Research and Application of Online Learning Mechanism Based on Data Drive

Zhao Liang[1], Xiaohua Xiao[1](✉), Ruixia Wu[2], Guili Zhang[1], Hanyue Zhang[1], Yinglong Yu[1], and Gang Cao[1]

[1] Network and Information Center, Southwest Petroleum University, Chengdu, Sichuan, China
{kathy,xiaoxh,zhgl,201999010104,201499010078,caogang}@swpu.edu.cn

[2] Marxism College, Sichuan Water Conservancy Vocational College, Chengdu, Sichuan, China

Abstract. With the rapid development of online teaching, teachers' and students' online teaching and learning activities have generated massive amounts of behavioral data. Through in-depth mining and analysis of these behavioral data information, such technologies as teacher-student clustering technology, user portrait technology, and academic performance prediction based on behavioral data, in order to improve the effect of online teaching, realize individualized teaching with thousands of people. And select a course for empirical application to help teaching staff master the learning trend of learners, gain a deeper understanding of learners' learning situation, and promote scientific and continuous improvement of online courses.

Keywords: Online Behavior Mechanism · Clustering of Teachers and Students · Academic achievement Prediction · Personalized Teaching

1 Foreword

"Internet + education" has fueled the rise of online teaching, while the COVID-19 epidemic has fueled unprecedented growth and development of online education in colleges and universities. Almost every college student now engages in online learning to varying degrees. The rapid development of online teaching has resulted in an increase in the number of teachers' "teaching" and students' "learning" online teaching behaviors, resulting in massive online teaching and learning data [1]. The question of how to use these data more effectively to help users better understand "teaching" and "learning" and improve the effectiveness of online teaching has emerged as a critical issue.

This paper thoroughly investigates and analyzes various process behavior data information of teachers' and students' online behavior mechanisms, yielding valuable information such as behavior rules and influencing factors of teachers' and students' teaching [2]. Assist teaching staff in mastering learners' learning trends and developing a deeper understanding of learners' learning situations in order to achieve scientific and continuous improvement in the quality of online course construction. Assist teaching administrators in recognizing the true application of the network teaching platform from

X. Yuan et al. (Eds.): ICEKIM 2023, AHCS 13, pp. 1695–1702, 2023.
https://doi.org/10.2991/978-94-6463-172-2_187

the top-level design level of informatization, so that targeted policies and strategies can be developed [3]. Promote the use of information technology and teaching in various regions, universities, disciplines, and specialties, vigorously promote the integration of information technology and education development, improve the quality of online education, and realize the organic combination of information technology, education, and teaching, as well as the deep application of educational technology.

2 Research on Data-Driven Online Learning Mechanism

The realization of "deep" learning in online teaching necessitates an examination of the online learning mechanism. Collect the operation and behavior traces of all teachers' and learners' teaching and learning behaviors on the online platform, and then conduct in-depth analysis and research [4]. Analyze the activity and behavior tendencies of teachers' and students' online teaching and learning behaviors, conduct an in-depth analysis of the characteristics of teachers' and students' online behaviors and their influencing factors, and propose a correlation and solutions.

2.1 Research on Evaluation of Learning Effect

Judgment of Attribute Correlation

The purpose of correlation analysis is to select differentiated attribute characteristics to express data laws. Teachers and students are divided into several categories based on historical online teaching behavior trajectory data using clustering technology, and then user portrait technology is used to analyze the categories of teachers with more online education workload and higher enthusiasm and students with higher success rate in online learning. Determine which teachers have "high value" in online teaching and which students have "high quality" in online learning. And compare the differences between different feature categories to determine the differences between different types of learners in each dimension.

Cluster Experiment of Teachers and Students

To begin, ten dimensions are set as analysis objects, including the number of courses, the number of tasks completed, the number of video tasks completed, the total time for students to watch videos, the number of chapter tests completed, the average score of chapter tests, the number of homework completed, the average score of exams completed, and the average score of exams. The closer each attribute's correlation to 1 is, the more relevant it is.

The correlation degree between "number of courses" and "number of tasks completed" and "number of video tasks completed" reaches 0.96, which is a similar feature, according to the research. The highest correlation between the average chapter test score and other attributes, on the other hand, is only 0.67. This phenomenon can be explained as learners' platform activity, which can reflect their "learning activity" to some extent, and there is a correlation between "activity" and achievement. However, the correlation between "number of exams completed" and "average exam score" is only 0.54, indicating that the number of exams has some influence on exam results. The score index is a comprehensive evaluation of the average score and exam completion rate, which reflects the students' overall performance. In addition, cluster and statistical analysis are used to determine which cluster has the highest average score and the highest average completion, in order to identify the best student group (as shown in Fig. 1). By analogy, it can analyze excellent curriculum groups with high homework completion, more activity forms, higher student participation, and better curriculum implementation effect by changing the dimension of clustering and setting different analysis objects from the teaching attributes of teachers.

Visualization of Clustering Results

A coding system for evaluating the value of courses was developed upfront. The course value is subdivided into 10 evaluation indicators: the total number of course materials and documents, the number of chapter videos, the total number of chapters, the total number of course resources, the total number of sign-in participants, the number of course activities, the number of interactive activities in the course, the total length of videos, the average completion of assignments and the number of course visits (course PV). By counting the historical data of the curriculum platform, the score values of each dimension are calculated, and these values are recorded in the curriculum evaluation coordinate system and connected with each other, forming an irregular broken-line closed-loop "course value judgment effect cloud".

To reflect the implementation effect of different courses, the normalized area, external perimeter, center of gravity, range, and dispersion coefficient are used. According to the clustering results, dynamic interactive visualization can be realized by using a radar chart to show the differences between clusters, and clicking on the corresponding clusters can control the hiding and displaying. A group's cloud image area is larger, indicating that its learning effect is stronger. Taking cluster 2 and cluster 3 as an example, through the comparison of the radar chart, it can be found that the student group of cluster 3 is better than the student group of cluster 2 in completing the course chapters, the number of

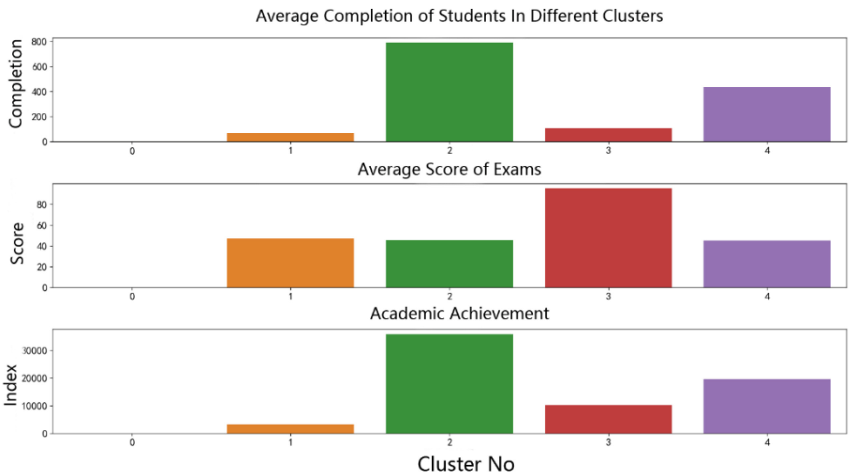

Fig. 1. Cluster analysis of teachers and students (Self-drawn)

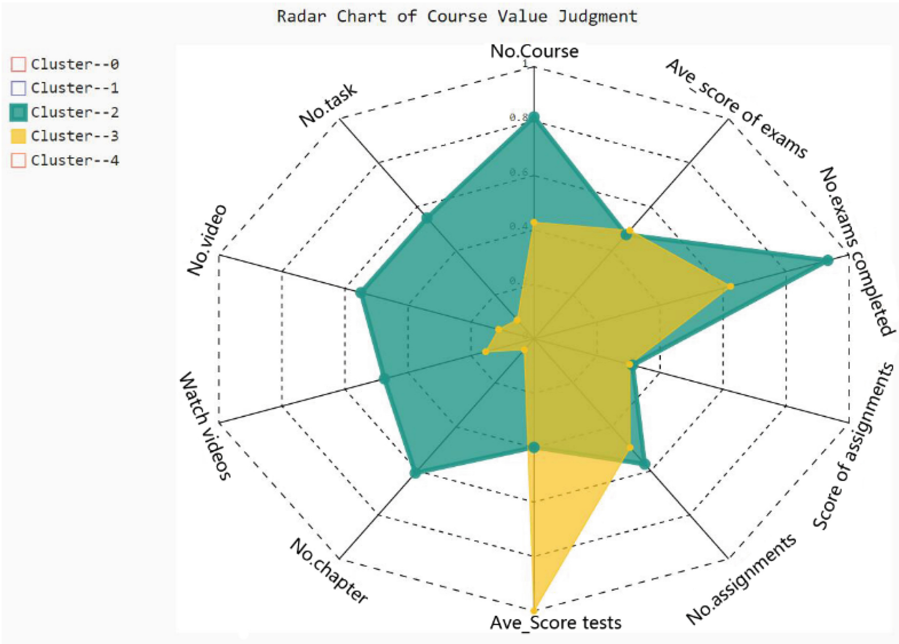

Fig. 2. Radar map showing differences between clusters (Self-drawn)

tasks completed, the number of videos watched, the number of exams completed, etc., but the average score of the test is lower than that of the student group of cluster 2, which shows that the students of cluster 3 have more general online behaviors, but the score is not high, It is characterized by "strong exploration and low scoring rate". While the students in cluster 2 have higher average scores in the test, they have fewer online learning behaviors and have a tendency of "chasing scores and ignoring the process". It is necessary for teachers to strengthen the management of daily teaching activities and increase the participation of cluster 2 in online teaching activities (as shown in Fig. 2). By analogy, we can objectively analyze the different learning characteristics of different learning groups, so as to help teachers provide targeted teaching methods and strategies, and help achieve personalized teaching effects.

2.2 Research on Academic Achievement Prediction Based on Behavioral Data

Despite the rapid development of MOOC, it emphasizes its unique advantages in overcoming time and space constraints and sharing high-quality educational resources [5]. Many people, however, question the effectiveness of online learning. This paper analyzes huge student data from Southwest Petroleum University's online learning platform, studies learner learning characteristics, analyzes learner types, and uses the supervised learning algorithm to predict students' test scores based on their historical behavior data. This research can encourage teachers to reflect on their teaching, improve teaching, make learners more suitable for online courses, reduce online learning drop-out rates, and achieve good teaching and learning [6].

First, the program is debugged using pycharm's console to obtain a general description of the data. The data available is analysed descriptively to gain an overall understanding of the data. Secondly, a predictive model was constructed by correlation analysis of the attributes. The correlation coefficient can be used to describe the relationship between the quantitative and variables and to determine whether there is a linear correlation between the dependent and explanatory variables [7]. It was determined that the mean score on the test was correlated with several items such as "mean score on homework" and "mean score on chapter quizzes", and a supervised learning algorithm was used to predict the mean score on the test based on these characteristics.

2.3 Relevance Analysis and Personalized Teaching Resources Push

Based on research into the differences between teachers' and students' using behaviors, this paper investigates the relationship between teachers' and students' online learning behaviors based on the characteristics of disciplines, majors, grades, and so on, specifically studies teachers' and students' behavior tendency on the platform, and comprehensively and deeply excavates users' needs. Through correlation analysis, it can assist teaching departments in carrying out precise resource construction, realizing personalized resource customization, and promoting the continuous evolution of digital education resources.

Combined with learners' learning style preference and learning group differentiation characteristics, we will push more appropriate explanation content and expanded materials; Through high-quality learning support services, we have achieved large-scale

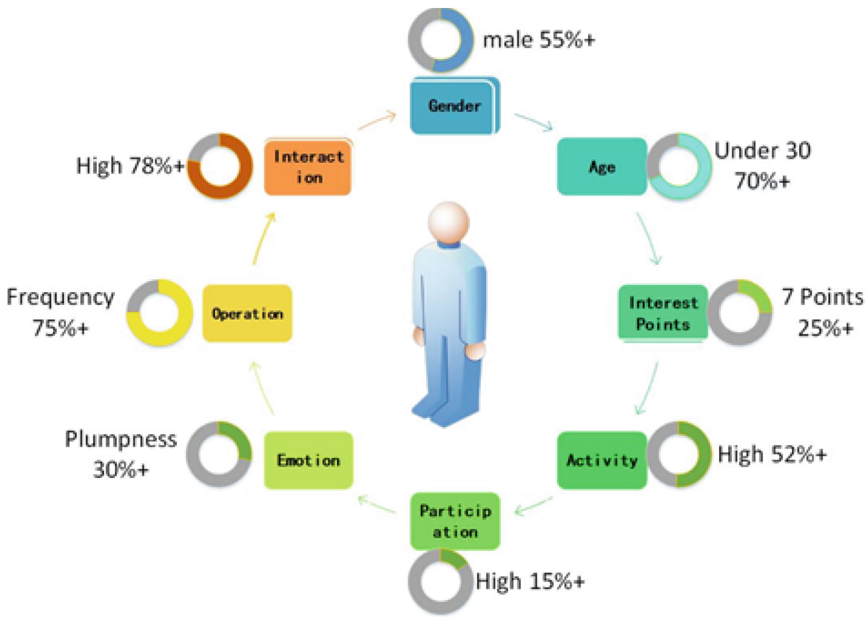

Fig. 3. Personal portrait of learning effect (Self-drawn)

individualized teaching of “thousands of people and thousands of faces” (as shown in Fig. 3).Realize the creation and integration of high-quality digital resources; Realize the intelligent push of high-quality digital education resources and the creation of a large number of distinctive resources; Realize the intelligent classification and dynamic aggregation of learning resources; Realize the improvement of digital education resource utilization rates and the creation of a high-quality digital education resource ecology [8].

3 Application of Data-Driven Online Learning Mechanism

3.1 Course Selection

An empirical study was conducted using Southwest Petroleum University’s cultural quality course “Digital Photography Technology and Art” as the research object. Learners of the course include not only Southwest Petroleum University undergraduates, but also online massive open online course teaching for social learners based on a domestic online teaching platform in 2018, with more than 30,000 students taking courses, indicating a large audience.

3.2 Application Process

To begin, different categories of online teaching and learning between teachers and students are obtained through clustering and user portrait analysis, and the differences of each category of people in each dimension are analyzed, in order to deeply understand the status quo, problems and challenges of online learning between teachers and students, and the influence of different dimensions on online learning activity and tendency.

Secondary, the study of student behavioural mechanisms based on behavioural data provides an in-depth analysis of student behavioural data. By studying the motivation of learners, the types of learners are derived, and corresponding prediction models are built for prediction. By studying the learning characteristics of learners, the types of learners are analysed and corresponding prediction models are built to predict the average completion rate of learners and generate early warnings of dropouts from online courses. It is important for teachers to reflect on their teaching, improve their teaching, make learners more comfortable with online courses, reduce online learning dropout rates, and help teachers to understand the real online learning trends of learners.

Third, it investigates the relationship between various influencing factors, discovers the relationship between learning behavior and learning effect by analyzing the interaction between learners and learning environment, and investigates the correlation of teachers’ and students’ online learning behaviors from the characteristics of disciplines, majors, grades, and so on, on the basis of studying the differences between teachers’ and students’ using behaviors, specifically It assists teaching departments in carrying out the precise construction of resources and realizing personalized customization of resources through correlation analysis. And recognize the push for personalized teaching resources, as well as promote educational resource ecological balance.

3.3 Application Results

By extracting and analysing online teaching behaviour data, the characteristics of teachers' and students' online behaviour and their influencing factors are derived, and the results are fed back into teaching. From the perspective of tutors, teaching reflection is carried out on learner learning process records, providing data support for the precise positioning of teaching objectives and teaching methods, supporting the improvement of online teaching design, and providing feasible suggestions for the implementation of personalised teaching and the reform of teaching models.

4 Conclusions

This data-driven research is based on the database of Southwest Petroleum University's online teaching platform, conducts a large-scale online learning mechanism research of teachers and students, and uses the log reflecting teachers and students' teaching behavior as the data object [9]. It is possible to improve the utilization rate of online teaching resources and the quality of digital teaching resource construction from the perspective of users by analyzing the activity and behavior tendency of teachers' and students' online teaching and learning behavior.

This paper presents the correlation and solutions in the post-epidemic era by analyzing the characteristics and influencing factors of teachers' and students' online behaviors. Predict, group, model, and monitor learning activities using teacher-student clustering, association rules, and visualization techniques [10]. Using data mining technology to improve decision-making quality promotes the development of online teaching. It will be the driving force and source of continuous development of online teaching through the above-mentioned in-depth study of online teaching by using learning analysis technology to provide a personalized and more interactive educational environment to improve the learning effect and teaching methods.

Acknowledgments. This study was supported by 2021 Sichuan Education Informatization Research Project of Sichuan Higher Education Association (grant no. GJXHXXH21-YB-23).

References

1. Li C. Y. Research on the mechanism of students' online learning behavior in the context of big data in education[J]. Computer Knowledge and Technology, 2021, 17(20):41–42.
2. Jing YJ, Li X, Jiang Xue. Analysis of factors influencing online learning behavior intention and educational inspiration in the post-epidemic era[J]. Chinese e-learning, 2021(06):31–38.
3. Weng J.L. Research on learning behavior based on user portraits and data mining [D]. Yunnan University of Finance and Economics, 2021.
4. Chen X, Xu H.L. Research on the effectiveness of online learning and improvement paths of undergraduate students in the context of epidemic--an analysis of a survey based on "stopping classes and studying" in a university[J]. China Education Informatization, 2021(09):76–80.
5. Liu J.Y, Huang L, Peng J. Analysis of online teaching behavior of teachers and students' satisfaction with online learning in higher education[J]. Education and Teaching Research, 2021, 35(04):79–88.

6. Liu Siqiang, He Hui. Effective teaching behavior in online teaching of university courses - empirical analysis based on large-scale online teaching during the epidemic [J]. Modern Education Management, 2022 (03): 66–73. DOI: https://doi.org/10.16697/j.1674-5485.2022.003.008
7. Gao Ya, Li Nannan. Analysis of the application of online teaching optimization strategy during the novel coronavirus pneumonia epidemic [J]. China Contemporary Medicine, 2021, 28 (19): 217–220
8. Sun Yan. Analysis of problems faced by online teaching and strategic thinking [J]. China Modern Education Equipment, 2021 (01): 64–65+69. DOI: https://doi.org/10.13492/j.cnki.cmee.2021.01.021
9. Li Shuang, Liu Zijing, Zheng Qinhua. Research on data-driven online teaching quality evaluation in the intelligent era [J]. Research on Audio-visual Education, 2022,43 (08): 36–42+76. DOI: https://doi.org/10.13811/j.cnki.eer2022.08.005
10. Liu Yan, Chen Jingbo. Resource application and data analysis of online teaching platform under the COVID-19 [J]. China Education Technology Equipment, 2021 (20): 58–60+65