Exploration and practice of data-driven student evaluation incorporating information technology

Wenjing Li[a*], Nan Wen[b], Jie Wang[c], Xiao Li[d]

Paris Curie Engineer School, Beijing University of Chemical Technology, Beijing, China

[a*]liwenjing@buct.edu.cn, [b]wennan@mail.buct.edu.cn
[c]wjcp@mail.buct.edu.cn, [d]lixiao1992@mail.buct.edu.cn

Abstract. With the continuous development of emerging technologies such as the Internet and big data, it is a trend for universities to use information technology for teaching. Blended learning relies on the advantages of the Internet and combines various types of data from the big data era with face-to-face offline teaching, which can effectively collect various types of data during students' learning process. By collecting, processing, and analyzing this data, it is possible to timely and effectively track students' learning progress. In this paper, information technology is introduced into various stages of pre-class, in-class, and post-class in blended learning practice. It records, collects, and analyzes students' learning and testing data from different platforms such as Youmuke Education Platform, Yuketang, and Microsoft Forms. These diverse data are integrated and utilized, and process-based assessment methods based on students' learning data are adopted. SPSS is used to analyze students' performance in different projects, and a model for analyzing students' exam scores is established based on data, predicting and warning students' final exam scores. By analyzing students' performance in various aspects during the learning process, a comprehensive understanding of students' learning situations is achieved, and targeted measures are taken to improve the quality of teaching.

Keywords: Information technology; Learning data; Process assessment; Modeling.

1 Introduction

Educational big data is the transformation pathway for revolutionizing education through abundant resources and scientific power. "China's Education Modernization 2035" proposes the construction of intelligent campuses, integrating the development of integrated intelligent teaching, management, and service platforms [1]. The Minister of Education, Huai Jinpeng, pointed out that education informatization should be regarded as a strategic high ground for development, using it to promote high-quality education and lead educational modernization [2].

M. Yu et al. (eds.), *Proceedings of the 2024 5th International Conference on Big Data and Informatization Education (ICBDIE 2024)*, Advances in Intelligent Systems Research 182,
https://doi.org/10.2991/978-94-6463-417-4_16

In recent years, universities nationwide have been actively implementing digital education strategies and developing “Internet + Education”. Based on modern information technologies such as 5G, artificial intelligence, and big data, they are constructing digital, intelligent, and comprehensive management and service platforms to promote the establishment of an information-based education system. Blended learning, relying on the advantages of the internet, combines various types of data from the era of big data with face-to-face offline teaching. This enables more effective collection of various types of data during students’ learning processes. Teachers can use the collection, processing, and analysis of this data to timely and effectively monitor students’ learning progress, predict and warn about their final exam scores, discover problems in advance, intervene in a timely manner, and enhance guidance, thereby improving students’ grades and learning outcomes.

This article introduces information technology into various stages of pre-class, in-class, and post-class in the practice of blended learning. It records, collects, and analyzes student learning and testing data from different platforms such as Youmuke Platform, Yuketang, and Microsoft Forms. Through a data-driven process-oriented assessment approach, it aimes to stimulate students’ learning enthusiasm, ignites their intrinsic motivation for active learning [3], to comprehensively understands students’ learning situations through the analysis of various performance indicators, and takes targeted measures to improve teaching quality. It mines, integrates, and utilizes the diverse data, adopts a process-oriented assessment approach based on student learning data, and establishes a model for analyzing students’ exam scores based on the data to predict and warn about their final exam scores.

2 Students Learning Data Sources

2.1 Youmuke Platform

The content of online course development on the Youmuke online education platform includes basic course information, unit learning, course resources, course activities, and in-class teaching. Through this development, teachers can upload excellent MOOC videos or their own recorded short course videos, exercises, and tests for students to access and learn. They can also implement in-class teaching activities and assign course assignments. Students can browse course resources repeatedly and complete course assignments.

Teachers can use the platform’s “course statistics” function to understand the duration and number of students accessing the course (Fig. 1). By using “resource access” and “view access details,” they can identify students who have not studied the course resources (Fig. 2). The frequency of resource access can be used to evaluate the difficulty of course resources and the attractiveness of course content. By using “assignment score distribution” and “assignment submission rate,” (Fig. 3) teachers can grasp the completion status of students’ assignments and identify outliers, allowing them to promptly discover the list of students who have not completed their assignments and provide timely supervision and guidance.

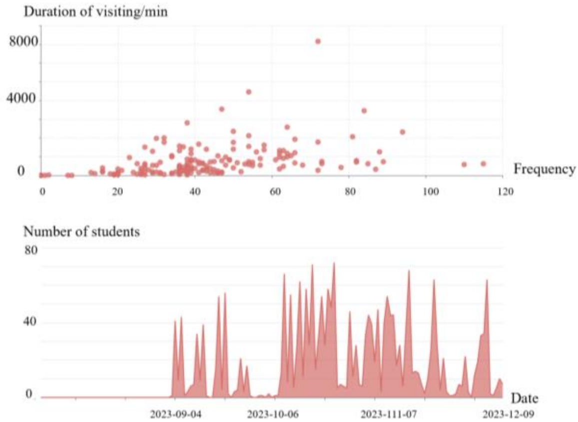


Fig. 1. Duration and number of students accessing the course

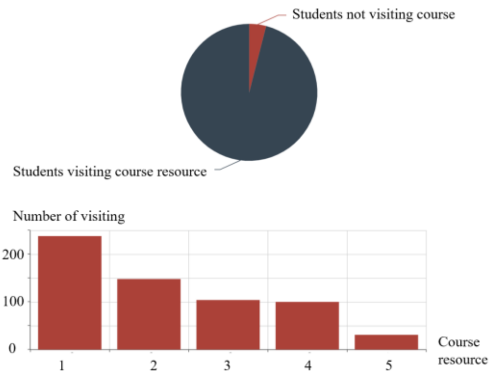


Fig. 2. Resource access data and requency of resource access

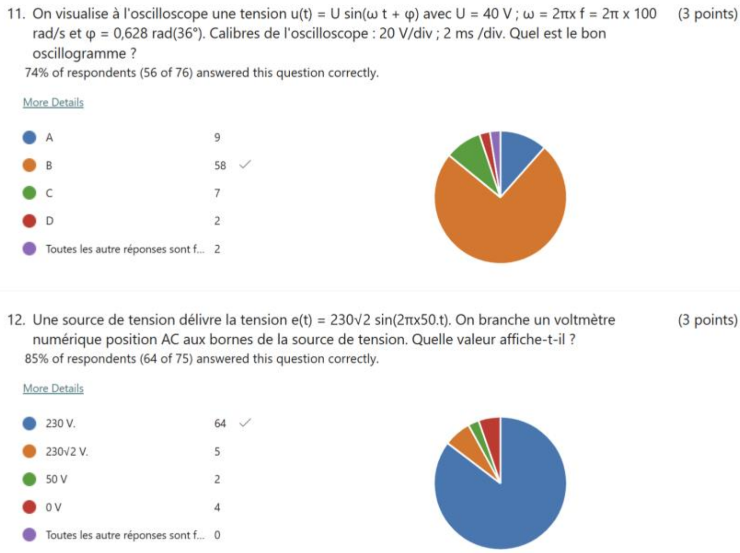
Fig. 3. Assignment score distribution and assignment submission rate

2.2 Yuketang

Yuketang is mainly used for in-class activities. It utilizes WeChat mini-program to enable teachers to push teaching resources, conduct classroom attendance, administer pre-tests, mid-tests, and post-tests, as well as provide features such as bullet screen comments and submissions. Yuketang can comprehensively track students' learning progress and data throughout the entire process. In addition to enhancing the classroom atmosphere, the introduction of Yuketang allows teachers to promptly understand the information of "unattended" students and promptly communicate with them to resolve attendance issues. It also enables teachers to assess students' understanding of classroom knowledge in real-time through tests. These student learning data serve as objective data support for teaching evaluations.

2.3 Microsoft Forms

Microsoft Forms offers a quiz format questionnaire where teachers can pre-set the point values for each question. After the quiz is submitted, the grades can be directly generated. The answers can be accompanied by explanations for each option. Users can check the explanations for the correct and incorrect answers after submission. For timed exams, time restrictions can be added in the settings to require completion within a specified time frame. It is a great tool for conducting formative assessments. The test results will display the distribution of options for each question, allowing teachers to understand students' learning progress and identify misconceptions (Fig. 4). After the test, the student's grade information can be directly saved as a subgrade for ongoing assessment.



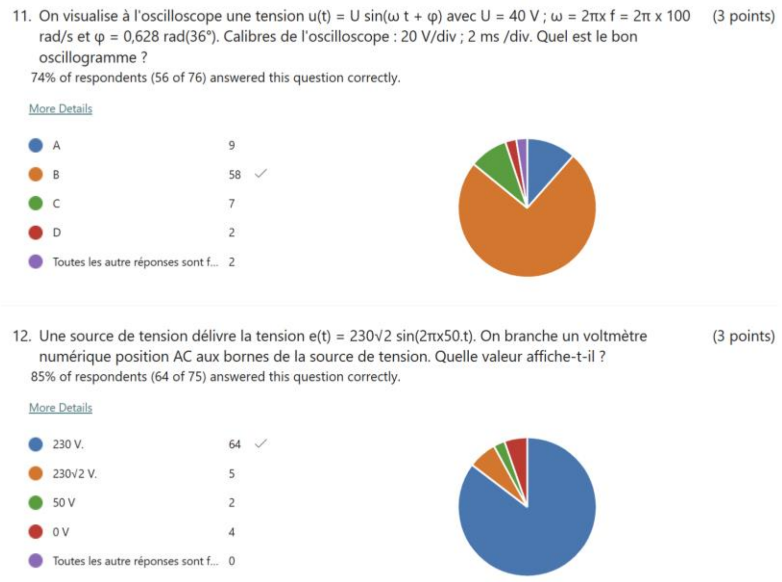


Fig. 4. Test results displaying the distribution of options for each question by Microsoft Forms

2.4 Mid-term and final exams data

Mid-term and final exams for students are conducted offline in traditional mode. The exam scores are recorded and organized, and the difficulty and discrimination of the mid-term and final exam questions are evaluated. This helps to understand the areas where students are weak and their grasp of different knowledge points, as well as examining the characteristics and rationality of the exam questions.

Difficulty and discrimination analysis is conducted. The difficulty coefficient is used to evaluate the difficulty of each question. The difficulty coefficient is calculated by dividing the score rate of each question by the maximum score. A higher difficulty coefficient indicates an easier question. Generally, the difficulty coefficient range and the judgment criteria for exam difficulty are as follows: 0.6 - 0.8 is considered easy, 0.3 - 0.6 is considered difficult. If the questions or the exam paper is too difficult or too easy, it will affect the reliability of the exam and fail to achieve the purpose of the exam. According to the data in the table below, the average difficulty coefficient is 0.545, indicating that the overall difficulty of the exam paper is high. The third question has the lowest difficulty coefficient, making it the most difficult question. By considering the knowledge points covered in the questions, it can help teachers understand the difficulties in student learning and provide more targeted explanations in the classroom.

The discrimination of the exam questions is evaluated to differentiate between high-level and low-level students. A low discrimination value indicates that the exam paper cannot assess the students' level effectively. A higher discrimination index value indicates better differentiation, shown in Table I.

Table 1. Test Discrimination indication

Discrimination Index	Effect	Appropriate Use
0.4-1	Excellent discrimination	Selection
0.2-0.4	Fair discrimination	Inspection
0-0.2	Poor discrimination	None
-1-0	Issues with the test	None

The methods used to calculate the discrimination index vary, and the method used in this paper is the extreme group method. The specific calculation steps are as follows:

(1) Arrange all student scores in descending order.

(2) The top 20 scores are considered the high group, and the bottom 20 scores are considered the low group.

(3) Determine the passing rate for each question by considering scores above 60% of the maximum score as passing and scores below as failing. Calculate the passing rate (H) for the high group and the passing rate (L) for the low group.

(4) Calculate the discrimination index (D) as D = H - L. The results of the calculation are shown in the table below. Table II is an example which summarizes the difficulty and discrimination index for the mid-term exam of Course1. The average discrimination index is 0.69, indicating that the discrimination of this test is very good and meets the standards of a selective exam. The second question has a discrimination index of 0.4 and serves as a foundational question for evaluating understanding. The first and fourth questions have high difficulty and high discrimination index, making them suitable for selecting high-performing students.

Table 2. difficulty and discrimination index for the mid-term exam of Course1

Question number	1	2	3	4	5
Full marks	30	15	15	20	20
Average score	19.437	7.736	6.655	12.782	9.575
Difficulty coefficient	0.648	0.516	0.444	0.639	0.479
Discrimination index	0.9	0.4	0.65	0.9	0.6
Average difficulty	0.545				
Average discrimination index	0.69				

3 Process-oriented assessment

Process evaluation covers the entire process of students' pre-class, in-class, and post-class activities. By mining students' learning process data, teachers can continuously improve teaching methods and enhance teaching effectiveness. Process evaluation provides continuous assessment of students' learning process and timely feedback to students [4]. Well-performing students can gain a sense of individual achievement and have their individual needs met, which motivates them to learn more actively and increases their engagement in learning. For students who are not performing well, timely and continuous evaluation helps them identify the gaps, recognize their weaknesses, and make timely adjustments to their learning plans and methods.

For teachers, process evaluation integrates teaching and assessment. Throughout the teaching process, teachers can evaluate and monitor students' learning outcomes in a timely manner. This real-time interactive feedback function is beneficial for teachers to continuously adjust and optimize their teaching plans and methods, thus improving the quality of teaching. In this study, blended learning was used to incorporate the use of information technology. Comprehensive

Table 3. Process Assessment of Blended Learning with the Integration of Information Technology

Assessment Type	Assessment Indicators	Assessment Content	Information Technology Tools/ Platforms	Weight
Continuous Process Assessment	Pre-class Preparation	Accessing course resource	Youmuke	5%
	Class Participation	Attendance, online quiz	Yuketang	10%
	Online homework	Homework quality	Youmuke	10%
	Section Tests	Understanding of chapter content	Microsoft Forms	10%
Mid-term Exam	Exam Score	Subjective questions	N/A	15%
Final Exam	Exam Score	Subjective questions	N/A	50%

collection of students' learning data was conducted, and process evaluation based on students' learning data was implemented for both Course 1 and Course 2. The types, indicators, contents, information technology platforms utilized, and respective weights for assessment are presented in Table III.

4 Grades tracking and analysis

Based on the above formative assessment indicators, the learning data from different sources will be integrated to obtain students' usual grades. Then, a deep analysis will be conducted on the usual grades of Course 1, the mid-term exam grades of Course 1, the final exam grades of Course 1, the usual grades of Course 2, and the final exam grades of Course 2. The statistical characteristics of these grades will be analyzed using SPSS software [5, 6]. The correlation between each grade will be examined, and a model will be established to predict the success or failure in subsequent exams based on the results of previous grades.

4.1 Basic statistical information of each score

SPSS is utilized to have a descriptive analysis and the normality test; the results are shown in Table IV. Based on the average scores of each category, it can be observed that the usual assessments have higher scores compared to the mid-term and final exam scores. This suggests that the exam questions were generally difficult, and the mid-term exam was slightly more challenging than the final exams for Course 1 and Course 2. The usual assessment scores for Course 1 are higher than those for Course 2, indicating a potential decline in student effort during the learning phase of Course 2.

Through the average scores of each category, it can be observed that the usual grades are higher than the mid-term and final exam scores, indicating that the exam questions are generally difficult. Additionally, the difficulty of the mid-term exam is slightly higher than that of the final exam for Course 1 and Course 2. The usual grades for Course 1 are higher than those for Course 2, suggesting that students may have been less diligent in their studies during the second course.

The standard deviation of the mid-term exam scores is smaller than that of the final exam scores, indicating that the final exams have a greater ability to differentiate scores. Furthermore, based on the negative skewness of the usual grades, mid-term exam scores, and final exam scores, it can be concluded that overall, the scores are skewed towards higher scores, with the usual grades showing a greater negative skewness compared to the mid-term and final exam scores, indicating that usual grades are easier to obtain.

In terms of kurtosis, the usual grades for Course 1 exhibit a high peakedness, while the final exam scores for Course 1 show a low peakedness. This suggests that the usual grades for Course 1 are more concentrated, indicating that there is not much disparity in the usual grades. This conclusion is consistent with the findings from the standard deviation (12.762 vs 18.560).

Table 4. Descriptive Analysis

Items	**N**	**Mean**	**Std.**	**Skewness**	**Kurtosis**
Usual performance-Course1	87	78.586	12.762	-2.172	6.492
Mid-term exam-Course1	87	56.184	15.949	-0.524	0.373
Final exam-Course1	87	62.046	18.560	-0.333	-0.139
Usual performance-Course2	87	72.609	19.590	-1.264	1.337
Final exam-Course2	87	60.632	17.968	-0.792	1.417

4.2 Pearson correlations analysis

To analyze the correlation between various scores, we use Pearson Correlations for testing [7], and the judgment method is shown in the table below. When the Pearson Correlation coefficient r value is ±1, it is called a perfect positive (negative) correlation, which is basically nonexistent in social research. The closer this number is to ±1, the stronger the correlation. When r = 0, there is no relationship between the two variables.

The results of Pearson correlation analysis for the correlation between the five scores using SPSS are shown in the Fig. 5. According to the Pearson correlation analysis, all scores are positively correlated, with r values ranging from 0.39 to 0.68. The correlation coefficient between the mid-term scores of Course 1 and the mid-term scores of Course 2 is 0.68, indicating that students' behavior has continuity. Most students are serious about the learning process in one course and also in the learning process of another course. The correlation coefficient between the Course 2 regular scores and the Course 1 mid-term scores is 0.39, which is lower than the correlation coefficient between the Course 2 regular scores and the Course 2 final scores (0.52).

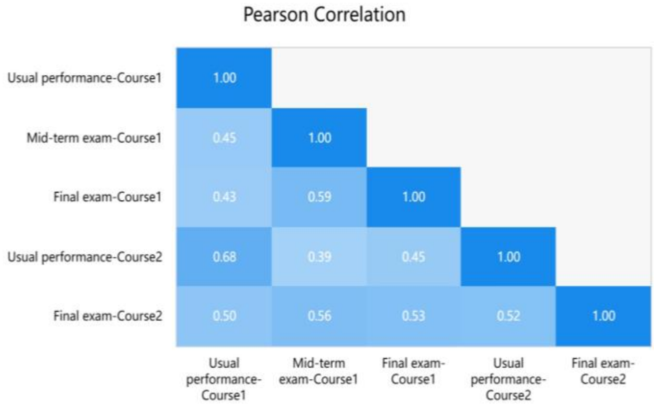


Fig. 5. The results of Pearson correlation analysis for the correlation between the five scores

4.3 Scores prediction based on naive Bayes model using SPSS

First, mark the scores above 60 in the Final exam-Course2 as Pass and the scores below 60 as Fail. Then, use Usual performance-Course1, Mid-term exam-Course1, Final exam-Course1, Usual performance-Course2 as independent variables, and Pass or fail course2 as the dependent variable. Set the training set ratio to 0.8 and analyze 87 samples using a naive Bayes model with a feature distribution of Gaussian distribution for Bayesian modeling [8].

Among them, accuracy is the proportion of correctly predicted samples to the total samples, and higher accuracy is better. Precision is the proportion of actual positive samples among the predicted positive results, and higher precision is better. Recall is the proportion of predicted positive samples among the actual positive samples, and higher recall is better. F1-score is a comprehensive evaluation metric that combines precision and recall, as it is the harmonic mean of precision and recall. Precision and recall are both better when higher, but they are often contradictory [9]. Therefore, F1-score is commonly used to evaluate the overall performance of a classifier [10]. Its value ranges from 0 to 1, where closer to 1 indicates better performance.

Table 5. Evaluation Results of naive Bayes model

Precision	Recall	F1-score	Sample	Count
0	0.63	0.83	0.71	6
1	0.90	0.75	0.82	12
Accuracy			0.78	18
Average	0.76	0.79	0.77	18
Average (Overall)	0.81	0.78	0.78	18

From the above Table V, we can see that the final model achieved an accuracy of 77.8%, a precision (composite) of 80.83%, a recall (composite) of 77.78%, and an f1-score (composite) of 0.78 on the test set. Based on the predicted results of the model, we can use the students' previous performance, mid-term exam scores, and final exam scores in relevant subjects to predict whether a student will pass the final exam, allowing us to identify students who are likely to fail in advance.

5 Conclusion

This study introduced information technology tools in the process of blended learning, mining, integrating, and analyzing students' learning data on various platforms throughout the learning process. Based on the students' learning data, a model for analyzing exam scores was established to predict and warn about students' future learning performance. Targeted measures were taken to improve the quality of teaching. However, there are still some issues, such as the lack of intelligence in the integration process due to multiple data sources, requiring manual extraction and integration. In the future, it is necessary to consider establishing an intelligent model for integrating data from multiple platforms to promote the deep implementation of information technology-assisted education.

Acknowledgement

2023 Education Reform Project of Beijing University of Chemical Technology "Construction and Practice of Internationalized Excellent Engineering Talent Training System based on the CTI Certification"

References

1. "China's Education Modernization 2035" [N]. People's Daily, 2019-02-24 (1).
2. Huai Jinping: Taking Education Informatization as a Strategic High Ground [J]. China Education Network, 2022 (Z1): 2.
3. Tinto V. Completing College: Re-thinking Institutional Action M. Chicago: The University of Chicago Press, 2012.
4. Soto-Estrada; Wellens; Gómez-Lizarazo, Student course evaluation: a process-based approach, Journal, [J] Australasian Journal of Engineering Education. Volume 23, Issue 2. 2018. PP 83-94.
5. Wang Xuehua. Management Statistics: Application Based on SPSS Software [M]. Beijing: Electronic Industry Press. 2011: 175-176.
6. Chen Rong, Analysis and Prediction of the Final Grade of Securities and Futures Based on SPSS. Journal, [J] Journal of Trend in Scientific Research and Development. Volume 4, Issue 4. 2020.
7. Dufera Abdisa G.; Liu Tiantian; Xu Jin, Regression models of Pearson correlation coefficient, Journal, [J] Statistical Theory and Related Fields. Volume 7, Issue 2. 2023. PP 97-106.
8. Qi Sheng et al. Comparison of machine learning approaches for radioisotope identification using NaI(TI) gamma-ray spectrum, Journal | [J] Applied Radiation and Isotopes. Volume 186, Issue. 2022. PP 110212-110212.
9. Fantinato Marcelo; Peres Sarajane Marques; Reijers Hajo A. X-Processes: Process model discovery with the best balance among fitness, precision, simplicity, and generalization through a genetic algorithm, Journal [J] Information Systems. Volume 119, Issue. 2023.
10. Cho Won Ki et al. Diagnostic Accuracies of Laryngeal Diseases Using a Convolutional Neural Network-Based Image Classification System, Journal | [J] The Laryngoscope. Volume 131, Issue 11. 2021. PP 2558-2566.