

Research on the innovation of education management mode and personalized development of students based on artificial intelligence

Weina Li^{1,✉}

¹ Faculty of Education and Liberal Arts, INTI International University, Nilai, Negeri Sembilan, 71800, Malaysia

ABSTRACT

Nowadays, “Artificial Intelligence + Education” is transforming teaching and learning. In this study, we employ AI-based data mining to innovate educational management by designing an academic monitoring system using K-means clustering and developing an early warning model through stacking multi-model superposition. Targeted management measures, including personalized video recommendations, are implemented based on the model’s predictions to promote individualized student development. By analyzing daily behavior data from 500 college students, the K-means algorithm effectively classified them into four groups, and the academic alert model achieved a prediction accuracy of 84.19%, outperforming single base models. The implementation of this personalized management method significantly improved student performance compared to traditional approaches, demonstrating its potential to enhance educational outcomes.

Keywords: artificial intelligence, k-means clustering, academic alert model, educational management

1. Introduction

In the context of the current global education reform and social development, the innovation and exploration of education management mode is particularly important. The traditional education management model is often teacher-centered and lacks attention to the individual differences of students, which limits the effective allocation of education resources and the improvement of education quality [17, 7, 15]. Therefore, we need to constantly explore and practice innovative education management mode to adapt to this rapidly changing era, and it is urgent to explore a new education

✉ Corresponding author.

E-mail address: lx257257@126.com (W. Li).

Received 23 July 2024; Accepted 10 November 2024; Published Online 16 March 2025.

DOI: [10.61091/jcmcc124-09](https://doi.org/10.61091/jcmcc124-09)

© 2025 The Author(s). Published by Combinatorial Press. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

management mode to meet the needs of the new era of education development and promote the overall development of students.

Based on the problems existing in the traditional education management mode, a summary is made here. First of all, the traditional education management mode is relatively rigid, and it is necessary to promote the conceptual change of education administrators, teachers, parents and other education stakeholders, from pure knowledge transfer to focus on the overall development of students and quality education [13, 8, 20]. Secondly, the innovative education management model needs to be supported by a more innovative and professional teacher team, but the current teacher training and development system still needs to be further improved, and also needs to provide diversified educational training and development opportunities [10, 3, 5, 4]. Finally, in the education management model, attention is paid to students' interests and outstanding characteristics, and rich and diverse educational resources are provided to meet students' needs, but there is still an imbalance in the allocation of educational resources, and it is necessary to open up educational resources through the efforts of the government and society in various aspects [19, 14, 6]. Among them, students' individualized demand is a prominent feature in the new education management model. The new education management model advocates a flexible learning environment and encourages students to personalize their learning and choose their own courses [1, 11]. Students can customize their learning plans according to their own interests and abilities, and improve their motivation and initiative in learning. Instead of relying on the traditional score evaluation system, multiple evaluation methods are introduced to emphasize the development of students' comprehensive quality and ability. In addition to academic performance, students' innovative thinking, teamwork and practical ability are also evaluated [9]. These aspects of the management model to achieve personalized development of students. In addition, the implementation of education informatization is a big trend, through the Internet technology and artificial intelligence and other means, to realize the intelligent distribution or intelligent sharing of teaching resources, the monitoring of the learning process and the intelligent management, to improve the efficiency and quality of education [21, 12, 16].

Research on data mining technology based on artificial intelligence for academic early warning, as a way to grasp students' learning progress and facilitate personalized education management. Firstly, K-means clustering is used for academic monitoring to classify students' learning level according to their daily behavior data. Then we construct the Stacking multi-model superposition of academic early warning model to predict students' academic performance in advance, and carry out personalized intervention according to the early warning information prompts. We also recommend learning videos based on students' video preference and cognitive level, and implement personalized teaching management for students. Finally, a controlled experiment is designed to verify the applicability of the above management methods according to the improvement of students' performance.

2. Innovation of education management mode based on artificial intelligence

2.1. Application of artificial intelligence in education management

Artificial intelligence is the intelligent behavior exhibited by a computer program. This intelligence can include abilities such as learning, reasoning, problem solving and autonomous action. Artificial intelligence technology applied to education management can help administrators better manage and analyze a large amount of student data, teaching data, and financial data, thus improving the science

and accuracy of decision-making [22]. Through data mining, administrators can better understand students' needs and performance, optimize curriculum and resource allocation, and improve teaching quality and student satisfaction.

2.2. Research on data mining based instructional management

Students' academic status is the core content of education management, through the monitoring and effective regulation of students' learning progress, it can help students successfully complete their studies. At present, the teaching management system is not intelligent enough to monitor the academic status and manage the students' academic records, the teacher can't have a comprehensive understanding of each student's academic status, and the academic administrators will only check whether each student meets the requirements of the academic level near the end of the semester. Therefore, the effective use of cluster analysis can monitor students' learning progress in stages, grasp the real learning status of students in a timely manner, form an early warning mechanism, help disadvantaged students, standardize the process of guiding the development of students and the results, and ensure that the requirements of the academic level are reached.

2.2.1. Cluster analysis. Cluster analysis refers to the process of decomposing a large number of data objects into multiple classes based on their many characteristics without knowing the class of the research object in advance, in the process of decomposition, according to the degree of affinity between the data objects are automatically classified, the sample objects within the class have similarity in characteristics, and the characteristics of the sample objects between the different classes differ greatly. SPSS software mainly provides hierarchical clustering and K-means clustering two methods, of which K-means is its classic algorithm.

K-means is also known as fast clustering method, it is mainly by constantly adjusting the center point of the K clusters until the cohesive point position reaches the convergence criteria, if the number of samples is large or the file is huge, then it is suitable for K-means clustering method. The core steps of K-means are as follows:

- 1) The user specifies the number of clusters and SPSS will derive a unique solution according to the user's needs.
- 2) Specify the initial class center point of these K classes, commonly used methods to specify the initial class center point: empirical selection method, random selection method, min-max method.
- 3) Calculate the Euclidean distance from each object to the center point of the K classes respectively, and assign all samples according to the principle of closest distance to form K classifications.
- 4) Redetermine the centroids of the K classes using the mean value method.
- 5) Determine whether the specified number of iterations has been reached or the degree of offset of the class center point has converged, the above two conditions to meet any one of the end of the clustering, if none of them are satisfied, then return to the third step. In cluster analysis, the class to which the sample belongs will be continuously adjusted until it is stabilized.

2.2.2. K-means clustering in academic monitoring. Through the academic monitoring and early warning mechanism, schools can count and analyze students' academic performance in each semester, grasp the real learning status of students, and remind different groups of students of the corresponding warning messages. At the same time, timely support measures are provided to intervene in students' learning status and standardize the process of guiding students' development, which will play an

important role in improving the quality of teaching and learning.

Cluster analysis can be a batch of sample data according to the nature of the degree of affinity, in the absence of a priori knowledge of the case of automatic classification, to produce multiple classification results, the samples within the class in the characteristics of the similarity, the characteristics of the samples between the different classes differ greatly, so the cluster analysis can be used in many cases to find different learning status of the student group or other groups. In cluster analysis, the degree of affinity between individuals is extremely important, which will directly affect the classification results, and this degree of affinity is measured by the distance between individuals. Since the K-means algorithm is concise and convenient, and is suitable for rapid clustering of large amounts of data, the K-means clustering algorithm will be used in this part to implement academic monitoring and early warning. Applying K-means clustering analysis to academic monitoring can realize that students can be automatically divided into different types of groups according to their average grades, attendance, etc., and there are obvious differences between the groups, so that different warning messages can be proposed for different groups of students [18].

2.3. Academic early warning modeling

This section provides academic early warning for struggling students through student behavioral characteristics. In order to improve the accuracy and reliability of prediction, a student academic warning model based on Stacking multi-model stacking is proposed. The model uses a variety of single prediction models, including logistic regression, support vector machine, XGBoost, GBDT, and integrates these models through the Stacking algorithm. The student academic alert model using Stacking multi-model stacking can overcome the situation of single-model prediction bias, and the Stacking technique can fully synthesize the advantages of each single model, so as to construct a more accurate prediction model.

2.3.1. Base regressors. When choosing base regressors, each base regressor is required to provide more accurate prediction results, and the similarity between base regressors should not be too high, in order to ensure that the reliability and robustness of the overall prediction results are improved. Only after selecting base regressors with diversity, independence and accuracy can the advantages of the Stacking algorithm be fully utilized to achieve more accurate and reliable predictions. By combining different regressors, a Stacking model is constructed to accomplish the final prediction task. Through the analysis and testing of multiple regressors, Random Forest, SVM, GBDT, and XGBoost algorithms are selected as the first layer of base classifiers, and the principle of each base regressor is described as follows:

1) Forest (RF) is an integrated learning algorithm that integrates multiple decision trees together and arrives at the final result by voting or averaging. Compared with a single decision tree, the random forest algorithm can better cope with problems such as data noise and overfitting, and has high accuracy and generalization ability [2]. In the process of training the random forest algorithm, it not only samples the data randomly, but also introduces a random selection mechanism for attributes, so that the attributes used for each training are different, thus increasing the diversity of the model, reducing the correlation between models, and having better generalization ability.

2) Support Vector Machine (SVM) is an algorithm that can solve linear and nonlinear classification problems. Its core idea is that in the original space when it is not possible to find a plane or hyperplane to divide the data into the desired categories, the data can be mapped into a high-dimensional space,

so as to distinguish the data differences more clearly and effectively avoid the overfitting problem. The main principle is to achieve the purpose of classification by determining the support vectors on the boundary to have the maximum edge. The geometric space structure of two support vector machines is shown in Figure 1.

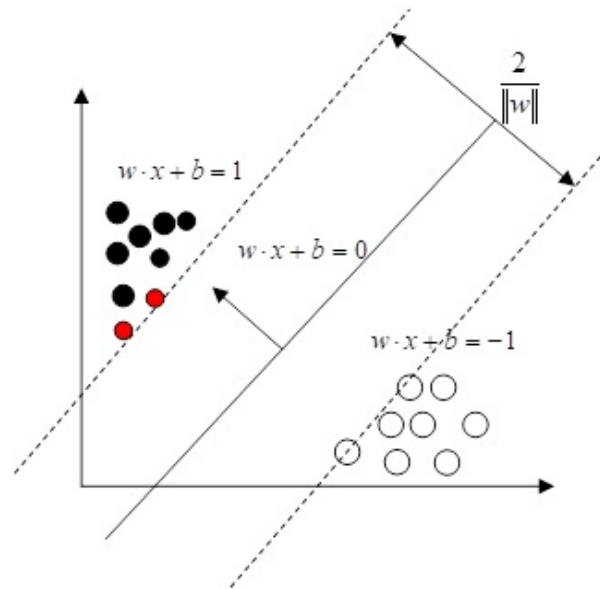


Fig. 1. Support vector machine geometry

3) The Extreme Boosting Decision Tree (GBDT) is an integrated learning algorithm whose core idea is to continuously improve the accuracy of the decision tree through Boosting techniques. Specifically, the algorithm evaluates the error between the prediction result of the previous round and the true value in each iteration, and adjusts the weights of the samples according to the error, so that the misclassified samples get higher weights in the next round of training. This method continuously adds new classifiers and, unlike traditional Bagging, it focuses more on learning from misclassified data. Compared with Random Forest and Neural Networks, it is able to excel in handling high-dimensional sparse data and can adapt to new data with stronger generalization ability.

4) Extreme Gradient Boosting (XGBoost) is a machine learning technique that aims to improve the accuracy and efficiency of a model by optimizing the basic idea of GBDT. Compared with GBDT, XGBoost is optimized in three aspects: first, it expands the choice of learner models to support not only decision trees, but also other models for training. Second, XGBoost adopts a finer regularization strategy to prevent the effects of overfitting and noise errors. It makes the model more robust and reliable, and can handle more complex data. Finally, XGBoost optimizes the loss function to make the gradient descent search more efficient and accurate, which can speed up the training of the model.

2.3.2. Meta-regressors. In this paper, Random Forest, SVM, GBDT and XGBoost algorithms are used in the Stacking multi-model stacking based academic alert model. In order to try to avoid overfitting between the base models, the second level meta-regressors need to be chosen to avoid such problems. When choosing the second meta-regressor, simple and easy to understand regressors are preferred. In this experimental model, the linear regression algorithm was chosen as the meta-regressor, both to avoid overfitting and to control the overall complexity of the model.

2.3.3. Stacking Multi-Model Stacked Academic Early Warning Models. The Stacking model is mainly composed of two layers, the first layer consists of four base regressors, namely Random Forest, SVM, XGBoost and GBDT. The four base regressors are trained using the same feature samples, and the training process of each base regressor does not affect each other. After the training is completed, each base regressor outputs a one-dimensional vector, and the output vectors of the four base regressors form a new input vector to be used as the input samples of the second layer of the Stacking model. In the second layer, the LR algorithm is used as a meta-regressor, and the output vectors of the four base regressors are used as inputs to predict students' grades. This constitutes the Stacking predictive model with a two-layer structure. The overall framework of the Stacking multi-model stacking academic alert model is shown in Figure 2. The first step is to input the original dataset and use the random forest model for feature selection, the second step is to input the data features into the four base regressors in the first layer of the Stacking model for prediction, and then the final prediction is made by the meta-regressor in the second layer based on the prediction results of the four base regressors.

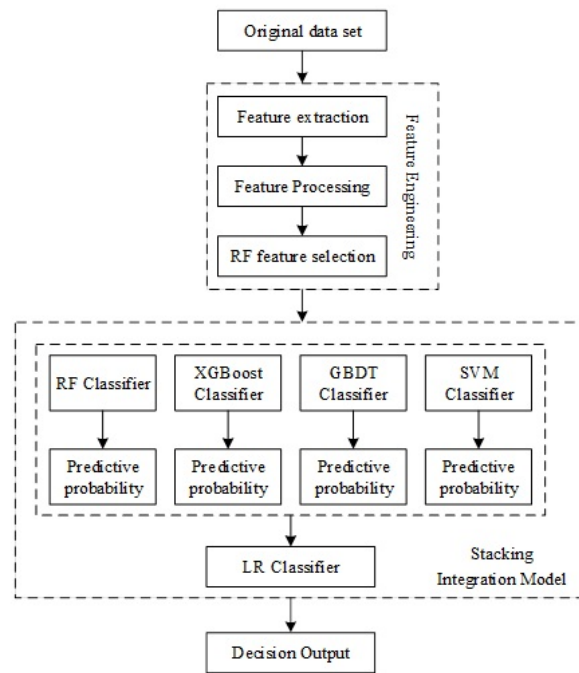


Fig. 2. The overall framework of the academic early warning model

2.3.4. Academic early warning model evaluation indicators. In this paper, the academic early warning model is evaluated using the metrics of accuracy, precision, recall and $F1$ -value to get a more comprehensive understanding of the model's performance from different perspectives. If the experimental results show that the model's accuracy, precision, recall and $F1$ -value are high, it indicates that the model's prediction results are reliable and have high accuracy and generalization. The formulas for accuracy and precision are shown in (1) and (2):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (2)$$

Recall indicates the proportion of correctly predicted positive to the total positive samples. The formula is shown in (3):

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (3)$$

The F1-score is a reconciled value that combines Precision and Recall and is used to reflect the overall metrics, calculated as shown in (4):

$$F_1 = \frac{2TP}{2TP + FN + FP}. \quad (4)$$

3. Design of individualized education management methods for students

In this paper, we set up different early warning information prompts and personalized intervention strategies for students with different achievement levels.

On the one hand, according to the academic early warning model constructed above to predict the academic performance of learners, students with scores of less than 60 points are given a high degree of early warning, the specific measures are to call their parents to inform them of the integration of home and school supervision, and at the same time, personalized learning guidance and similar learning peers are recommended for the learners. For learners with scores between 60 and 70, a medium level of warning will be given, and specific measures will be taken to send out e-mails and study progress notes to urge learners to study in a timely manner. For learners with scores between 70-80, a low alert is given by sending a verbal reminder and a visual learning report to the learner. For students with scores above 80, no warning is given, and a visual learning report is sent directly to allow students to view the details of their mistakes and summarize them.

On the other hand, personalized learning video recommendations are made to all learners. Mining learners' cognitive level and interest preferences, and combining the difficulty of knowledge points, to build a personalized learning video recommendation model. We recommend learning videos with personalized preferences for learners with different levels of achievement that meet the learners' mastery of knowledge points.

3.1. Video rating prediction

The quantification of the ancillary domain rating data can be calculated in terms of both the learner's preference for the video and the cognitive level, as shown in Eq. (5):

$$R_a(u_i, v_k) = 5 \times (\varepsilon P_v(u_i, v_k) + (1 - \varepsilon) F_c(u_i, v_k)), \quad (5)$$

where $R_a(u_i, v_k)$ denotes the implicit rating and ε denotes the proportion of weights assigned to interest preferences and cognitive levels.

The learner's rating of the video is weighted by the auxiliary domain rating data and the target domain rating data, as shown in Eq. (6):

$$R(u_i, v_k) = \alpha R_{ik1}(u_i, v_k) + (1 - \alpha) R_{ik2}(u_i, v_k), \quad (6)$$

where, $R(u_i, v_k)$ is the value of the final learner's rating of the video and α is the fusion ratio of implicit and explicit ratings.

However, when choosing videos, learners mostly habitually choose videos with higher popularity and score such videos higher. However, it is difficult for this type of videos to reflect the differences in

interests among learners, instead, cold videos better reflect the interests among learners. Therefore, in order to prevent the negative impact of the above long-tail phenomenon on the similar results, this paper adds the item popularity penalty factor, as shown in Eq. (7):

$$weight(i, j) = \frac{1}{\log(1 + s(i, j))}, \quad (7)$$

where $weight(i, j)$ denotes the popularity penalty factor for video i and video j , and $s(i, j)$ denotes the number of learners who rated both video i and video j .

The rating data of all learners for video i is taken as a vector, which can be expressed as $I_i = \{R_{1i}, R_{2i}, \dots, R_{ni}\}$, and similarly for video j as $I_j = \{R_{1j}, R_{2j}, \dots, R_{rij}\}$, removing some of the learners' rating habits and reducing the effect of popularity. The specific calculation is shown in Eq. (8):

$$sim(i, j) = weight(i, j) \times \frac{\sum_{u \in U_{R_i} \cap U_{R_j}} (R_{ui} - \bar{R}_i) (R_{uj} - \bar{R}_j)}{\sqrt{\sum_{u \in U_{R_i}} (R_{ui} - \bar{R}_i)^2} \sqrt{\sum_{u \in U_{R_j}} (R_{uj} - \bar{R}_j)^2}}, \quad (8)$$

where $sim(i, j)$ denotes the similarity between video i and video j , R_{ui} denotes the rating of video i by learner u , \bar{R}_i and \bar{R}_j denote the average rating of video i and video j , respectively, and U_{R_i} denotes the set of learners who have rated video i .

3.2. Semantic similarity calculation

For most online courses, there is a one-to-many correlation between videos and knowledge points and one or more paths between any two knowledge points, and the distance between these paths can be used to characterize the semantic similarity between videos.

The entity-relationship triad data in the video knowledge graph is known, and in this paper, we utilize the PTransE model to embed the existing entity-relationship triad information into a low-dimensional vector space and generate the corresponding semantic vectors. The semantic relationship path consists of directly connected knowledge points and indirectly connected knowledge points. In order to avoid the path is too long, resulting in reduced computational efficiency, this paper chooses the length of 3 and the reliability score is greater than 0.01 of the relationship path. In this paper, each video v_i is represented as a vector: $V_i = (v_{1i}, v_{2i}, \dots, v_{di})^T$, and a weighting factor is added when calculating the similarity of knowledge points between two videos to improve the accuracy of the similarity.

Firstly, function $Y_1(v_i, v_j)$ is utilized to represent the number of direct relationships between videos v_i and v_j , and then function $Y_2(v_i, v_j)$ is utilized to represent the number of indirect relationships between videos v_i and v_j , and the similarity calculation is specifically shown in Eq. (9):

$$sim_{PTransE}(v_i, v_j) = \frac{1}{E(h, P, t) + 1} [\lambda Y_1(v_i, v_j) + \lambda^2 Y_2(v_i, v_j)], \quad (9)$$

where λ is used as the relationship weight factor, in practice, the two videos with direct relationship have more weight than those with indirect relationship, so the direct relationship weight is set to λ and the indirect relationship weight is set to λ^2 .

$E(h, P, t)$ is used to model the correlation between triples in multiple relationship paths, and the specific calculation is shown in Eq. (10):

$$E(h, P, t) = \frac{1}{Z} \sum_{p \in p(h, t)} R(p|h, t) E(h, p, t), \quad (10)$$

where $E(h, P, t) = \|h + p - t\|$ denotes the energy function formed by the triad, i.e., the similarity measure $d(h + p, t)$ of the two learning video knowledge points under the semantic relation p , and d is the Euclidean distance. P denotes the set of multiple paths that can be reached between two knowledge points, $R(p|h, t)$ denotes the probability that the head entity h reaches the tail entity t through the relationship path r , and Z is the normalization factor.

For a path $P = (r_1, r_2, \dots, r_g)$ between two knowledge points, where $e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \dots e_g$, for any knowledge point $m \in e_i$, the probability of reaching m along a semantic relation p is specified as shown in Eq. (11):

$$R_p(m) = \sum_{n \in e_{i-1}(\cdot, m)} \frac{1}{|e(n, \cdot)|} R_p(n), \quad (11)$$

where $e(n, \cdot)$ denotes the direct successor of e_{i-1} , and $R_p(n)$ denotes the probability of following the semantic relation path p and wandering to the knowledge point n .

A large number of experimental results have proved that the semantic effect is best when using the additive model in the PTransE algorithm for combining multi-path relations. Therefore, this paper adopts the additive model in the semantic combination, such as the path from e_0 to e_1 is: $P = r_1 \circ r_2 \circ r_3 \circ \dots \circ r_g$, and the semantic combination is specifically shown in Eq. (12):

$$P = r_1 + r_2 + r_3 + \dots + r_g. \quad (12)$$

3.3. Generating a list of video recommendations

Traditional collaborative filtering algorithms are less likely to simultaneously consider the sparsity of scoring data, project scalability, and cold start when similarity is calculated. In view of this, this paper aims to improve the traditional collaborative filtering algorithm by considering both the learner's personalized behavioral information and the semantic information contained in the videos when similarity is calculated. Therefore, the similarity of learners' ratings of learning videos and the semantic similarity between learning videos are calculated by linear fusion, as shown in Eq. (13):

$$\text{sim}(v_i, v_j) = (1 - \alpha)\text{sim}_{CF}(v_i, v_j) + \alpha\text{sim}_{PTransE}(v_i, v_j). \quad (13)$$

The top n learners with the highest similarity to the target learner are treated as the nearest neighbor learners, and the learning videos watched by the nearest neighbor learners are used as a collection of videos to be recommended, and the target learner's rating for each video in the collection is predicted, and the videos are sorted according to the size of the ratings to generate a recommendation list. The target learner predicts the rating of a particular video as shown in Eq. (14):

$$P_{\text{prediction}}(u_i, v_i) = \bar{R}_{u_i} + \frac{\sum_{u_j \in \text{neighbor}(u_i)} \text{sim}(v_i, v_j) (R_{j k}(u_j, v_i) - \bar{R}_{u_j})}{\sum_{u_j \in \text{neighbor}(u_i)} |\text{sim}(v_i, v_j)|}, \quad (14)$$

where $P_{\text{prediction}}(u_i, v_i)$ denotes the predicted score of learner u_i for video v_i , $\text{sim}(v_i, v_j)$ denotes the similarity of video v_i, v_j , and $\text{neighbor}(u_i)$ denotes the set of near-neighbor videos of learner u_i .

4. Empirical analysis of education management optimization based on artificial intelligence

The student daily behavior data studied in this section were obtained from 500 undergraduate students of a college in the class of 2022 in University M. The timeframe was two full semesters of

the academic year 2022-2023, in which the specific daily behavior data included the usual classroom performance behavior data, the one-card consumption behavior data, and the library borrowing behavior data, and the student's final exam score data were also obtained as a prediction object.

4.1. Cluster analysis of student learning

Using K-means algorithm for clustering analysis of student performance data and library access data, the clustering results of student learning data are shown in Figure 3. There are 4 groups of students, each group accounts for 21%, 26%, 13% and 40% respectively.

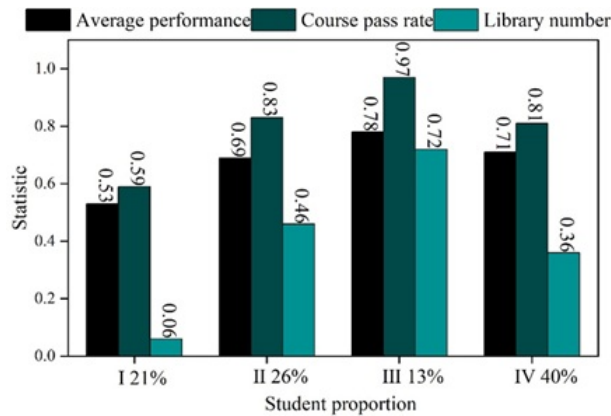


Fig. 3. Learning data clustering results

The average grades and course pass rate of students in Group I are low, and the number of times they go to the library is very small, which indicates that this group of students do not study hard enough, and the teacher should correct the learning attitude of this group of students.

The average grade of students in Group II is in the middle to upper level, the course passing rate is relatively high, and the number of times they go to the library is also high, which indicates that this group of students is particularly hardworking in their study, and belongs to the practical and diligent type of students. However, their performance is not very good because their learning methods are not easy to understand. Teachers should instruct this group of students on learning methods to improve their learning efficiency.

The average score of students in group III is very good, and the passing rate of the course is almost 100%, which means that this group of students belongs to the “bully” type of students who study very hard.

Students in group IV are in the middle of the scale compared to other types of students. Teachers should give encouragement to this group of students to increase their motivation to study.

4.2. Performance analysis of academic early warning models

In this section, the performance of Stacking's multi-model stacked academic alert model is tested through experiments, and after dividing the training set and test set, support vector machine, GBDT, XGBoost, random forest, and logistic regression are used to predict whether a student needs to be alerted or not, respectively, and the prediction results are shown in Fig. 4. It can be seen that Stacking's multi-model stacked academic warning model has an accuracy, precision, recall, and F1 score of 0.8419, 0.8391, 0.8437, and 0.8522, respectively, which are better than a single base model in the problem of predicting student performance. It shows that the method of this paper can effectively warn the students who cannot reach the academic level, which is important for the educational

management of students.

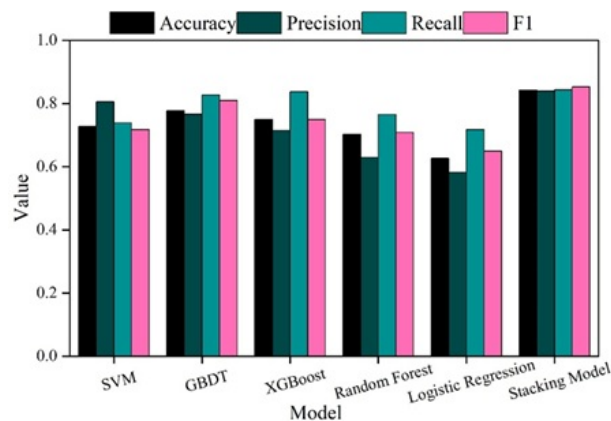


Fig. 4. Performance comparison of various model algorithms

4.3. Analysis of the effects of individualized student development

In this section, classes (1) and (2), with 45 students in each of the two classes, of a comparable level of specialization in a college in M, were selected for a controlled experiment. Class (1) as an experimental class uses the personalized management method designed in this paper for teaching management. Class (2) as the control class uses the traditional management mode. Before the beginning of the experiment, students' test scores in specialized courses A, B, C and D were counted. The students' examination scores in the four specialized courses are counted again at the end of one semester.

4.3.1. Pre-test differences between experimental and control classes. Before the beginning of the experiment, the grades of the four specialized courses of the two classes were counted separately and the paired samples t-test was performed. The differences between the experimental class and the control class on the pre-test are shown in Table 1. The P-values of the two classes in the four specialized courses are 0.253, 0.224, 0.218, and 0.209, respectively, indicating that there is no significant difference between the pretest scores of the experimental class and the control class, which can be used for the control experiment.

Table 1. Test differences between the experimental class and the comparison class

Specialized course	N	Laboratory class	Cross-reference class	T	P
A	90	65.32±9.72	65.22±8.64	0.22	0.253
B	90	67.63±6.33	67.59±6.67	0.16	0.224
C	90	69.12±7.19	69.33±6.87	0.34	0.218
D	90	64.36±6.55	64.28±6.32	0.16	0.209

4.3.2. Pre- and post-test differences in experimental classes. The data from the pre- and post-tests of the test scores of the experimental class were analyzed, and the average pre- and post-test scores obtained are shown in Table 2. The students' scores in the four professional courses before the experiment (1) were low, indicating that the students' knowledge of the courses was low, but after the adoption of the personalized management approach to teaching management, the scores of the

experimental class rose relatively significantly. From the table, it can be learned that the P-values of the test scores of the four specialized courses before and after the experiment are 0.006, 0.003, 0.004, 0.000, respectively, which are all significantly different, and the performance of the experimental class students' performance improvement is significant.

Table 2. Contrast between the experimental class

Specialized course	N	Pretest	Posttest	T	P
A	45	65.32±9.72	83.96±4.51	3.26	0.006
B	45	67.63±6.33	86.49±5.63	2.34	0.003
C	45	69.12±7.19	89.63±4.11	3.11	0.004
D	45	64.36±6.55	84.74±6.51	1.96	0.000

4.3.3. Pre- and post-test differences between control classes. A comparison of the differences between the pre- and post-tests of the control class is shown in Table 3. The semester test scores of the control class in the four specialized courses and also their mean scores do not improve very much, indicating that the students' knowledge and ability have not been greatly changed, and generally remain at a similar level. In addition, the paired-sample t-test was conducted on the pre- and post-test data of the semester test scores as well as the case test scores of the control class, and it can be learned from the comparison of pre- and post-test differences of the test scores of the control class that the p-values of the pre- and post-test scores before and after the experiment are all greater than 0.05, and there is no significant difference, and although the scores of the control class students have been slightly enhanced, the enhancement effect is not very obvious.

Table 3. Contrast between the comparison of the comparison

Specialized course	N	Pretest	Posttest	T	P
A	45	65.22±8.64	68.16±7.56	1.36	0.126
B	45	67.59±6.67	68.42±6.33	1.24	0.059
C	45	69.33±6.87	68.51±6.42	1.11	0.076
D	45	64.28±6.32	68.32±6.44	2.03	0.234

4.3.4. Differences in posttests between experimental and control classes. The differences between the experimental and control classes are shown in Table 4. The average posttest scores of the experimental class ranged from 83.96 to 89.63 on the four specialized courses. The average posttest scores of the control class ranged from 68.16 to 68.51. As analyzed above, the two pre-test scores are close to each other, but there is a significant gap between the scores after different teaching modes. The paired samples t-test of the pre and post-test scores of the experimental and control classes showed that there was a significant difference between the scores of the two classes in the four specialized courses $P=0.000 > 0.05$. This indicated that the two classes were at a comparable level of learning before the experiment began. However, in the post-test, there was a significant difference between the scores of the two classes, which fully proves that the personalized education management method implemented in the experimental class has achieved significant results in classroom efficiency. Within the same teaching time, this teaching mode can effectively promote the improvement of students' performance.

Table 4. The experimental class and the comparison class were measured

Specialized course	N	Laboratory class	Cross-reference class	T	P
A	90	83.96±4.51	68.16±7.56	3.56	0.000
B	90	86.49±5.63	68.42±6.33	3.22	0.000
C	90	89.63±4.11	68.51±6.42	4.25	0.000
D	90	84.74±6.51	68.32±6.44	5.33	0.000

5. Conclusion

Research on the use of K-means clustering for academic monitoring to grasp the status of students' learning progress. Construct Stacking multi-model superimposed academic warning model and learning video recommendation model to guide students' personalized development in a targeted way. Use K-means clustering method to classify 500 student categories in college M. The method reasonably divides students into 4 categories. The accuracy, precision, recall, and F1 score of the designed academic alert model are 0.8419, 0.8391, 0.8437, and 0.8522, respectively, which outperform the single base model. After utilizing the designed academic early warning model and learning video recommendation model for education management, students' performance is significantly improved compared to the traditional education management model. The method of this paper plays an obvious promotion effect for students' personalized development.

References

- [1] A. J. Bingham, J. F. Pane, E. D. Steiner, and L. S. Hamilton. Ahead of the curve: implementation challenges in personalized learning school models. *Educational Policy*, 32(3):454–489, 2018. <https://doi.org/10.1177/0895904816637688>.
- [2] V. H. Chau, A. T. Vo, and H. P. Ngo. Retracted: powerlifting total score prediction based on an improved random forest regression algorithm. *Journal of Intelligent & Fuzzy Systems*, 46(4):9999–10004, 2024. <https://doi.org/10.3233/JIFS-230032>.
- [3] V. Ellis, S. Steadman, and Q. Mao. 'come to a screeching halt': can change in teacher education during the covid-19 pandemic be seen as innovation? *European Journal of Teacher Education*, 43(4):559–572, 2020. <https://doi.org/10.1080/02619768.2020.1821186>.
- [4] L. Florian and D. Camedda. Enhancing teacher education for inclusion. *European Journal of Teacher Education*, 43(1):4–8, 2020. <https://doi.org/10.1080/02619768.2020.1707579>.
- [5] E. Garzón Artacho, T. S. Martínez, J. L. Ortega Martin, J. A. Marin Marin, and G. Gómez García. Teacher training in lifelong learning—the importance of digital competence in the encouragement of teaching innovation. *Sustainability*, 12(7):2852, 2020. <https://doi.org/10.3390/su12072852>.
- [6] R. Huang, A. Tlili, T.-W. Chang, X. Zhang, F. Nascimbeni, and D. Burgos. Disrupted classes, undisrupted learning during covid-19 outbreak in china: application of open educational practices and resources. *Smart Learning Environments*, 7:1–15, 2020. <https://doi.org/10.1186/s40561-020-00125-8>.
- [7] S. E. Kaymakamoglu. Teachers' beliefs, perceived practice and actual classroom practice in relation to traditional (teacher-centered) and constructivist (learner-centered) teaching (note 1). *Journal of Education and Learning*, 7(1):29–37, 2018. <https://doi.org/10.5539/jel.v7n1p29>.

-
- [8] J. Li. Analysis of innovative strategies for college student education management based on "three comprehensive education". *Curriculum Learning and Exploration*, 2(1):88–90, 2024. <https://doi.org/10.18686/cle.v2i1.3788>.
- [9] W. Lu, C. Jia, and J. Zuo. Application of fuzzy comprehensive evaluation in comprehensive quality evaluation of higher education students. *International Journal of Emerging Technologies in Learning*, 16(12):201–214, 2021. <https://doi.org/10.3991/ijet.v16i12.23327>.
- [10] K. Mohanna. The all-round flexible and adaptable teacher. In *Your Teaching Style*, pages 23–29. CRC Press, 2018.
- [11] C. Müller and T. Mildenerger. Facilitating flexible learning by replacing classroom time with an online learning environment: a systematic review of blended learning in higher education. *Educational Research Review*, 34:100394, 2021. <https://doi.org/10.1016/j.edurev.2021.100394>.
- [12] W. C. Ng, W. Y. B. Lim, J. S. Ng, Z. Xiong, D. Niyato, and C. Miao. Unified resource allocation framework for the edge intelligence-enabled metaverse. In *ICC 2022-IEEE International conference on Communications*, pages 5214–5219. IEEE, 2022. <https://doi.org/10.1109/ICC45855.2022.9838492>.
- [13] M. Oda Abunamous, A. Boudouaia, M. Jebril, S. Diafi, and M. Zreik. The decay of traditional education: a case study under covid-19. *Cogent Education*, 9(1):2082116, 2022. <https://doi.org/10.1080/2331186X.2022.2082116>.
- [14] S. M. Reis and J. S. Renzulli. The schoolwide enrichment model: a focus on student strengths & interests. In *Systems and Models for Developing Programs for the Gifted and Talented*, pages 323–352. Routledge, 2023.
- [15] A. Rohman, S. Muhtamiroh, A. Imron, and N. Miyono. Integrating traditional-modern education in madrasa to promote competitive graduates in the globalization era. *Cogent Education*, 10(2):2268456, 2023. <https://doi.org/10.1080/2331186X.2023.2268456>.
- [16] Y. Sun, L. Tian, M. Huang, and L. Yang. Exploration of artificial intelligence technology in the application of academic monitoring for college students. In *Proceedings of the 2024 International Symposium on Artificial Intelligence for Education*, pages 546–551, New York, NY, USA. Association for Computing Machinery, 2024. <https://doi.org/10.1145/3700297.3700391>.
- [17] A. Vallée, J. Blacher, A. Cariou, and E. Sorbets. Blended learning compared to traditional learning in medical education: systematic review and meta-analysis. *Journal of Medical Internet Research*, 22(8):e16504, 2020. <https://doi.org/10.2196/16504>.
- [18] H. Wang. Optimization of ideological and political education management strategies under k-means algorithm in big data environment. *Security and Communication Networks*, 2022:1–14, Apr. 2022. <https://doi.org/10.1155/2022/6120230>.
- [19] Q. Wang, H. Zhu, X. Liu, and D. L. Caseldo. The understanding and analysis of education management reform. *The Educational Review, USA*, 6(8):438–445, 2022. <http://dx.doi.org/10.26855/er.2022.08.017>.
- [20] X. Xiu, W. Song, and Y. Chen. Train new talents in the era of all-round development through labor education. *Frontiers in Educational Research*, 4(1):82–87, 2021. <https://doi.org/10.25236/FER.2021.040117>.
- [21] L. Yufeia, S. Salehb, H. Jiahuic, and S. M. Syed. Review of the application of artificial intelligence in education. *Integration (Amsterdam)*, 12(8):1–15, 2020. <https://doi.org/10.53333/IJICC2013/12850>.

-
- [22] J. Zhao and Y. Li. Construction of higher education management cloud space based on machine learning and artificial intelligence. *Journal of Artificial Intelligence Practice*, 7(3):146–155, 2024. <https://doi.org/10.23977/jaip.2024.070318>.