20] A. Hesse, L. Glenna, C. Hinrichs, R. Chiles, and C. Sachs, "Qualitative research ethics in the big data era," *Amer. Behav. Scientist*, vol. 63, no. 5, pp. 560–583, May 2019.
- [21] A. Pashazadeh and N. J. Navimipour, "Big data handling mechanisms in the healthcare applications: A comprehensive and systematic literature review," *J. Biomed. Informat.*, vol. 82, pp. 47–62, Jun. 2018.
- [22] S. Li, F. Jiao, Y. Zhang, and X. Xu, "Problems and changes in digital libraries in the age of big data from the perspective of user services," *J. Academic Librarianship*, vol. 45, no. 1, pp. 22–30, Jan. 2019.
- [23] H. S. Munawar, S. Qayyum, F. Ullah, and S. Sepasgozar, "Big data and its applications in smart real estate and the disaster management life cycle: A systematic analysis," *Big Data Cognit. Comput.*, vol. 4, no. 2, p. 4, Mar. 2020.
- [24] L. T. Wright, R. Robin, M. Stone, and D. E. Aravopoulou, "Adoption of big data technology for innovation in B2B marketing," *J. Bus. Bus. Marketing*, vol. 26, nos. 3–4, pp. 281–293, Oct. 2019.
- [25] M. D. B. Castro and G. M. Tumibay, "A literature review: Efficacy of online learning courses for higher education institution using meta-analysis," *Educ. Inf. Technol.*, vol. 26, no. 2, pp. 1367–1385, Mar. 2021.
- [26] X.-W. Zhang and M. Zeng, "An empirical study on big data stock volatility forecasting algorithm based on multivariate hybrid criterion fuzzy model," *Int. J. Electr. Eng. Educ.*, vol. 13, no. 7, pp. 1–11, Sep. 2021.
- [27] Z. Jiang, T. Li, W. Min, Z. Qi, and Y. Rao, "Fuzzy c-means clustering based on weights and gene expression programming," *Pattern Recognit. Lett.*, vol. 90, pp. 1–7, Apr. 2017.



**TENGYUN GONG** was born in Fuzhou, Jiangxi, China, in 1978. She received the master's degree from East China University of Political Science and Law, Shanghai, China. Currently, she is with the Scientific Research Department, Huaiyin Normal University. Her research interests include education administration and sports law.



**JUNMIN WANG** was born in Ganyu District, Jiangsu, China, in 1975. He received the Ph.D. degree in Physical Education Training from Fujian Normal University, Fuzhou, China. Currently, he is with the Institute of Physical Education, Huaiyin Normal University. His research interest includes physical activity promotes health.

• • •