Intelligent Evaluation and Optimization of Postgraduate Education Comprehensive Ability Training under the Mode of "One Case, Three Systems"

Yong Xiang[1], Zeyou Chen[2], Liyu Lu[3], Yao Wei[4]*
School of Civil Engineering, Architecture and Environment, Xihua University, Chengdu, Sichuan, 610039, China[1, 2, 3]
School of Construction Engineering, Sichuan Technology and Business University, Chengdu, Sichuan, 611745, China[4]

***Abstract*—This study aims to explore the intelligent evaluation and optimization methods for the comprehensive ability training of graduate students under the mode of "one case, three systems" to improve the quality and effect of graduate training. Firstly, a weighted clustering algorithm for mixed attributes is designed. Secondly, an evaluation model of postgraduate training quality based on sampling method and ensemble learning is established. Finally, the algorithm's and the model's performance are compared and tested. The test results show that with the increase in the number of experiments, the accuracy of the proposed weighted clustering algorithm can reach more than 90%, which is improved by 10%. The average number of iterations is 276, and the accuracy and F1 value can achieve the highest level with fewer iterations and stable algorithm performance. Compared with the R1 model, F1 and the accuracy of the model proposed in this study are enhanced by 3.29% and 6.75%, respectively. The feature-weighted clustering algorithm and the training quality evaluation model designed here complement each other and jointly construct a more elaborate and comprehensive training system. The feature-weighted clustering algorithm oriented to mixed attributes for the first time combines sampling methods and ensemble learning in the education ability training. Moreover, a multi-dimensional and intelligent postgraduate training evaluation framework is constructed, which provides a new idea for improving the quality of postgraduate training.**

Keywords*—Comprehensive ability training of graduate students; one case; three systems; feature weighted clustering algorithm; sampling method; ensemble learning

I. INTRODUCTION

With society's continuous development and change, higher education has faced increasingly complex and diverse challenges in recent years. Improving the quality and comprehensive ability of postgraduate education has become one of the key objectives of the reform and development of higher education [1], [2]. However, the traditional postgraduate training mode has gradually revealed its limitations in meeting the growing social needs and personnel training requirements. Higher education in China is actively exploring and promoting the reform of the education system [3], [4].

In higher education, especially in postgraduate education, intelligent evaluation, and optimization research has gradually attracted widespread attention. Jiang (2021) utilized data analysis and mining technology to comprehensively evaluate the academic achievements, scientific research activities, and practical experience of graduate students to understand the characteristics and potential of students better and provide a basis for personalized training programs [5]. He (2020) discussed the critical issues of cultivating the comprehensive ability of graduate students. For example, how to reasonably integrate elements such as case teaching, curriculum system, mentor system, and practical system and establish effective connections and interactions between these elements to achieve the best effect of comprehensive ability training [6]. Liu (2023) focused on how to use big data analysis (BDA), machine learning (ML), and artificial intelligence (AI) technology to quantitatively evaluate the academic performance, innovation ability, and practical experience of graduate students [7].

The importance of this study cannot be ignored, as it deeply explores intelligent evaluation and optimization methods in the cultivation of postgraduate comprehensive abilities. In the current field of education, although research on intelligent evaluation and personalized learning is increasing, how to effectively integrate these concepts and technologies into postgraduate education models, especially in the application of the "one case, three systems" model, is still an area that has not been fully explored. In addition, the practical application of music education and virtual reality (VR) technology in the education system and how these technologies affect the quality and effectiveness of postgraduate ability development also require more systematic research and evaluation. Hence, this study aims to fill the gap in existing research by introducing intelligent technology and advanced algorithms, enriching understanding and comprehension of intelligent evaluation, personalized learning paths, music education integration, and VR technology application in higher education. Through the implementation of this study, a new postgraduate training evaluation framework based on data-driven and intelligent analysis can be constructed. Moreover, the effectiveness and advantages of intelligent evaluation and optimization methods in practical applications can be more clearly revealed through empirical research. As a result, the quality of postgraduate education has been improved while providing educators with more accurate training decision support. Additionally, through the in-depth analysis and empirical results provided by this study, it is possible to better understand the design principles of personalized learning pathways and how music education and VR technology can be creatively integrated into modern education systems. It can offer new theoretical and practical perspectives for future research in related fields. Compared to previous studies, this study adopts an advanced weighted clustering algorithm for mixed attributes and the evaluation model based on sampling method and ensemble learning, a significant extension of existing educational evaluation methods. The limitation of this study is that the effectiveness of intelligent evaluation and optimization methods may be limited by factors such as data quality, algorithm adaptability, and diversity of teaching environments. Future research must validate the universality and stability of the proposed model in a broader educational context, thus enhancing comparative fairness and the universality of conclusions. Through continuous iteration and optimization, this study aims to improve the intelligent evaluation and optimization framework for cultivating graduate comprehensive abilities, making it an indispensable part of a high-quality postgraduate training system.

In short, with the rapid development of information technology, many colleges and universities began to explore the use of AI, data mining (DM), ML, and other technical means to achieve the evaluation of graduate students' comprehensive ability and the optimization of training programs. Section I summarizes the research background, current situation, and purpose. Section II proposes a weighted clustering algorithm for mixed attributes to solve the attribute evaluation problem. Then, an evaluation model of postgraduate training quality based on sampling method and ensemble learning is established. Section III tests the performance of the algorithm and model proposed here, describes the test data, environment, and results, and compares and discusses the proposed method with the traditional method. Section IV looks forward to the research contribution and future direction, explaining this study's theoretical and practical value and development space. Section V concludes the study.

II. Literature Review

In the research field of "One Case, Three Systems," under the intelligent evaluation and optimization of postgraduate education comprehensive ability training mode, the work of multiple scholars provides key background and support for the current research. Li et al. (2020) introduced the entropy weight method and grey clustering analysis, providing a novel method for evaluating the quality of online teaching [8]. This reflected the scholars' understanding of using multi-dimensional evaluation when facing the problems of online teaching quality. Based on the proposed evaluation model, scholars also proposed a series of strategies to improve the quality of online teaching, thus enriching the practical application of the research. However, this study had limitations, such as the representativeness and feasibility of the datasets used in empirical analysis. Future research can further validate and expand the conclusions of this study through broader data collection and deeper empirical research. Lee et al. (2021) focused on a holistic music education approach for children jointly developed by music therapists and experts, combining technology and music, integrating local culture, and constructing a holistic education framework [9]. The research results indicated that implementing a holistic music education approach can significantly enhance children's abilities with developmental delays, while supportive training has a positive effect. In addition, decision trees explored and developed an intelligent evaluation model with high learning effectiveness. The sensitivity of this model within the sample reached 90.6%, while the comprehensive indicator F was 79.9%, which had a high reference value. In the future, educators can use DM to assist decision-making systems as evaluation tools to evaluate children participating in education in the early and middle stages of the curriculum. It can also predict their continuous implementation and learning effectiveness, help decide whether to continue investing and adjusting the curriculum, and use educational resources more effectively. Yang et al. (2022) investigated the evolution from Assessment of Learning (AoL) to Assessment as Learning (AaL) from four aspects: participants, testing format, process-based multivariate data, and measurement models for multivariate data. They proposed suggestions for interdisciplinary collaboration, integrating education, psychology, and information technology into the theories and methods of educational measurement [10]. In addition, the study emphasized that measurements' effectiveness, ethics, and fairness should also be considered vital issues. Researchers and practitioners in educational measurement must adhere to the pursuit of the substantive significance of measurement and provide unique experience and guidance for the theoretical and practical development of educational evaluation in this tremendous revolution. Wang et al. (2022) constructed a personalized learning model based on distributed computing methods of the Internet of Things (IoT) and clustering algorithms of deep learning (DL). The research results showed that the accuracy of this model on personalized learning platforms based on the IoT and DL algorithms reached 85% [11]. Compared with the latest research models, this model performed better in score prediction and customized recommendations. This model had significant practical value in promoting the development of the IoT and DL in professional learning. This exploration innovatively divided the understanding level of learners at a hierarchical level. It provided personalized learning resources aligned with their cognitive abilities to enhance their knowledge level and achieve personalized learning. These studies focused on improving algorithm performance, offering relevant research for the performance improvement of weighted clustering algorithms in this study. Lee and Hwang (2022) highlighted sustainable education from the perspective of VR technology application, providing innovative insights into the educational research field [12]. By delving into the experience of pre-service teachers in VR content design, they highlighted the importance of this transformative experience in enhancing teachers' technical readiness, digital citizenship, and perceived educational benefits. The study emphasized the importance of introducing emerging technologies into education to promote sustainable education development.

Although the research of the above scholars provided valuable insights into the current fields of intelligent evaluation, personalized learning, music education, and the application of VR technology in education, there are certain limitations to each research. Firstly, many studies were influenced by specific backgrounds, sample sizes, and dataset limitations, which limited their universality and generalization ability. Secondly, the empirical analysis of some studies lacked sufficient breadth and depth, failing to validate the applicability of the proposed model. Besides, some studies still lacked practical application cases in real-world environments, thus limiting their guiding role in practice. However, examining previous research provided this study with a profound understanding of different aspects of the field of education. Still, these findings must be validated in a broader and more diverse context. In addition, there was a relative lack of comprehensive research on cultivating graduate students' extensive abilities under the "One Case, Three Systems" model. Previous research aimed to fill some gaps in previous studies and expand understanding of the practical applications of personalized learning, intelligent evaluation, music education, and VR technology in education. By delving deeper into the limitations of previous research, there was an urgent need for broader, deeper, and more empirical research to enhance the practicality of existing theoretical frameworks and promote their practical application. This study’s importance was applying the feature-weighted clustering algorithm and training quality evaluation model to build a more comprehensive framework for postgraduate training evaluation. Compared with the previous research, this study emphasized the total consideration of multi-dimensional characteristics while improving the evaluation model’s algorithm performance and accuracy.

III. METHOD

A. Feature-weighted Clustering Algorithm for Mixed Attributes

The feature-weighted clustering algorithm for mixed attributes proposed in this study is a data clustering technology that deals with numerical and discrete (mixed attributes) data [13,14]. The flow of the clustering algorithm is displayed in Fig. 1:

In Fig. 1, the proposed algorithm’s core idea is to give appropriate weights to different attributes in the clustering process to accurately reflect the importance of attributes when calculating the distance or similarity between data points [15]. The specific process is as follows:

Step 1: Improvement of dissimilarity of classification attributes;

Sample distribution information is described by covariance matrix, and the Mahala Nobis distance D_m^2 between sample x_i and cluster center a_j is defined by Eq. (1):

$$D_m^2 = (x_i - a_j)\Sigma^{-1}(x_i - a_j)^T \tag{1}$$

Σ represents covariance matrix, which Eq. (2) can calculate:

$$\Sigma = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T - a_j a_j^T \tag{2}$$

On this basis, the proportional coefficient P is introduced to increase the statistical probability of classified samples and unclassified samples in clustering, as illustrated in Eq. (3):

$$P = diag\left(\sqrt{\gamma_1}, \sqrt{\gamma_2}, \cdots, \sqrt{\gamma_m}\right), \gamma = 1 - |C_l| \,/\, \left|C_{lij}\right| \tag{3}$$

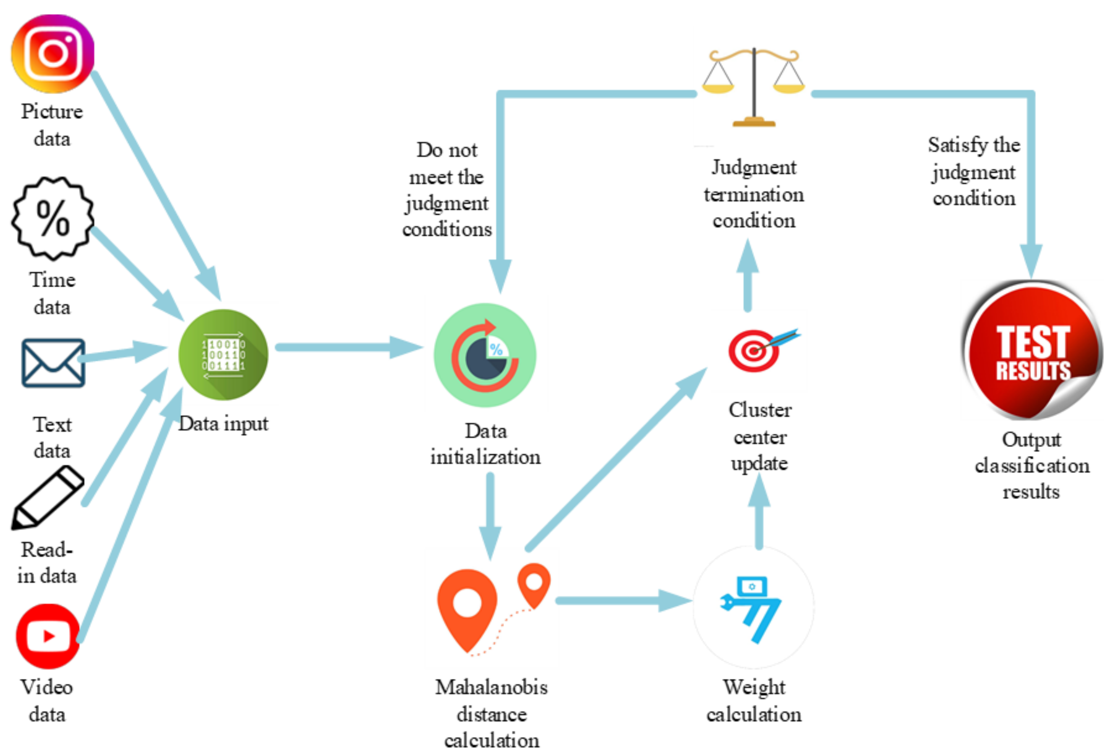

Fig. 1. Flow chart of feature-weighted clustering algorithm for mixed attributes.

$|C_l|$ refers to the number of classified samples, and $|C_{lij}|$ denotes the frequency of the j-th attribute of unclassified sample x_i in C_l. Redefine Mahalanobis distance after orthogonal decomposition of positive definite matrix Σ, as listed in Eq. (4):

$$D_m^2 = \left[PQ\Lambda^{-\frac{1}{2}}Q^T(x_i - a_j)\right]^T \left[PQ\Lambda^{-\frac{1}{2}}Q^T(x_i - a_j)\right] \quad (4)$$

Q is the inverse of the covariance matrix, the off-diagonal element of Λ is 0, and the diagonal element is the eigenvalue of the covariance matrix.

Step 2: Calculation of the weights of the numerical attributes;

The K nearest neighbor sample of x_i reads:

$$KN(x_i) = \{x_j | d(x_i, x_j) \le d(x_i, Near(x_i))\} \quad (5)$$

$Near(x_i)$ represents the nearest K-th sample point to the Mahala Nobis of x_i. The definition of intra-cluster dispersion d_1 is as follows:

$$d_1 = \sum_{j=1}^{k} \frac{|x_{ij} - KN(x_{ij})|}{max(A_l) - min(A_l)} \quad (6)$$

A_l represents the cluster's feature set of sample points. The definition of cluster dispersion is expressed by Eq. (7):

$$d_2 = \sum_{C} \frac{p(C)}{1 - p(x_i)} \sum_{j=1}^{k} \frac{|x_{ij} - KN(x_{kj})|}{max(A_l) - min(A_l)} \quad (7)$$

$C \ne class(x_i)$. The update of the weights of samples and X on the feature set A_l is Eq. (8):

$$w_l^n = w_l^n - \frac{d_1}{mk} + \frac{d_2}{mk} \quad (8)$$

The weights of numerical attributes are mainly determined by d_1 and d_2. The smaller the intra-cluster dispersion is, the greater the inter-cluster dispersion is, and the higher the similarity of samples in changing attributes. On the contrary, the lower the discrimination of attributes for clustering, the greater the weight.

Step 3: Calculation of the weight of the classification attribute [16,17];

$D(A_l) = \{a_l | a_l = x_{il}, 1 \le i \le n, P \le l \le m\}$ is defined as the set consisting of the first attribute of the sample set X, and its distribution function is as follows:

$$\begin{cases} p(a) = \frac{|D(A_{l=v})|}{|D(A_l)|} \\ p(r) = \frac{|D(R_{l=t})|}{|D(R)|} \\ p(a,r) = \frac{|D(A_{l=v})|}{|D(R)|} \end{cases} \quad (9)$$

$p(a)$ and $p(r)$ are the marginal probability distribution function of classification attribute A and clustering result R; $p(a,r)$ is the joint probability distribution function of A and R. The mutual information M between A and R can be written as Eq. (10):

$$M(A_l, R) = \sum_{r \in R} \sum_{a \in A_l} p(a,r) \log 2 \left(\frac{p(a,\ r)}{p(a) * p(r)}\right) \quad (10)$$

The weight of the classification attribute is calculated as Eq. (11):

$$w_l^C = \frac{M(A_l, R)}{\sum_{l=p+1}^{m} M(A_l, R)} \quad (11)$$

Based on these calculations, the revised objective function $E(U, C)$ is obtained, as indicated in Eq. (12):

$$E(U,C) = \sum_{j=1}^{k} \sum_{i=1}^{n} u_{ij}{}^a (w_l^n d_1(x_i - a_j) + \gamma w_l^C d_2(x_i - a_j)) \quad (12)$$

Step 4: Update of the cluster center;

The updating equation of the l-th numerical feature c_{jl} of the cluster center c_j is:

$$c_{jl} = \frac{\sum_{i=1}^{n} (u_{ij})^a w_l^n x_{il}}{\sum_{i=1}^{n} (u_{ij})^a} \quad (13)$$

The equation for updating the l-th classification feature c_{il} of the cluster center c_i reads:

$$c_{il} = a_l^s \in D(A_l) \quad (14)$$

s satisfies the condition shown in Eq. (15):

$$\sum_{i=1}^{n} (u_{ij}{}^a | x_{ij} = a_l^s) \ge \sum_{i=1}^{n} (u_{ij}{}^a | x_{ij} = a_l^t), s \ge 1, t \le |D(A_l)| \quad (15)$$

$|D(A_l)|$ is the number of values of the classification attribute A_l.

Step 5: Initialization of parameters, cluster center matrix, and iteration times;

All variables start from 0, and the maximum number of iterations is set to t.

Step 6: The weighted clustering is completed according to the process shown in Fig. 1.

B. *Evaluation Model of Postgraduate Training Quality based on Sampling Method and Ensemble Learning*

The evaluation model of postgraduate training quality based on sampling method and ensemble learning is an analytical tool for evaluating the process and effect of postgraduate training [18]. The specific structure is presented in Fig. 2.

Fig. 2 signifies that postgraduate training involves data of multiple dimensions, such as academic achievements, scientific research activities, and practical experience. To reduce computational complexity and improve efficiency, the model extracts feature from different training dimensions and gives appropriate weights to different features in the selected samples, accurately reflecting their importance in training quality evaluation.

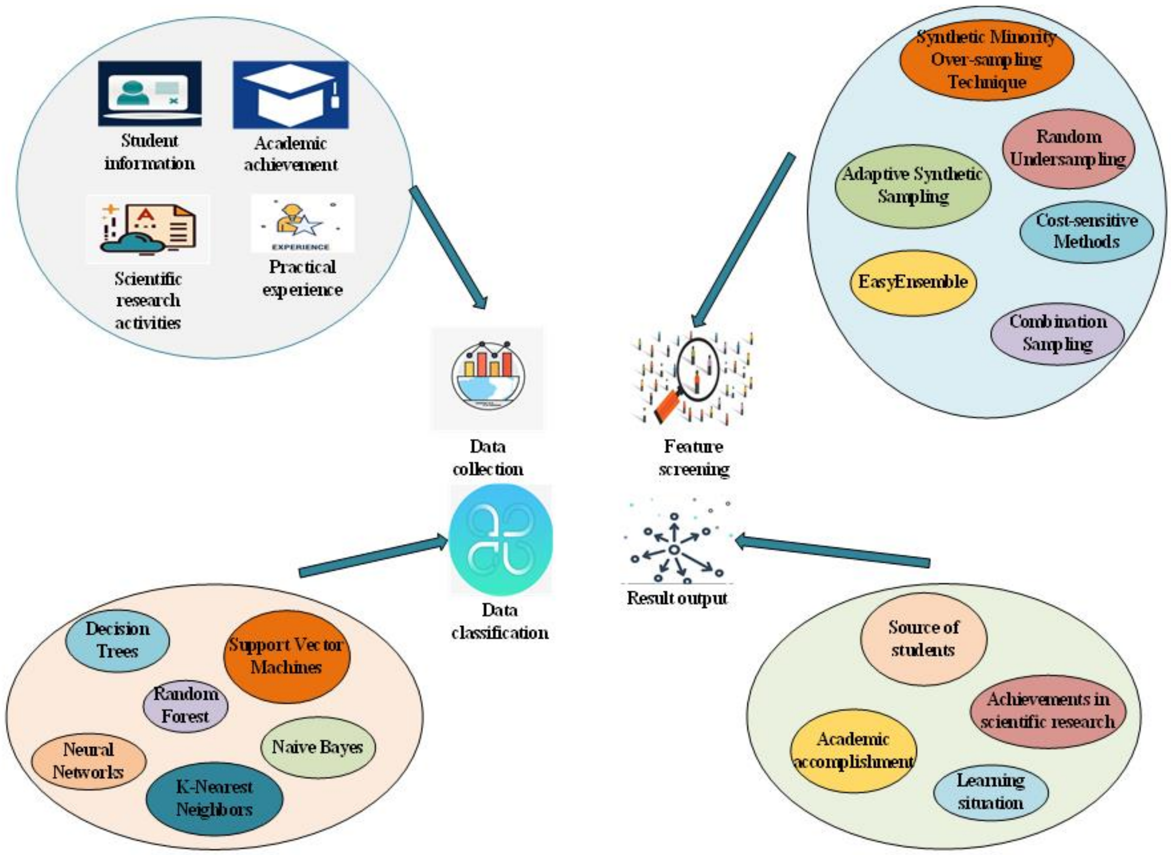

Fig. 2. Model structure of postgraduate training quality evaluation based on sampling method and ensemble learning.

IV. Results and Discussion

A. Data Collection and Experimental Environment

This study employs the Student Performance Data Set data from the University of California, Irvine's UCI ML Repository database as the algorithm performance test's test set and data source. The UCI data set is an open-source dataset proposed by the University of California, Irvine, which is suitable for pattern recognition and ML direction. Many scholars choose to use the data set on UCI to verify the correctness of their proposed algorithms. These data sets are divided into binary classification, multi-classification, and regression fitting problems. The UCI data set provides the main attributes of each data set. It can be used to demonstrate the rationality of various algorithms proposed by oneself through experimental results on its data set. Data address: https://archive.ics.uci.edu/datasets. The Student Performance Data Set aims to predict the performance of students in two secondary schools in Portugal, covering two different subjects: Mathematics (mat) and Portuguese (por). The data collection methods include school reports and questionnaire surveys, encompassing multiple attributes such as student performance, demographics, and social and school-related characteristics [19]. The entire data set contains 649 instances, mainly used for classification and regression tasks. Two independent data sets correspond to two subjects: mat and por. In both data sets, the target attribute G3 refers to the final year grade, while G1 and G2 represent the first and second-semester grades, respectively. It is important to note that there is a strong correlation among G3, G2, and G1. This is because G3 is the final grade (released in Semester 3), while G1 and G2 correspond to Semester 1 and 2 grades. Although it is more challenging to predict G3 without G2 and G1, this prediction is more useful in practical applications. The attributes of this data set include school, gender, age, place of residence, family size, parental cohabitation status, parental education levels, parental occupation, the reason for choosing a school, guardians, school time, study time, number of past classroom failures, additional educational support, family educational support, extra paid courses, and extracurricular activities. Additionally, it also involves whether to attend daycare, whether to plan for higher education, whether there is a romantic relationship, family network access, quality of family relationships, free time after school, frequency of outings with friends, alcohol consumption on workdays and weekends, current health status, and school absences. Data address: https://archive.ics.uci.edu/dataset/320/student+performance.
This study tests the evaluation model of postgraduate training quality based on sampling method and ensemble learning. The ratio of the training set to the data set is 4:1, and the number of training set samples is 622, divided into three levels: high, medium, and low, with the numbers 68, 480, and 74, respectively. The test set's data volume is 156, and the number of high, medium, and low samples is 17, 121, and 18, respectively, which meets the proportional requirements.

The experimental equipment environment of this study is Dell notebook equipment; the CPU model is Intel (R) Core

(TM) i5-7300HQ CPU @ 2.50GHz; the memory is 8G; the graphics card model is GTX1050; the operating system is Windows 11 system, and the software environment is C++ language environment to ensure the stability of the algorithm and model in the process of training and testing.

B. Indicator Setting

1) Feature-weighted clustering algorithm for mixed attributes: This study uses the accuracy F-Measure to evaluate the algorithm's performance. F-Measure is the weighted harmonic average of accuracy and recall. Accuracy indicates the proportion that objects in the same cluster are divided into the same category, and recall reflects the proportion that objects in the same category are assigned to the same cluster. The fuzzy factor is set to 2 in the experiment, the clustering threshold is 0.0001, and the maximum number of iterations T is set to 10000.

When the algorithm is used to verify the comprehensive ability training of postgraduate education, it is verified by the evaluation results of the postgraduate training information data of S University from 2019 to 2022 according to the indicators exhibited in Table I:

Table I describes that 11 secondary refinement indicators are set under the classification of four primary indicators: student source, tutor title, scientific research achievements, and learning situation, and weighted cluster analysis is carried out.

a.

2) Evaluation model of postgraduate training quality based on sampling method and ensemble learning: Aiming at the designed evaluation model, this study evaluates it through three standards: accuracy, recall, and F1 value. F1 value, as a calculation index of the reconciliation of accuracy and recall, can fully reflect the classification ability of the model and is the most representative.

C. Experimental Result

To verify the accuracy of the proposed feature-weighted clustering algorithm, this algorithm is compared with the fuzzy clustering, the goal-Okumoto-kapur with k-prototypes, synthetic minority over-sampling technique, the Adaptive Synthetic Sampling, and the improved genetic fuzzy K-Prototypes algorithms. They are named A1, B2, C3, D4, E5, F6 in turn. Fig. 3 depicts various algorithms' clustering accuracy and clustering effect on data sets.

In Fig. 3(a), as the number of experiments increases, the proposed weighted clustering algorithm can achieve over 90%, which is 10% higher. Fig. 3(b) denotes that the average number of iterations of the algorithm proposed here is 276, and the accuracy and F value can reach the highest level with fewer iterations and stable algorithm performance. Fig. 4 demonstrates the proposed algorithm's weighted clustering results for postgraduate training quality and various algorithms' contour change curves with the change of fuzzy factors.

TABLE I. QUALITY TRAINING INDEX OF POSTGRADUATE EDUCATION

First Index	**Source of students**			
Second index	Undergraduate colleges A#	Learning mode B#	Entrance examination marks C*	
First Index	**Academic literacy**			
Second index	Tutor title D#		Tutor academic achievements E*	
First Index	**Achievements in scientific research**			
Second index	Number of papers F*	Paper quality #G*	Number of patents H*	Number of awards I
First Index	**Learning conditions**			
Second index	Average scores J*		Type of scholarships K#	

Note: * indicator is a numerical indicator, # indicator is a classified indicator.

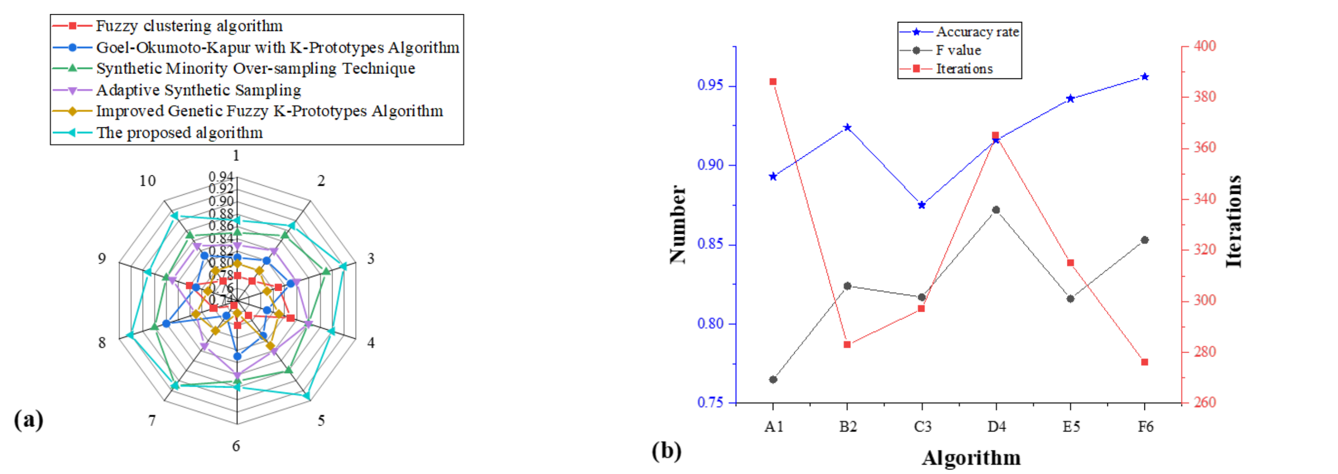

Fig. 3. Clustering accuracy and clustering effect of different algorithms on data sets ((a) Clustering accuracy; (b) Clustering effect).

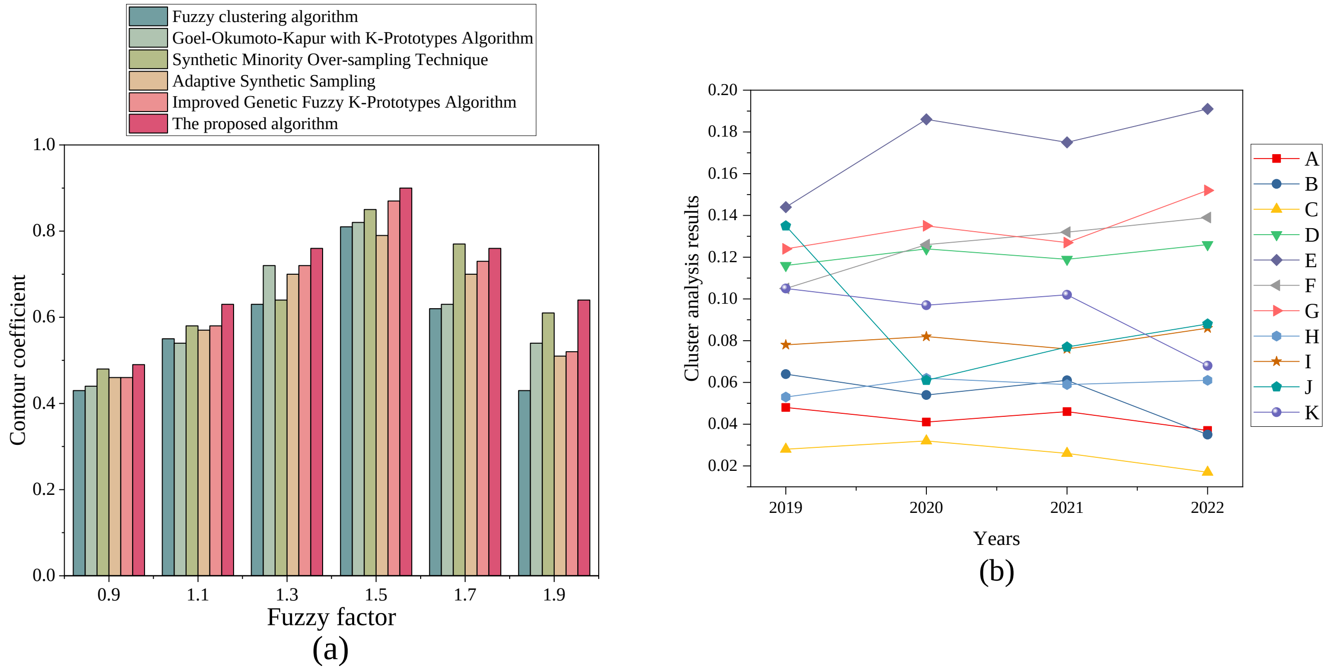

Fig. 4. Evaluation results of postgraduate training quality (a) Changes of contour coefficients of various algorithms; (b) Weighting cluster analysis results for evaluation elements)

In Fig. 4(a), with the fuzzy factor gradually increasing from 0.8 to 2.0, the contour coefficient of the proposed algorithm is higher than other algorithms, and the average value is around 0.2. It illustrates that the weighted clustering algorithm designed in this study can guarantee the data features to the greatest extent and has great advantages for feature extraction. In Fig. 4(b), the weight of evaluation indicators has not changed significantly in recent years. The academic achievements of tutors and the quality of papers published by students have become the most influential factors in the cultivation of postgraduate education ability, accounting for a relatively high proportion, with a comprehensive rate of over 46%. It can be found that the guidance of mentors to students and the students' research results are the core content of current postgraduate education ability cultivation. Fig. 5 presents the comparison of diverse classification and sampling algorithms' evaluation performance:

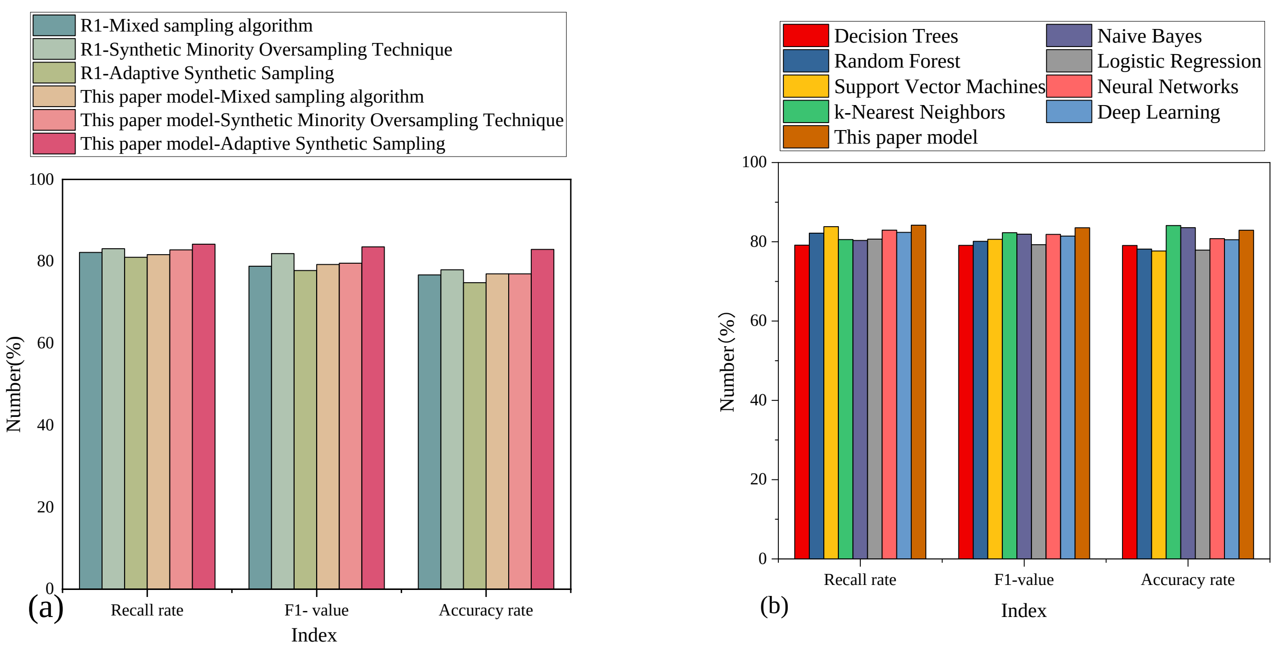

Fig. 5. Comparison of evaluation performance of different sampling and classification algorithms ((a) Comparison of sampling algorithms; (b) Performance comparison of classification algorithms).

In Fig. 5, this study sets the sampling comparison algorithm as a mixed sampling algorithm, synthetic minority oversampling technique algorithm, and Adaptive Synthetic Sampling algorithm. It constructs the R1 model based on a decision tree and random forest as base classifiers. As a meta-classifier, the Gradient Boosting Decision Tree is a quality evaluation model of postgraduate education based on sampling method and ensemble learning. Then, the model's performance is tested. In Fig. 5(a), the F1 value of the model proposed in this study is 83.54%. The accuracy is 82.91%, and its practical application performance is good. Compared with the R1 model, the proposed model’s F1 value and accuracy increase by 3.29% and 6.75%, respectively. In Fig. 5(b), compared with other algorithms, the proposed model has a recall of 84.18%, an F1 value of 83.54%, and an accuracy of 82.91%. This model has higher accuracy and stability and can effectively solve the problems of subjectivity and data imbalance in the evaluation process.

The performance evaluation results of the feature-weighted clustering algorithm on different data sets are outlined in Table II. The EduNLP Student Performance Data Set in Table II analyzes students' literal expressions and evaluates their academic performance and needs through natural language processing technology. Student Performance in Online Learning Data Set is employed to evaluate students' performance in online learning, including learning time, interaction, and other characteristics. Student Academic Performance Data Set (Kaggle) is used to analyze students' academic performance, involving grades, subject preferences and other characteristics. The results reveal that among all data sets, the feature-weighted clustering algorithm performs best on the UCI data set, at 90%, while the performance is relatively low on the Online Learning Data Set, at 81%. On different data sets, the average number of iterations of the algorithm fluctuates in a relatively reasonable range, and the average number of iterations on UCI data set is the lowest, at 276. The algorithm on the UCI data set performs the best in terms of contour coefficients, at 0.2, while the Online Learning Data Set is at 0.16. In all data sets, feature-weighted clustering algorithms use evaluation indicators such as F1 value, accuracy, and recall, providing researchers with a comprehensive performance evaluation. These evaluation results provide a reference for applying feature-weighted clustering algorithms in different backgrounds, and the performance on the UCI data set offers potential best practices for other datasets. The performance on the EduNLP Student Performance Data Set is second, while the performance on Online Learning Data Set is relatively poor, requiring more in-depth research and improvement. The Student Academic Performance dataset performed well, approaching the performance level of the UCI data set.

The performance evaluation results of various intelligent evaluation methods on UCI data sets are suggested in Table III. The results of Table III indicate that the performance of the proposed weighted clustering algorithm for mixed attributes is the best, with a clustering accuracy of 90%, an average number of iterations of 276, and a contour coefficient of 0.2. In contrast, although the STEAM Education+Smart Greenhouse VR [20] algorithm performs relatively stable in clustering accuracy, its overall performance is poor, with a clustering accuracy of 80%, an average number of iterations of 300, and a contour coefficient of 0.15. The Analytic Hierarchy Process+Delphi method [21] achieves good clustering accuracy, reaching 85%, but slightly inferior to the proposed algorithm in other indicators. The neural fuzzy method with multiple inputs and outputs [22] has relatively low clustering accuracy and F1 value, at 78% and 73.25%, respectively. However, its average number of iterations is 350, which may have some stability issues. The binary genetic algorithm [23] and the educational data classification method of swarm intelligence [24] obtain 82% and 87% clustering accuracy, respectively. Furthermore, their performances are relatively balanced, but slightly inferior to the proposed algorithm in other indicators. In general, the proposed algorithm has excellent clustering accuracy, average number of iterations, and F1 value, proving its effectiveness in students' comprehensive ability evaluation.

TABLE II. PERFORMANCE EVALUATION RESULTS OF FEATURE WEIGHTED CLUSTERING ALGORITHMS ON DIFFERENT DATA SETS

Data Set	Cluster accuracy (%) (Mean)	Average number of iterations	Contour coefficient (Mean)	Evaluation indicators	F1 value (%)	Accuracy (%)	Recall rate (%)
Student Performance Data Set (UCI)	90	276	0.2	Feature-weighted clustering	83.54	82.91	84.18
EduNLP Student Performance Data Set	88	290	0.18	Feature-weighted clustering	82.21	83.12	81.49
Student Performance in Online Learning Data Set (Kaggle)	81	360	0.16	Feature-weighted clustering	75.12	76.05	74.76
Student Academic Performance Data Set (Kaggle)	89	280	0.22	Feature-weighted clustering	83.02	84.11	82.58

TABLE III. PERFORMANCE EVALUATION RESULTS OF DIFFERENT INTELLIGENT EVALUATION METHODS ON THE STUDENT PERFORMANCE DATA SET (UCI)

Algorithm	Cluster accuracy (%) (Mean)	Average number of iterations	Contour coefficient (Mean)	Evaluation indicators	F1 value (%)	Accuracy (%)
The proposed algorithm	90	276	0.2	83.54	82.91	84.18
STEAM Education+Smart Greenhouse VR	80	300	0.15	76.21	78.34	75.89
The Analytic Hierarchy Process+Delphi method	85	320	0.18	80.56	81.42	79.73
The neural fuzzy method with multiple inputs and outputs	78	350	0.13	73.25	75.11	72.46
The binary genetic algorithm	82	280	0.22	78.92	79.87	78.36
The educational data classification method of swarm intelligence	87	300	0.24	81.67	82.53	81.12

D. Discussion

Zhu (2023) paid attention to using intelligent methods to design personalized learning paths and course recommendation systems to improve graduate students' comprehensive ability training effect according to their interests, abilities, and backgrounds [25]. Thurzo (2023) used BDA, ML, and DM technology so researchers could analyze students' learning behavior, performance, and progress and dig out meaningful information for cultivating graduate students' comprehensive ability to evaluate and optimize their intelligence [26]. Shan (2023) focused on building a complete comprehensive ability evaluation system, including academic ability, innovation ability, practical ability, and other indicators and standards, to conduct intelligent evaluation and optimization more effectively [27]. The algorithm and model proposed in this study can achieve an accuracy of over 90% with an increase in the number of experiments, an improvement of 10%, and an average iteration of 276 times. Therefore, the accuracy and F-value can reach the highest level. With the fuzzy factor increasing from 0.8 to 2.0, the proposed algorithm's contour coefficient is larger than other algorithms, and the average value is about 0.2. The proposed model's F1 value is 83.54%, and the accuracy is 82.91%. Compared with the R1 model, the proposed model's F1 value and accuracy are improved by 3.29% and 6.75%, and the recall and F1 values are 84.18% and 83.91%. Compared with the traditional methods, the feature-weighted clustering algorithm for mixed attributes can better consider the relationship between different attributes and improve the accuracy and interpretation of clustering. The evaluation model of postgraduate training quality based on sampling method and ensemble learning can make personalized evaluation and training suggestions according to the situation of individual students to better meet their needs and development direction.

V. CONCLUSION

The proposed weighted clustering algorithm based on mixed attributes can better handle multi-attribute data and accurately evaluate graduate students' comprehensive ability. The evaluation model of graduate students' training quality based on sampling method and ensemble learning provides new ideas and tools, and personalized training suggestions can better meet the needs of graduate students. Applying the algorithm and model to emergency management helps improve the quality and effect of graduate students' training in this field. However, the proposed methods and models still need adaptive adjustment in other fields, and their versatility is limited. In future research, people can consider introducing multiple data sources, such as text, image, and voice, to carry out multimodal data fusion to evaluate graduate students' comprehensive ability. In addition, applying the proposed methods and models in other professional fields has been explored to verify their applicability and effectiveness in different fields.

ACKNOWLEDGEMENT

This work was sponsored in part by the Graduate School of Xihua University: Research on Interdisciplinary Education Model and Innovation Reform for graduate students of Emergency major based on One case and three systems (No: YJG2020008); Sichuan Department of Education: Research on the Cultivation and realization Path of National Identity Concept in College Students' Innovation and Entrepreneurship Education under the guidance of Party Building (No: JG2021-938).

COMPETING OF INTEREST

The authors declare no competing of interests.

AUTHORSHIP CONTRIBUTION STATEMENT

Yao Wei: Conceptualization, Investigation, Methodology, Writing, Project administration.

Yong Xiang: Formal analysis, Language review

Zeyou Chen: Software, Methodology

Liyu Lu: Validation

REFERENCES

[1] Y. Cao, J. Shan, Z. Gong, J. Kuang, and Y. Gao, "Status and challenges of public health emergency management in China related to COVID-19," Front Public Health, vol. 8, p. 250, 2020. https://doi.org/10.3389/fpubh.2020.00250.

[2] S. Wang, L. Jiang, J. Meng, Y. Xie, and H. Ding, "Training for smart manufacturing using a mobile robot-based production line," Frontiers of Mechanical Engineering, vol. 16, pp. 249–270, 2021. https://doi.org/10.1007/s11465-020-0625-z.

[3] A. Bhutoria, "Personalized education and artificial intelligence in the United States, China, and India: A systematic review using a human-in-the-loop model," Computers and Education: Artificial Intelligence, vol. 3, p. 100068, 2022. https://doi.org/10.1016/j.caeai.2022.100068.

[4] M. Liu and R. Su, "Research on the Path of Digital Transformation of Postgraduate Education in Chinese Universities under the Background of Digital Education Strategy," Intell Inf Manag, vol. 15, no. 5, pp. 339–349, 2023. https://doi.org/10.4236/iim.2023.155016.

[5] Y. Jiang and B. Li, "Exploration on the teaching reform measure for machine learning course system of artificial intelligence specialty," Sci Program, vol. 2021, pp. 1–9, 2021. https://doi.org/10.1155/2021/8971588.

[6] H. He, H. Yan, and W. Liu, "Intelligent teaching ability of contemporary college talents based on BP neural network and fuzzy mathematical model," Journal of Intelligent & Fuzzy Systems, vol. 39, no. 4, pp. 4913–4923, 2020. https://doi.org/10.3233/JIFS-179977.

[7] C. Liu, X. Li, and J. Zhang, "Construction of Students' Innovation Ability Portrait for Cultural Intelligence Computing," International Journal of Crowd Science, vol. 7, no. 2, pp. 77–86, 2023. https://doi.org/10.26599/IJCS.2023.9100001.

[8] M. Li, and Y. Su, "Evaluation of online teaching quality of basic education based on artificial intelligence," International Journal of Emerging Technologies in Learning (iJET), vol. 15, no. 16, pp. 147-161, 2020. https://www.learntechlib.org/p/217942/.

[9] L. Lee, and Y.-S. Liu, "Training effects and intelligent evaluated pattern of the holistic music educational approach for children with developmental delay," International Journal of Environmental Research and Public Health, vol. 18, no. 19, pp. 10064, 2021. https://doi.org/10.3390/ijerph181910064.

[10] L. P. Yang, and T. Xin, "Changing Educational Assessments in the Post - COVID - 19 Era: From Assessment of Learning (AoL) to Assessment as Learning (AaL)," Educational Measurement: Issues and Practice, vol. 41, no. 1, pp. 54-60, 2022. https://doi.org/10.1111/emip.12492.

[11] M. Wang, and Z. Lv, "Construction of personalized learning and knowledge system of chemistry specialty via the internet of things and clustering algorithm," The Journal of Supercomputing, vol. 78, no. 8, pp. 10997-11014, 2022. https://doi.org/10.1007/s11227-022-04315-8.

[12] H. Lee, and Y. Hwang, "Technology-enhanced education through VR-making and metaverse-linking to foster teacher readiness and sustainable learning," Sustainability, vol. 14, no. 8, pp. 4786, 2022. https://doi.org/10.3390/su14084786.

[13] A. R. Haniah, A. Aman, and R. Setiawan, "Integration of strengthening of character education and higher order thinking skills in history learning," Journal of Education and Learning (EduLearn), vol. 14, no. 2, pp. 183-190, 2020. https://doi.org/10.11591/edulearn.v14i2.15010.

[14] D. Denny and I. Iskandar, "The mastery of teacher emotional intelligence facing 21st century learning," International Journal of Education and Teaching Zone, vol. 1, no. 1, pp. 50–59, 2022. https://karya.brin.go.id/id/eprint/13167.

[15] Y. Baashar et al., "Evaluation of postgraduate academic performance using artificial intelligence models," Alexandria Engineering Journal, vol. 61, no. 12, pp. 9867–9878, 2022. https://doi.org/10.1016/j.aej.2022.03.021.

[16] S. Grabowska and S. Saniuk, "Assessment of the competitiveness and effectiveness of an open business model in the industry 4.0 environment," Journal of Open Innovation: Technology, Market, and Complexity, vol. 8, no. 1, p. 57, 2022. https://doi.org/10.3390/joitmc8010057.

[17] W. Wei and Y. Jin, "A novel Internet of Things-supported intelligent education management system implemented via collaboration of knowledge and data," Mathematical Biosciences and Engineering, vol. 20, no. 7, pp. 13457–13473, 2023. https://doi.org/10.3934/mbe.2023600.

[18] C. Malamateniou et al., "Artificial intelligence: guidance for clinical imaging and therapeutic radiography professionals, a summary by the Society of Radiographers AI working group," Radiography, vol. 27, no. 4, pp. 1192–1202, 2021. https://doi.org/10.1016/j.radi.2021.07.028.

[19] L. H. Son, and H. Fujita, "Neural-fuzzy with representative sets for prediction of student performance," Applied Intelligence, vol. 49, no. 1, pp. 172-187, 2019. https://doi.org/10.1007/s10489-018-1262-7.

[20] C.-Y. Huang, B.-Y. Cheng, S.-J. Lou, and C.-C. Chung, "Design and Effectiveness Evaluation of a Smart Greenhouse Virtual Reality Curriculum Based on STEAM Education," Sustainability, vol. 15, no. 10, pp. 7928, 2023. https://doi.org/10.3390/su15107928.

[21] H.-C. Tsai, A.-S. Lee, H.-N. Lee, C.-N. Chen, and Y.-C. Liu, "An application of the fuzzy Delphi method and fuzzy AHP on the discussion of training indicators for the regional competition, Taiwan national skills competition, in the trade of joinery," Sustainability, vol. 12, no. 10, pp. 4290, 2020. https://doi.org/10.3390/su12104290.

[22] L. H. Son, and H. Fujita, "Neural-fuzzy with representative sets for prediction of student performance," Applied Intelligence, vol. 49, no. 1, pp. 172-187, 2019. https://doi.org/10.1007/s10489-018-1262-7.

[23] S. S. Shreem, H. Turabieh, S. Al Azwari, and F. Baothman, "Enhanced binary genetic algorithm as a feature selection to predict student performance," Soft Computing, vol. 26, no. 4, pp. 1811-1823, 2022. https://doi.org/10.1007/s00500-021-06424-7.

[24] A. A. Yahya, "Swarm intelligence-based approach for educational data classification," Journal of King Saud University-Computer and Information Sciences, vol. 31, no. 1, pp. 35-51, 2019. https://doi.org/10.1016/j.jksuci.2017.08.002.

[25] Z. Zhu and L. Zhang, "Artificial Intelligence Empowers Postgraduate Education Ecologically Sustainable Development Model Construction," Sustainability, vol. 15, no. 7, p. 6157, 2023. https://doi.org/10.3390/su15076157.

[26] A. Thurzo, M. Strunga, R. Urban, J. Surovková, and K. I. Afrashtehfar, "Impact of artificial intelligence on dental education: a review and guide for curriculum update," Educ Sci (Basel), vol. 13, no. 2, p. 150, 2023. https://doi.org/10.3390/educsci13020150.

[27] X. Shan, J. Cao, and T. Xie, "The Development and Teaching of the Postgraduate Course 'Engineering System Modeling and Simulation' in Combination with Essentials Taken from Research Projects," Systems, vol. 11, no. 5, p. 225, 2023. https://doi.org/10.3390/systems11050225.