International Journal for Housing Science and Its Applications
Publish September 9, 2025. Volume 47, Issue 1 Pages 708-720

https://doi.org/10.70517/ijhsa47162

Development and Promotion of Multi-source Data Fusion Technology in Smart Classroom Construction

Bei Zhang[1] and Lin Zhang[2,*]

[1] College of Teacher Education, Xingtai University, Xingtai, Hebei, 054000, China
[2] Research and Development Center, Hebei Institute of Machinery & Electricity, Xingtai, Hebei, 054000, China

Corresponding authors: (e-mail: zhanglin1992401@163.com).

Abstract The current educational landscape is undergoing subtle changes, and smart classrooms urgently require upgrades and innovations. This study takes School A, which is undergoing educational reforms, as an example. First, using a weekly time unit, we collected a series of data on students' consumption behavior (including consumption frequency, consumption time, and consumption amount), online behavior (including online frequency and online time), and daily routines (including start time of activities and preparation time for rest) over two semesters, thereby constructing a multi-source dataset of students' daily behaviors. The data is preprocessed and corresponding features are extracted. Then, a multi-feature combination-based online course performance prediction model is proposed. Finally, a student academic performance management system based on multi-source information fusion technology is constructed. Through experiments, the accuracy rate of the student performance prediction model based on feature combinations can reach over 90%. The student academic performance management system constructed based on this model is accepted by most students and can significantly improve student performance compared to traditional models. Additionally, applying multi-source information fusion technology to teaching, learning, and management processes can promote the joint development of schools and students.

Index Terms smart classroom, multi-source data fusion, academic performance management system, performance prediction model

I. Introduction

With the rapid development of information technology, traditional teaching models can no longer meet the learning needs of modern students. Smart classrooms have emerged in response to this need, aiming to transform traditional classrooms into digital, intelligent learning environments by introducing advanced educational technology equipment and software [1]. Smart classrooms provide a wealth of diverse teaching resources, interactive methods, and personalized learning support, stimulating students' interest in learning, creativity, and thinking abilities, while enhancing teaching effectiveness and student experience [2], [3]. With the support of technologies such as the Internet of Things (IoT), 5G, and edge computing, smart classrooms can embody the three layers of the IoT (application layer, network layer, and perception layer). By utilizing sensors, radio-frequency identification (RFID), and other technologies, information sensing devices can real-time perceive any required information, transmit it via possible networks (such as Wi-Fi-based wireless local area networks, mobile communications, and telecommunications networks) to connect any object to the internet, enabling information exchange and communication, achieving ubiquitous connectivity between objects and between objects and people, and enabling intelligent identification, tracking, monitoring, and management of objects [4]-[8].

In smart classrooms, various technical devices capture data on classroom environments, student behavior, and teacher behavior. Combined with device operation data and intelligent teaching management system data, these form multi-source information data. However, this data is often used individually or in pairs, leading to a lack of precision in decision-making regarding student management and teaching methods [9]-[11]. For example, most studies utilize teacher-student behavior data for classroom management to enhance teacher-student interaction and student engagement; they also combine student behavior data with intelligent teaching management system data to provide personalized teaching guidance, but they overlook students' physical and mental factors as well as environmental factors [12]-[13]. The application of device operation data and environmental data in smart classrooms is limited, and the phenomenon of data silos is severe. Multi-source data fusion technology is a key solution to addressing data silos.

Currently, the development of multi-source data fusion technology is in a phase of rapid advancement, with limited research specifically focused on smart classrooms. Liu et al. [14] converted multi-source data into a standardized format based on the Resource Description Framework (RDF), introduced data fusion algorithms, thereby achieving the fusion of heterogeneous multi-source data, and also addressed semantic conflicts during the fusion process. Sun and Ren [15] shared a multi-source heterogeneous data fusion method for intelligent systems, combining IoT technology, hybrid information gain strategies, system computing and storage capabilities, analyzing the correlation between different types of data, and integrating different datasets to achieve data fusion. Cai [16] utilized fuzzy neural network algorithms, the Delphi method, and predictive missing value estimation methods to construct a multi-source sensor data fusion system, which has been widely applied in IoT systems. Qin et al. [17] represented single-source information data using intervals, processed the heterogeneity of fuzzy data under the developed dynamic interval standardization algorithm, and improved the quality of single-source fuzzy data under the standardized interval fusion representation model, effectively achieving dynamic fusion of multi-source fuzzy data. Liu et al. [18] fused two sets of road network sensor data using a three-layer backpropagation neural network. They further combined particle swarm optimization algorithms and genetic algorithms to develop a multi-source data fusion model, thereby improving the accuracy of decision estimates.

The study constructed a multi-feature combination-based academic performance prediction model using a multi-source dataset of students' daily behaviors. First, the model utilizes deep neural networks to automatically perform feature engineering, reducing the need for manual intervention. Second, the model uses factorization machines and two neural networks to simultaneously consider the effects of first-order, second-order, and higher-order features, fully learning the relationship information between each feature and academic performance. Integrating this model into a student academic performance management system can provide learners with more efficient academic performance management strategies. The analysis section of the paper conducts experimental tests on model performance and system application effectiveness to validate the scalability of multi-source data fusion technology.

II. Application of multi-source data fusion technology in smart classrooms

Most existing smart classroom practices focus on flexible classroom layout, intelligent control of the physical environment, cloud platform construction for teaching resources, and cloud recording and broadcasting system construction for the teaching process [19]. There is a general lack of support for enhancing teaching—the ability to deeply apply immersive, interactive, and engaging learning experiences—and especially a lack of support for comprehensively recording students' daily behavioral data for academic performance evaluation. This urgently requires upgrading smart classrooms to support current teaching practices.

Multi-source data fusion is a comprehensive data processing paradigm that integrates, processes, extracts, and manages heterogeneous data sources. This technology aims to break through the data silo effect by constructing a multi-dimensional correlation analysis model, enabling in-depth mining, systematic analysis, and intelligent integration of cross-source information. This ultimately forms an organically coordinated data ecosystem, helping information analysts establish a global cognitive framework [20].

This study focuses on student academic performance management and proposes to use multi-source information fusion technology to predict student grades. By developing a student academic performance management system, it aims to provide insights for the construction and upgrading of smart classrooms.

II. A. *Construction of a multi-source dataset on students' daily behavior*

II. A. 1) Data Source Selection

School A has been fully implementing a smart teaching model since 2020, achieving significant progress in teaching effectiveness. This experiment collected data generated by School A students in their daily lives, including consumption data, internet usage data, daily routine data, basic student data, and mental health data. Consumption data primarily comes from sensors distributed across the campus, such as card readers and campus card recharge machines in canteens, bathrooms, and water rooms. These sensors efficiently and comprehensively collect the majority of students' consumption behaviors on campus and transmit the collected data to the database of the relevant administrative department for storage. Internet usage data is collected through the campus network system. Daily routine data is analyzed through statistical analysis of students' multifaceted behavioral data. Student basic information is primarily obtained through the orientation system. Student mental health information is primarily managed by counselors, health center staff, and doctors. To ensure efficient operation of subsequent experiments, the datasets collected from different data sources are stored in a unified database.

II. A. 2) Data preprocessing

During data collection, issues such as data redundancy and loss often arise. If these problematic data are not addressed promptly, it can reduce the validity of the dataset and impact the performance of subsequent experimental models. The data preprocessing required for this experiment can be categorized into the following types:

(1) Standardization

In the various databases managed by different departments, each database has its own standards, which may not be consistent. Therefore, after data collection, the data formats must be standardized, such as converting all time data to time type and all floating-point data to single-precision floating-point type.

(2) Feature Embedding

Feature embedding involves converting real-world text, images, and other data into a format that computers can recognize while preserving the original information. Since the input data for the model in the experiment must be in vector form that the network can learn and compute, and the collected data includes text information such as housing type and family type that cannot be directly input into the model, feature embedding is used to convert this information into vectors that the model can recognize, enabling these details to participate in the model's training and learning processes.

(3) Missing data

For missing data issues, it is first necessary to determine whether the data does not exist or whether it was not recorded successfully for various reasons. For example, some students may not like to eat in the cafeteria and have a low frequency of cafeteria consumption, resulting in many blank records in their consumption records. These records are normal data that reflect the student's dining preferences and are not considered missing data. If a student's consumption record table contains an entry with a consumption time but no consumption amount, this is considered a missing data case and requires handling. For the scale of missing data, if the missing data exceeds 50%, the data is deemed unusable and should be discarded. Otherwise, linear interpolation is applied to the data to ensure temporal consistency.

(4) Noise data processing

There will be some abnormal data in the collected data, and this abnormal data will become noise that affects normal data.

(5) Data redundancy

Since data is collected from multiple data sources, a piece of data may exist in multiple data sources, and when merging data sources, this data will be repeated multiple times. For example, consider a student's cafeteria consumption. If a student makes a purchase at the cafeteria at noon, but this record exists in multiple data platforms, then when integrating the data sources, it may appear as though the student made multiple purchases at noon. Therefore, it is necessary to use other fields in the data to determine whether the record is redundant. If redundancy is detected, it should be removed promptly.

(6) Data normalization

After addressing the above data issues, the final step in data processing is normalization, which involves mapping the data to the range of 0 to 1 using relevant functions. This reduces the complexity of data input into the model for computation, shortens computation time, and accelerates convergence to the optimal solution. According to statistics, this experiment collected a dataset of 8,053 students' daily behavior data. The sizes of the various sub-datasets in the dataset are shown in Table 1.

Table 1: Data set size

Data set	The size of the data set
Student basic information set	8053
Student consumption data set	1952053
Student data set	1235452
Student routine data set	262452
Mental health information set	2014
The life is based on the coefficient	10452
Learning relational data set	187563

II. A. 3) Feature extraction

After data preprocessing, the dataset has been preliminarily constructed. The next step is to extract relevant features from the data and input them into the model for training. Based on the characteristics of the data, it is divided into static data and dynamic data. The former does not change over time, such as students' majors and places of origin. The latter changes over time, such as students' dining hall spending amounts and campus internet usage duration. Therefore, static features and dynamic features are extracted from the data in the dataset, as detailed below:

(1) Student static feature data

Extracting student static feature data helps analyze the relationship between students' mental health status and their personal circumstances, family economic conditions, and other static features. This enables a better understanding of the characteristics of students with mental health issues, allowing for early identification and timely intervention and psychological counseling. The static information features extracted in this experiment are shown in Table 2. The left column lists the feature names, and the right column lists the corresponding style values.

Table 2: Student static feature data sheet

Feature	Meaning	style
bh	Numbering	0001
xfly	Tuition	Household income
myshf	Monthly living expenses	1500
jkzk	Health status	health
zfxz	Occupancy	self-building
jtlx	Family type	Single parent
jtsrlx	Family income type	Family income
nsr	Annual income	250000
jtrks	Family population	4
jtldls	Family number	2
jtzxss	Family enrollment	2
fzje	Amount of liability	0
sfzrzh	Natural disaster	No

(2) Student dynamic feature data

Extracting dynamic data features can analyze the correlation between daily behavior and mental health. The student dynamic features extracted in this experiment are shown in Table 3. These are the daily behavior features of students over two semesters, with a time unit of one week. For consumption data, since the cafeteria is generally only open at fixed times, the cafeteria consumption data is divided into three categories based on the cafeteria opening hours: morning, noon, and evening. The consumption time, frequency, and amount for each of these three categories are then extracted. Additionally, evening consumption in bathrooms or washrooms is often for pre-sleep hygiene, so these consumption records are approximated as students' sleep times. Morning cafeteria consumption times, which correspond to breakfast times, can be approximated as students' wake-up times. For campus network data, extracting students' online duration and frequency features enables analysis of their online dependency.

Table 3: Student dynamic feature data sheet

Feature	Meaning	style
zsxfsj	Morning consumption time	8:10
zsxfcs	Morning consumption	3
zsxfje	Morning amount	3
zwxfsj	Noon time	11:40
zwxfcs	Noon consumption	5
zwxfje	Noon amount	13.5
wsxfsj	Night time	17:30
wsxfcs	Night consumption	5
wsxfje	Evening amount	10
swcs	Internet number	9
swsc	Internet length	2:10
kshdsj	Start time	7:30
zbxxsj	Prepare for a break	23:00

For subsequent experimental processing, students' dynamic characteristics were divided into consumption data, internet usage data, and daily routine data. Consumption data mainly includes meal times, meal frequencies, and meal amounts for three meals a day. Internet usage data includes the number of times online and the duration of internet usage. Daily routine data includes the start time of activities and the preparation time for rest. All feature classifications are shown in Figure 1.

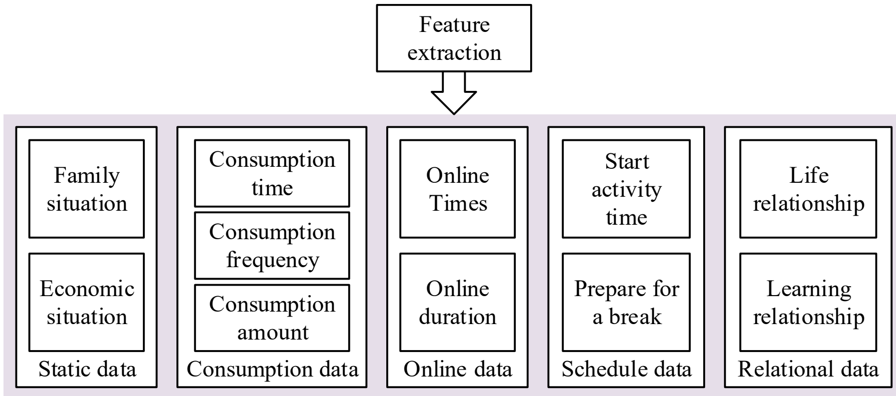

Figure 1: Feature classification map

II. B. Student performance prediction model based on feature combinations

II. B. 1) Problem Definition

Given a set of student characteristics F, where F consists of static characteristics $Static = \{s_1, s_2, \cdots, s_m\}$ and dynamic characteristics $Dynamic = \{d_1, d_2, \cdots, d_k\}$, i.e., $F = \{Static, Dynamic\}$, where m and k represent the number of features, respectively. For this student, his final course grade is y_i, and $y = \{0,1\}$ is the set of categories into which student grades are divided, where 0 indicates that the student failed the course and 1 indicates that the student passed the course. The goal of student grade prediction is to predict the grade category y_i based on the student's features F.

II. B. 2) Model Framework

This paper combines the static and dynamic characteristics of students extracted in the previous section to propose a feature-based performance prediction model (FDPN), whose framework is shown in Figure 2. FDPN consists of three layers.

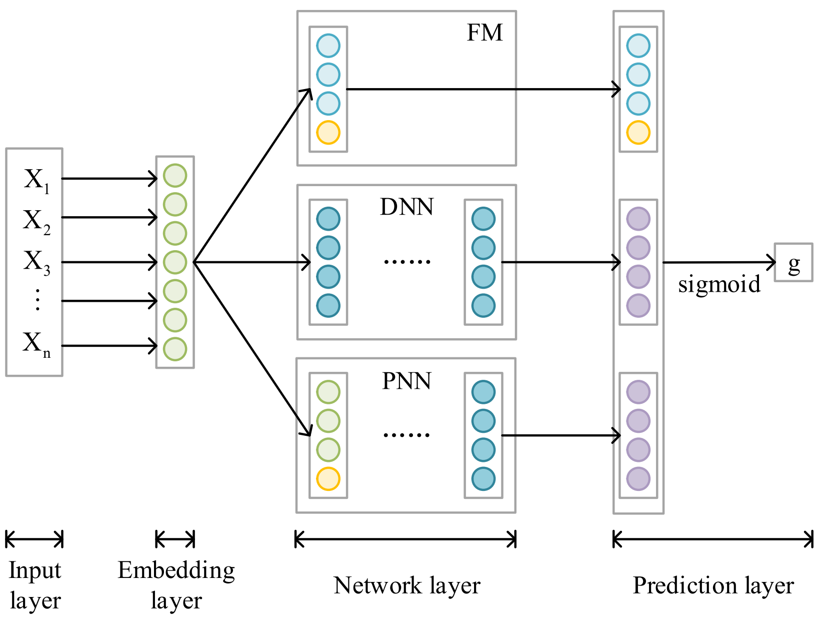

Figure 2: FDPN model framework

(1) Embedding layer

Since the original data is relatively sparse, it is reduced in dimension to obtain a low-dimensional representation of the features. Converting the initial features into low-dimensional vector representations makes the data relatively dense and reduces the computational workload. The structure of the embedding layer is shown in Figure 3.

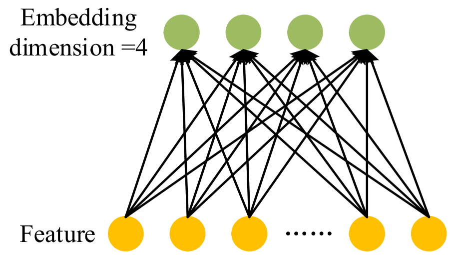

Figure 3: Embedded layer structure

The mapping output of the embedding layer can be expressed as:

$$a = \left[e_1, e_2, ..., e_p\right] \tag{1}$$

Among them, a represents the embedded features, e_i represents the i th embedded feature, and p refers to the number of embedded features, where $p \le (m+n+k)$.

(2) Network layer

a) Factorization Machine

Factorization Machine (FM) is used to learn feature interactions. Specifically, as shown in formula (2):

$$y_{fm} = w_0 + \sum_{i=1}^{p} w_i e_i + \sum_{i=1}^{p} \sum_{j=i+1}^{p} w_{ij} e_i e_j \tag{2}$$

Among them, w_0, w_i, w_{ij} are the weights of each feature, respectively. The factorization machine learns the first-order representation of features through logistic regression and learns second-order feature information through the dot product of vectors. The final output value y_{fm} of the FM layer is directly transmitted as input to some nodes in the prediction layer.

b) Deep Neural Network

A Deep Neural Network (DNN) is a deep neural network that can fit complex nonlinear features and learn the interaction information between higher-order features, i.e., it has stronger learning capabilities [21]. The output of the embedding layer is the input to the first hidden layer of the DNN, and the calculation formula for the first hidden layer is shown in Equation (3):

$$h_1 = f(W_0 e_p + b_0) \tag{3}$$

Assuming that the DNN part has a total of l hidden layers, whose output y_{dnn} goes directly to the input part of the prediction layer, the calculation formula for the final output value of the DNN network is shown in Equation (4):

$$y_{dnn} = f(W_{l-1} h_{l-1} + b_{l-1}) \tag{4}$$

Among them, $f(\cdot)$ is the activation function of the hidden layer, and the ReLU activation function is used.

c) Product Neural Network

The Product Neural Network (PNN) is a feedforward deep neural network that includes a Product layer. In PNN, the input information not only contains first-order feature-related information, but also second-order feature information. Therefore, the Product layer enriches the information input to the deep neural network. The calculation of the second-order feature is shown in Equation (5). where p denotes the inner product of the embedding layer feature vectors e_i and e_j:

$$p = \langle e_i \cdot e_j \rangle = \left[e_i^1 e_i^2 \cdots e_i^p\right] \begin{bmatrix} e_j^1 \\ e_j^2 \\ \vdots \\ e_j^p \end{bmatrix} = \sum_{i=1}^{p} \sum_{j=i+1}^{p} w_{ij} e_i e_j \tag{5}$$

The input vector of PNN is composed of the first-order feature vector output by the embedding layer and the second-order feature vector generated by their pairwise interaction. The calculation formula is formalized as in Equation (6):

$$x_{pnn} = [a; p] \tag{6}$$

The final output value y_{pnn} of the PNN is calculated as shown in Equation (6), differing from the DNN in that the input feature vectors from the embedding layer to the first hidden layer are different. The output value y_{pnn} of the last hidden layer node in the PNN is directly transmitted as input to a portion of the nodes in the prediction layer.

(3) Prediction Layer

The primary task of the prediction layer is to concatenate the low-order and high-order feature representations output by the FM, DNN, and PNN layers in the network and predict the target student's grade category. The purpose of feature fusion is to utilize these features more comprehensively and accurately to predict student grades.

This paper uses the concatenation method for feature fusion, directly concatenating the features y_{fm}, y_{dnn}, y_{pnn}. This process can be formally represented as:

$$f = \left[y_{fm}; y_{dnn}; y_{pnn} \right] \tag{7}$$

f is the final feature obtained after feature fusion of y_{fm}, y_{dnn}, y_{pnn}. Finally, the feature f is input into a perceptron with a sigmoid activation function to obtain the probability of the student's course grade category.

As can be seen from the above, FDPN consists of three parts: FM, DNN, and PNN. The final grade prediction result is obtained using formula (8):

$$g = Sigmoid(f) \tag{8}$$

II. B. 3) Loss function

This paper uses the cross-entropy loss function and the L_2 regularization parameter. The loss function of the model is as follows:

$$loss = -\frac{1}{n}\sum_{i=1}^{n} y_i \log g + \lambda \| \theta \|^2 \tag{9}$$

In this context, n denotes the total number of training data points, y_i represents the performance category of the i th data point, g denotes the predicted probability of the performance category for the i th data point, $\lambda \| \theta \|^2$ is the L_2 regularization term, and θ is the set of all model parameters.

II. C. Design of a student learning management system based on multiple sources of information

Based on the requirements for student academic performance management, this system builds a central platform that integrates consumption data, internet usage data, daily routine data, basic student data, mental health data, and other daily learning and living data. This platform supports data sharing, data analysis and mining, unified management, and analytical decision-making. Based on breakthroughs in key technologies, model construction, and visualization research, a system is developed for academic performance management analysis and evaluation in educational institutions. The system has high scalability and can be expanded with additional functional modules as business needs grow.

The core application functional modules of the system integration will include the following aspects. First, academic performance information collection and management, including online information (such as academic affairs information, consumption information, and access control information collection) and offline information (such as activity information and questionnaire information collection). Second, academic information preprocessing management. The collected academic information is preprocessed using data cleaning, merging, transformation, and reduction methods to achieve data classification and integration, retaining key information for further data mining. Third, academic information analysis and evaluation management. FDPN and clustering analysis techniques are used to analyze and process academic data, and evaluations and recommendations are made based on the results.

III. Model testing and system application effect analysis

This chapter conducts performance testing on a grade prediction model based on feature combinations and conducts a questionnaire survey on the smart teaching model under a student learning management system based on multi-source information, comparing it with the traditional smart teaching model to analyze the system's scalability.

III. A. Model testing

III. A. 1) Evaluation Indicators

This paper uses accuracy, precision, recall, F1, and other metrics to evaluate the predictive performance of the model:

$$Accuarcy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \tag{13}$$

The meanings of TP, FP, FN, and TN are shown in Table 4.

Table 4: Confusion matrix

True situation	Prediction	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

III. A. 2) Analysis of Results

To validate the effectiveness of the proposed method, the proposed FDPN was compared with four traditional machine learning classification prediction methods—Support Vector Machine (SVM), Logistic Regression (LR), Gaussian Naive Bayes (GaussionNB), and Decision Tree—in two publicly available educational datasets: the English grade dataset and the mathematics grade dataset, both from the “student performance” dataset. The experiments were conducted to verify the effectiveness of the proposed method.

The prediction results on the English dataset are shown in Figure 4. In the English score dataset, among the four traditional classification prediction algorithms—support vector machines, logistic regression, Gaussian Naive Bayes, and decision trees—the decision tree achieved the best classification prediction performance, with a prediction accuracy rate of 92.7%, while the support vector machine achieved the highest prediction precision, reaching 89.9%. Overall, the decision tree performed the best, followed by the support vector machine, with both significantly outperforming the Gaussian Naive Bayes and logistic regression algorithms on this dataset. The proposed feature combination-based student performance prediction method achieved better results than traditional classification methods across four evaluation metrics: Accuracy, Precision, Recall, and F1-Measure, with a prediction accuracy rate of 92.8%. It can be concluded that the proposed feature combination-based student performance prediction method can improve the accuracy of predicting students' final English grades.

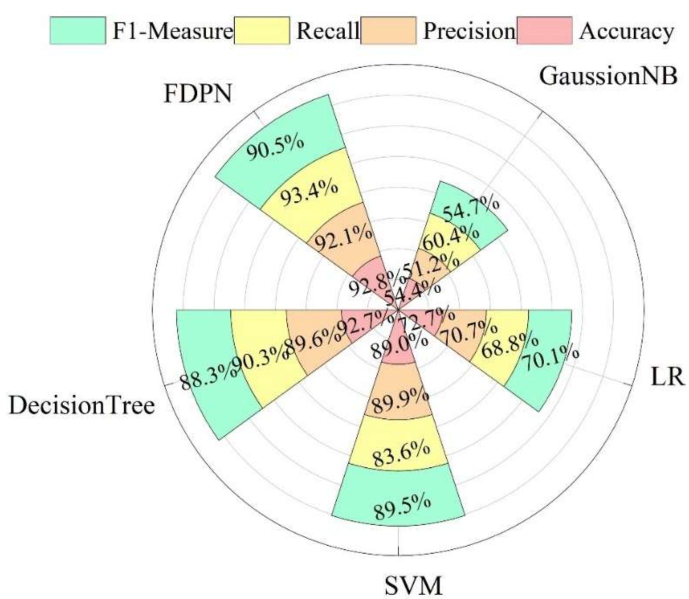

Figure 4: Prediction of the English data set

The prediction results on the mathematics dataset are shown in Figure 5. The experimental results of different methods on the mathematics performance dataset indicate that: Among the four traditional classification prediction algorithms—support vector machines, logistic regression, Gaussian naive Bayes, and decision trees—the decision tree demonstrates a significant advantage in all four measurement metrics: prediction accuracy, precision, recall, and F1-Measure. It achieves the best classification prediction performance, with a prediction accuracy rate of 92.9%, far exceeding that of support vector machines, logistic regression, and Gaussian naive Bayes. The proposed student performance prediction method based on feature combination outperforms the four traditional methods in terms of four evaluation metrics: Accuracy, Precision, Recall, and F1-Measure. with a prediction accuracy rate of up to 93.2%. Compared to the four traditional methods, the accuracy rate of the relatively poorer-performing Gaussian Naive Bayes method was improved by 31.1%, and the prediction accuracy rate of the best-performing decision tree among the four traditional methods was improved by 0.3%. Precision, recall, and F1-Measure were improved by 1.3%, 2.9%, and 0.7%, respectively.

In summary, the proposed feature combination-based grade prediction method can effectively improve the prediction accuracy of students' final mathematics grades.

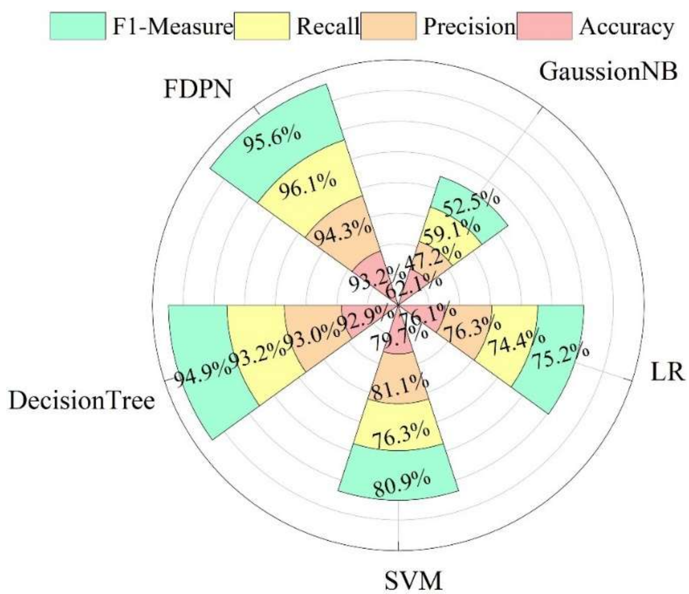

Figure 5: Prediction of the Math data set

The proposed student performance prediction model based on feature combination achieves higher prediction accuracy than four traditional machine learning methods. The reason for this may be that traditional methods do not extract more feature information specific to the performance target, but instead directly input each attribute feature as a classification feature into the model for learning and training. In this process, the influence of each attribute feature on the final performance score is treated equally. By introducing feature combinations, the model can fully learn the relationship information between each feature and academic performance. Compared to using a single feature for learning, this approach enhances the model's predictive effectiveness, thereby significantly improving its predictive capabilities.

III. B. Analysis of the Application Effectiveness of the Student Information Management System

The study subjects were students from Class 1 and Class 2 of the Accounting Program at School A in the 2024 cohort, totaling 120 participants. Class (1) incorporated the student learning management system designed in this paper, which utilizes multi-source data fusion technology, into the smart teaching process, while Class (2) adopted the traditional smart teaching model. After a 12-week comparative teaching period, a professional course examination was conducted. Daily behavioral data of students from both classes were collected, and the FDPN model was used for student performance prediction. A questionnaire was designed to assess students' evaluations of the new and traditional teaching models. Finally, an independent samples t-test was performed on the performance of the two classes.

III. B. 1) Analysis of predicted results

The comparison between the predicted grades based on the FDPN model and the actual course grades is shown in Figure 6. The horizontal axis represents the individual student ID numbers, and the vertical axis represents the grade values. The trends in the two sets of data are consistent, and the specific data values differ only slightly, indicating that course grade prediction based on multi-source data is relatively accurate.

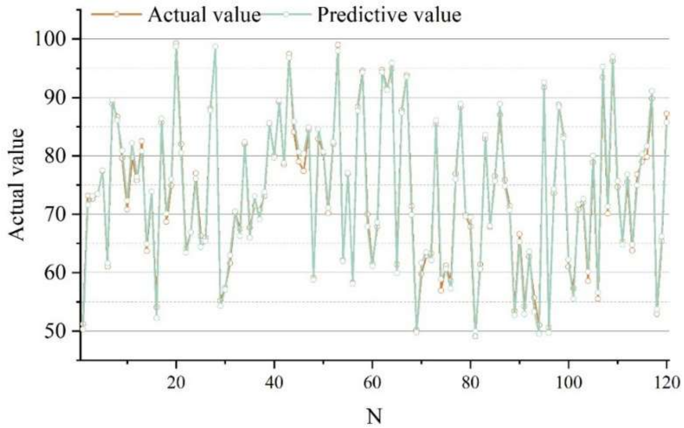

Figure 6: Results distribution

III. B. 2) Analysis of Questionnaire Survey Results

Before distributing the questionnaire, this paper first tested its reliability, primarily using Cronbach's alpha reliability index for analysis. After analysis, it was found that the questionnaire coefficient was 0.933, passing the validity test, indicating that the questionnaire had good reliability, and the questionnaire results could truly reflect the expected objectives, with the collected data having analytical value.

The questionnaire primarily consists of three sections: students' acceptance of the new teaching model, students' recognition of the classroom learning process, and students' classroom participation and evaluation of learning outcomes.

(1) Students' acceptance of the new teaching model

This section primarily examines whether students recognize and accept the new teaching model. Statistical analysis revealed that in the basic accounting course, the new teaching model was favored and recognized by 79% of students, with only 3.2% expressing resistance. When compared to traditional smart classrooms, 81% of students believed that the student learning management system could overcome the shortcomings of traditional smart classrooms and address the teaching issues present in traditional smart classrooms. Regarding the use of the student learning management system, 75% of students demonstrated a high level of acceptance.

(2) Students' recognition of the classroom learning process is shown in Figure 7. The questions included: "Very interesting" (T1), "Improved learning efficiency" (T2), "Enhanced teacher-student communication" (T3), and "Improved my pre-class preparation and review effectiveness" (T4). During classroom instruction, 80% of students found the new teaching model to be relaxed and enjoyable, not only enhancing teacher-student interaction but also improving students' sense of learning efficacy. Regarding the use of the Student Learning Management System as a teaching tool, approximately 82.5% of students believe it enhances communication between teachers and students, indicating that the system improves upon the one-way output model of traditional teaching.

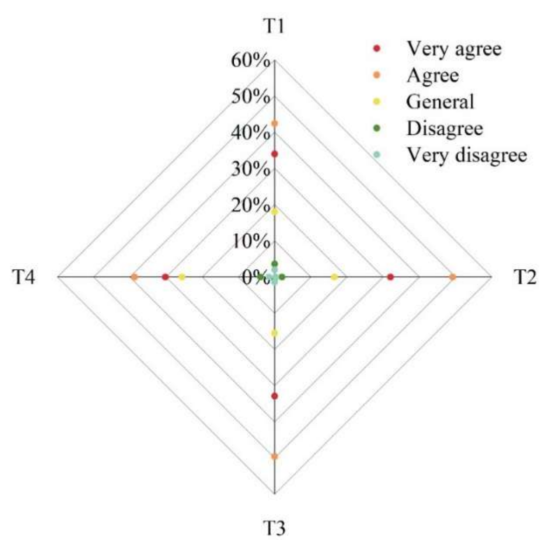

Figure 7: The recognition of the learning process

(3) The evaluation of classroom participation and learning outcomes for Class 8 students is shown in Figure 8. The questions included: can enhance practical skills (T5), can stimulate innovative thinking (T6), can improve independent learning abilities (T7), and can improve problem-solving abilities (T8). Respectively, 76.6% and 72.2% of students believed that the new approach played a positive role in improving independent learning abilities and problem-solving abilities.

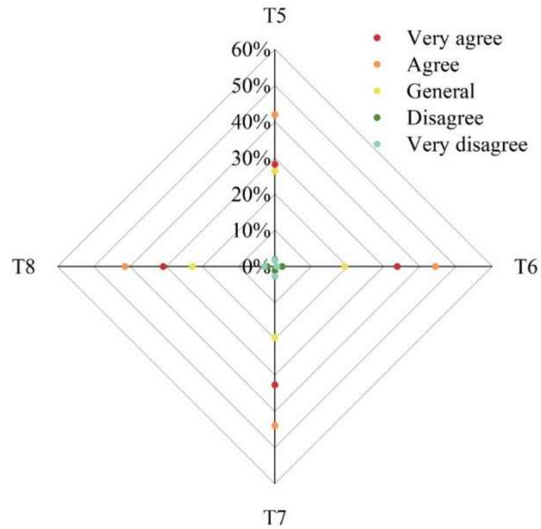

Figure 8: Student participation and learning effect evaluation statistics

III. B. 3) Comparative Analysis with Traditional Wisdom Teaching

To more intuitively analyze the differences in teaching effectiveness between the new teaching model and the traditional smart classroom teaching model, an independent samples t-test was conducted on the professional course grades of the two classes. The variables such as the instructors and the semester of the two classes (Class 1 and Class 2) were consistent, meeting the experimental conditions. The test results are shown in Table 5. Upon examining the academic performance of the two groups of classes, it was found that there was a significant difference in the professional course grades between Class (1) and Class (2). Specifically, the mean scores for Class (1) and Class (2) were 87.06 and 80.36, respectively, with Class (1) achieving significantly better results than Class (2). This indicates that the new teaching model has a significant effect in improving student performance.

Table 5: Independent sample t test results

		Levene test		T test					
		F	Sig	t	Sig	MD	SE	95% confidence interval	
								Lower limit	Upper limit
Grade	Equal variance assumed	1.483	0.309	6.214	0	5.998	0.934	3.981	7.763
	Equal variance not assumed	-	-	6.237	0	5.975	0.933	3.633	7.752

IV. Intelligent teaching services based on multi-source data fusion technology

The aforementioned student academic performance management system was designed based on multi-source data fusion technology, promoting the development of smart classrooms. In addition, the application of multi-source data fusion technology in personalized learning services, efficient teaching services, and classroom management services is also worth promoting.

IV. A. Personalized learning services based on multi-source data fusion technology

By collecting and analyzing learners' multimodal data, it is possible to accurately evaluate their cognition, behavior, emotions, etc., promote connections between learners and elements of the educational context, clarify the mechanisms of learning, and thereby adjust relevant elements of the teaching context, change learning paths, and provide learners with "customized" learning support.

IV. B. Efficient teaching services based on multi-source data fusion technology

By collecting and analyzing data throughout the entire learning process, teachers can fully understand students' learning characteristics and cognitive levels, thereby adjusting teaching strategies, changing teaching models, and improving classroom teaching effectiveness. At the same time, based on the generated classroom records and teacher growth curves, teachers' teaching ability changes can be tracked in a timely manner, providing a basis for teachers' self-reflection and improvement of professional teaching levels.

IV. C. Classroom management services based on multi-source data fusion technology

By collecting classroom teaching behavior data from "multiple spaces, multiple subjects, and multiple stages" using smart devices, a unified classroom activity arrangement model can be established, which to a certain extent meets the structured requirements of data-driven precision teaching activities and promotes the efficient management of classroom teaching [22]. Additionally, multi-source data fusion technology can comprehensively summarize learners' cognition, thinking, abilities, and emotions, helping teachers accurately grasp students' learning situations and states, and provide targeted instruction to improve the quality of classroom teaching management services.

In summary, multi-source data fusion technology brings new concepts and evaluation methods to classroom teaching, including presenting phenomena that are difficult to characterize or observe visually in a data-driven form, thereby promoting scientific and effective teaching decisions and classroom management.

V. Conclusion

Based on the current state of smart classroom construction, this study identifies the key elements and design principles for the development of smart classrooms, designs a student learning management system, and provides references for the construction and upgrading of smart classrooms.

A comparative experiment was conducted using publicly available educational data on student performance to verify the effectiveness of a student performance prediction model based on feature combinations. The accuracy rate of the student performance prediction model based on feature combinations can reach over 90%, outperforming the predictive performance of four traditional machine learning classification prediction models. After the model was integrated into the student academic performance management system and put into use, student performance improved significantly compared to the traditional model. By utilizing multi-source information fusion technology, it is possible to deeply mine various academic performance data of students on campus, which is conducive to promoting the common development of schools and students. Additionally, the application of this technology in personalized learning services, efficient teaching services, and classroom management services is also worth promoting.

References

[1] Ren, F. (2024). The construction and application of smart classrooms in China: A literature review based on 93 studies. Science Insights Education Frontiers, 23(1), 3705-3724.

[2] Phoong, S. Y., Phoong, S. W., Moghavvemi, S., & Sulaiman, A. (2019). Effect of smart classroom on student achievement at higher education. Journal of Educational Technology Systems, 48(2), 291-304.

[3] Tuo, P., Bicakci, M., Ziegler, A., & Zhang, B. (2025). Measuring Personalized Learning in the Smart Classroom Learning Environment: Development and Validation of an Instrument. Education Sciences, 15(5), 620.

[4] Ojo, O., Kareem, M. K., Odunuyi, S., & UGWUNNA, C. (2022). An internet-of-things based real-time monitoring system for smart classroom. Journal of the Nigerian Society of Physical Sciences, 297-309.

[5] Rong, J. (2022). Innovative research on intelligent classroom teaching mode in the "5G" era. Mobile information systems, 2022(1), 9297314.

[6] Dai, Z., Zhang, Q., Zhao, L., Zhu, X., & Zhou, D. (2023). Cloud-edge computing technology-based internet of things system for smart classroom environment. International Journal of Emerging Technologies in Learning (iJET), 18(8), 79-96.

[7] Alassery, F. (2019, March). A smart classroom of wireless sensor networks for students time attendance system. In 2019 IEEE Integrated STEM Education Conference (ISEC) (pp. 324-331). IEEE.

[8] Jiang, D. (2022). Research on remote monitoring method of smart classroom based on internet of things. International journal of autonomous and adaptive communications systems, 15(3), 220-234.

[9] Kaur, A., & Bhatia, M. (2022). Smart classroom: A review and research agenda. IEEE Transactions on Engineering Management, 71, 2430-2446.

[10] Wang, Z., Li, L., Zeng, C., & Yao, J. (2023). Student learning behavior recognition incorporating data augmentation with learning feature representation in smart classrooms. Sensors, 23(19), 8190.

[11] Li, Y., & Chen, L. (2020). Improved LSTM data analysis system for IoT-based smart classroom. Journal of Intelligent & Fuzzy Systems, 39(4), 5141-5148.

[12] Savov, T., Terzieva, V., Todorova, K., & Kademova-Katzarova, P. (2019, September). Smart classroom, internet of things and personalized teaching. In CBU International Conference Proceedings (Vol. 7, pp. 1001-1007).

[13] Hou, G. (2024). Correlation among teacher ICT teaching, teacher immediacy behaviors, teacher–student rapport, and student engagement in smart classroom teaching. Sustainability, 16(21), 9592.

[14] Liu, W., Zhang, C., Yu, B., & Li, Y. (2019, February). A general multi-source data fusion framework. In Proceedings of the 2019 11th international conference on machine learning and computing (pp. 285-289).

[15] Sun, R., & Ren, Y. (2024). A multi-source heterogeneous data fusion method for intelligent systems in the Internet of Things. Intelligent Systems with Applications, 23, 200424.

[16] Cai, Y. (2021). Research on data fusion method of multi-source complex system. Journal of web engineering, 20(5), 1553-1572.

[17] Qin, C., Guo, B., Zhang, Y., & Shen, Y. (2023). The dynamic fusion representation of multi-source fuzzy data: C. Qin et al. Applied Intelligence, 53(22), 27226-27248.
[18] Liu, J., Huang, J., Sun, R., Yu, H., & Xiao, R. (2020). Data fusion for multi-source sensors using GA-PSO-BP neural network. IEEE Transactions on Intelligent Transportation Systems, 22(10), 6583-6598.
[19] Fengjie Ren. (2024). The Construction and Application of Smart Classrooms in China: A Literature Review Based on 93 Studies. Science Insights Education Frontiers,23(1),
[20] Dongfang Zhu. (2025). Knowledge creation in vocational education using multi-source data fusion under big data environment. International Journal of Information and Communication Technology,26(5),69-83.
[21] Mahmoud Abdasalam,Ahmad Alzubi & Kolawole Iyiola. (2024). Student grade prediction for effective learning approaches using the optimized ensemble deep neural network. Education and Information Technologies,30(8),1-25.
[22] Zhang Fangfang & Liu Qiang. (2024). Innovative Analysis of Student Management Path Based on Artificial Intelligence and Big Data Integration. International Journal of e-Collaboration (IJeC),20(1),1-20.