International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 6370-6381

https://doi.org/10.70517/ijhsa463547

AI large model-driven OMO teaching model intelligent learning path planning and recommendation algorithm

Fangli Li[1,2,*] and Qinying Li[1,3]
[1] School of Information Engineering, Jiangxi University of Technology, Nanchang, Jiangxi, 330098, China
[2] Faculty of Social Science, Arts and Humanities, Lincon University College, Selangor, 47301, Malaysia
[3] Faculty of AI Computing and Multimedia, Lincon University College, Selangor, 47301, Malaysia
Corresponding authors: (e-mail: jxlily202502@163.com).

Abstract Internet technology promotes the innovation of education mode, and OMO teaching mode deeply integrates online and offline teaching and lacks personalized learning path planning. The accumulation of large-scale educational data provides a foundation for learner behavior analysis, and the construction of knowledge point difficulty model and learner state model through data analysis algorithms realizes personalized learning path planning for different learning styles, improving learning effect and satisfaction. In this study, cognitive network analysis (CNA), social network analysis (SNA) and content analysis (CA) are used to analyze and process educational data, focusing on learners' online learning behaviors and practice test scores, and establishing a model for judging learners' learning status. The experimental results show that the learning path planned based on the CNA method is more in line with the user's learning style, and the similarity with the user 2 learning style reaches 0.90, which is higher than the SNA and CA methods. The data analysis showed that the number of library behaviors of students with good grades (about 19.73% of the total number of students) reached 52 times, which was significantly higher than that of students with average grades (46 times) and students with poor grades (40 times). In addition, seven key factors were extracted through principal component analysis, which could explain 69.945% of the overall variance, effectively reflecting the correlation between students' behaviors and academic performance. The study proves that personalized learning path planning based on large-scale educational data analysis can effectively meet the needs of users with different learning styles, improve learning efficiency and user satisfaction, and provide effective methodological support for the practical application of OMO teaching mode.

Index Terms OMO teaching mode, large-scale educational data, data analysis algorithm, personalized learning path, cognitive network analysis, learning style

I. Introduction

With the development of information technology and the popularization of intelligent terminal equipment, under the development trend of “Internet+Education”, the pure traditional classroom teaching mode or pure online teaching can no longer adapt to the development of the times, and the online-offline integration (OMO) teaching mode can combine the advantages of both organically [1]-[4]. The OMO teaching model does not refer to a simple substitution of online and offline teaching, but rather a fitting ratio in which both online and offline teaching can maximize their advantages [5]-[7]. The OMO teaching mode can not only popularize high-quality educational resources rapidly on a large scale, but also satisfy the needs of students' personalized learning and evaluation on a large scale [8]-[10]. Therefore, in the actual teaching process, teachers should rationally and scientifically carry out the design of OMO teaching mode, so as to better further enhance the teaching effect and efficiency [11], [12].

However, in the OMO teaching mode, the planning of learning paths has become a problem for teaching [13]. Learning path planning is crucial to the quality of teaching and learning, and a scientific and reasonable learning planning can help students acquire knowledge and improve their abilities efficiently in OMO teaching mode [14]-[16]. And with the development of artificial intelligence, intelligent algorithms are gradually throwing their weight around in learning path planning [17]. Large-scale educational data analysis algorithms are algorithms that are often used to process and analyze large amounts of educational data, and their purpose is to extract valuable information from the data, so as to support educational decision-making, optimize the teaching process, and improve the quality of education, which includes data mining, data analysis, machine learning, and data visualization, etc [18]-[21]. In the OMO teaching mode, the focus of large-scale educational data analysis algorithm learning path planning can be multi-dimensional data collection and intelligent analysis, so as to build a dynamic navigation system with personalization and realize the integration of online and offline teaching elements [22]-[24].

The development of education informatization has given rise to a variety of teaching mode innovations, and the scale of education data has been growing explosively, which contains rich information on learner behavior and learning characteristics. Currently, Internet learning resources continue to enrich, personal terminal promotion, development of application tools and connectivity learning space and other content has made online and offline blended learning presents mutual penetration, seamless switching and diverse docking characteristics, online and offline “blended” learning is stepping into the “fusion” stage. OMO teaching mode, as a new type of teaching mode, will be the depth of the integration of online and offline teaching, to create the overall redesign of the teaching and learning system. OMO teaching mode as a new teaching mode will deeply integrate online and offline teaching, create an overall redesign of the teaching and learning system, and realize the best synergy between face-to-face teaching and online learning. Professor Li Jiahou pointed out that OMO teaching mode is a teaching mode that optimizes the selection and combination of all teaching elements to achieve teaching goals. Zhu Zhiting et al. clearly pointed out that the core of OMO teaching mode is the learning of learners, and with the help of technology to open up the online and offline, virtual and real learning scenarios, to form an integrated online and offline learning situation, and to meet the personalized learning needs of learners. However, there is a lack of effective personalized learning path planning methods in the existing OMO teaching model, which makes it difficult to meet the needs of learners with different learning styles. Personalized learning path planning aims to provide learners with different learning states with learning sequences for efficiently mastering knowledge, and the accumulation of educational big data provides the possibility of realizing personalized learning path planning. Currently, online education platforms in colleges and universities generate huge and rich data, including display data and implicit data, structured and unstructured data, etc. How to effectively use these data to extract learner behavioral characteristics, construct knowledge point difficulty models and learner state models, and then realize personalized learning path planning has become a key issue in the practice of OMO teaching mode.

This study explores the learning path planning method in OMO teaching mode based on large-scale educational data analysis algorithm. First, through in-depth analysis of the concepts and characteristics of OMO teaching mode, we construct the OMO model for large-scale course teaching in colleges and universities; second, combining data analysis methods such as cognitive network analysis (CNA), social network analysis (SNA), and content analysis (CA), we mine and analyze the data of learners' online learning behaviors, and construct the difficulty model of the knowledge points and the state model of the learner; and then, based on the learners' learning state and style characteristics, personalized learning paths are designed by algorithms; finally, the effects of different methods are verified through experiments, and the advantages and disadvantages of the three algorithms, CNA, SNA and CA, are analyzed and compared in personalized learning path planning. This study will provide effective methodological support for the practical application of OMO teaching mode, promote learners' personalized development, and improve the quality and effect of teaching.

II. Relevant concepts

II. A. OMO teaching model

At present, the Internet learning resources are constantly enriched, and the promotion of personal terminals, the development of application tools, and the connection of learning spaces have made online and offline blended learning manifested in the mutual penetration, seamless switching, and diversified docking of the two learning modes, which, at the same time, also indicates that online and offline “blended” learning has stepped into a mutual “integration” stage. At the same time, it also indicates that online and offline “blended” learning has stepped into a stage of mutual “integration”. Currently, scholars at home and abroad have different opinions on the definition of OMO teaching mode, which implies the overall redesign of the teaching and learning system to create the best synergy between face-to-face teaching and online learning, rather than randomly applying skills to teach. Professor Li Jiahou pointed out that OMO teaching mode is a teaching mode that optimizes the selection and combination of all teaching elements to achieve teaching goals.

Explaining the problem of “integration” from the different levels of complexity of blended learning, it is concluded that at the level of “integration” blended learning, online and offline teaching and learning are fully integrated, with no obvious boundaries, and in the process of teaching and learning, it is more dependent on online learning platforms, online courses and learning tools, and some classes are even mainly online. The teaching process is more dependent on online learning platforms, online courses and learning tools, and some classes are even mainly online learning activities. Zhu Zhiting et al. clearly pointed out that the core of OMO teaching mode is learners' learning, and with the help of technical means, it opens up the online and offline, virtual and real learning scenarios to form an online and offline integrated learning situation, which meets the personalized learning needs of learners, realizes the learning objectives, and achieves the new pattern of personalized teaching and service.

In summary, although domestic and foreign scholars do not have a unified definition of OMO teaching mode for the time being, in general, OMO teaching mode is carried out in the open and shared, virtual and real integration of the integrated learning space, providing multiple teaching scenarios and teaching organization to meet the various teaching needs of teachers and learners, promoting personalized learning of learners, and fostering creative thinking of learners [25].

II. B.Large-scale OMO teaching model construction

According to the analysis of the above two theoretical models, elements are extracted for the construction of OMO teaching mode. First of all, by analyzing the construction of hybrid teaching system, it can be seen that the OMO teaching process can be divided into three main steps: first, online learning, learners can learn through the online learning platform on subject knowledge points and stage teaching tasks for independent learning; second, online and offline integration, mainly reflected in the teaching evaluation, Q&A and other aspects of the classroom, the teacher can be through the online learning platform to assist in the teaching process, the learner's learning situation Diagnostic evaluation is to guide learners to solve problems, so as to more accurately grasp the learning situation of learners and timely adjustment of the offline teaching design; Third, offline teaching, building a learning community, through cooperative learning, teachers and students, students and students, human-computer interaction and other ways to create a more immersive learning atmosphere, to enhance the teaching of interest, so that learners take the initiative to learn. Secondly, by analyzing the design principles of OMO teaching mode, it can be seen that the main consideration of three factors: first, the teaching environment, before the start of teaching, should build OMO teaching environment, determine the online teaching platform and offline learning place, can support the learners to participate in all aspects of learning; second, the information technology, in the whole teaching process, should be considered in the main body of the teaching activities whether there is a good use of experience? Whether it can be based on teaching services to help teachers teach? Third, the teaching process, OMO teaching mode should design a variety of teaching activities to drive teaching, cultivate learners' independent learning ability, and promote learners' personalized development.

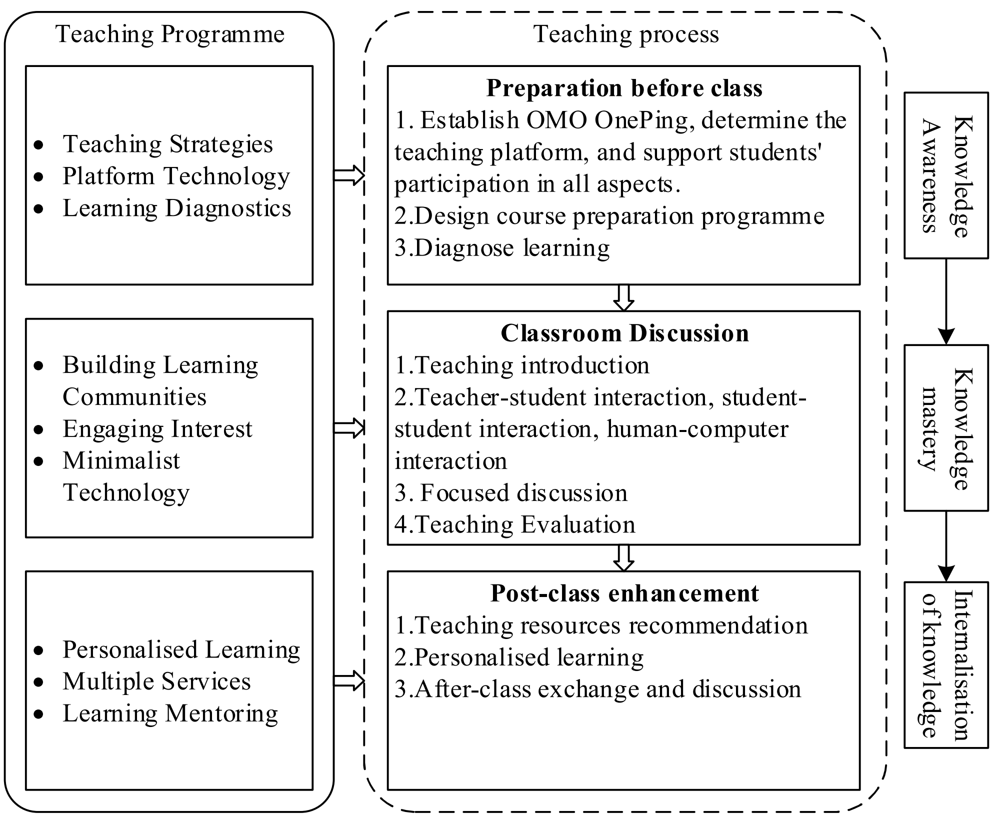

Figure 1: University curriculum teaching OMO mode

This study combines the construction of hybrid teaching theory system and the design principles of OMO teaching model, and at the same time, refers to the “double helix teaching mode”, “dual-teacher classroom” and other new teaching modes, to construct the “large-scale course teaching in colleges and universities”. The OMO model is shown in Figure 1. The model is mainly divided into two modules: teaching plan and teaching process, and the construction of the model is aimed at personalized teaching with the learner as the main body and improving the quality of university course teaching. The pre-study stage before class is the awareness of knowledge, the classroom exploration stage is the mastery of knowledge, and the post-course enhancement stage is the internalization of knowledge [26].

II. C.Personalized Learning Pathway Planning

The output of personalized learning path planning is a sequence of rearranged knowledge points. This learning path is designed to plan a series of learning sequences for learners with different learning states to meet their needs for efficient knowledge acquisition. Currently, in existing personalized learning path planning studies, the input parameter is generally the learner's personal trait information, and the output is a series of sequences of knowledge points that meet the learner's traits. Based on existing research, the personalized path planning involved in this study cannot be separated from three main steps, which are.

(1) Knowledge point difficulty calculation based on the learner's online behavior to prepare for the subsequent arrangement of personalized path guidance.

(2) Constructing a corresponding learner state model based on the learner's online learning behavior information, judging the learner's learning state, and updating the learner's state.

(3) To be able to plan the appropriate learning content sequence for different learners according to their learning status, i.e., personalized learning path [27].

The personalized path planning framework mainly relies on the learner's online learning behavior to construct the learner state model, knowledge point difficulty model to finally realize the personalized learning path as shown in Figure 2.

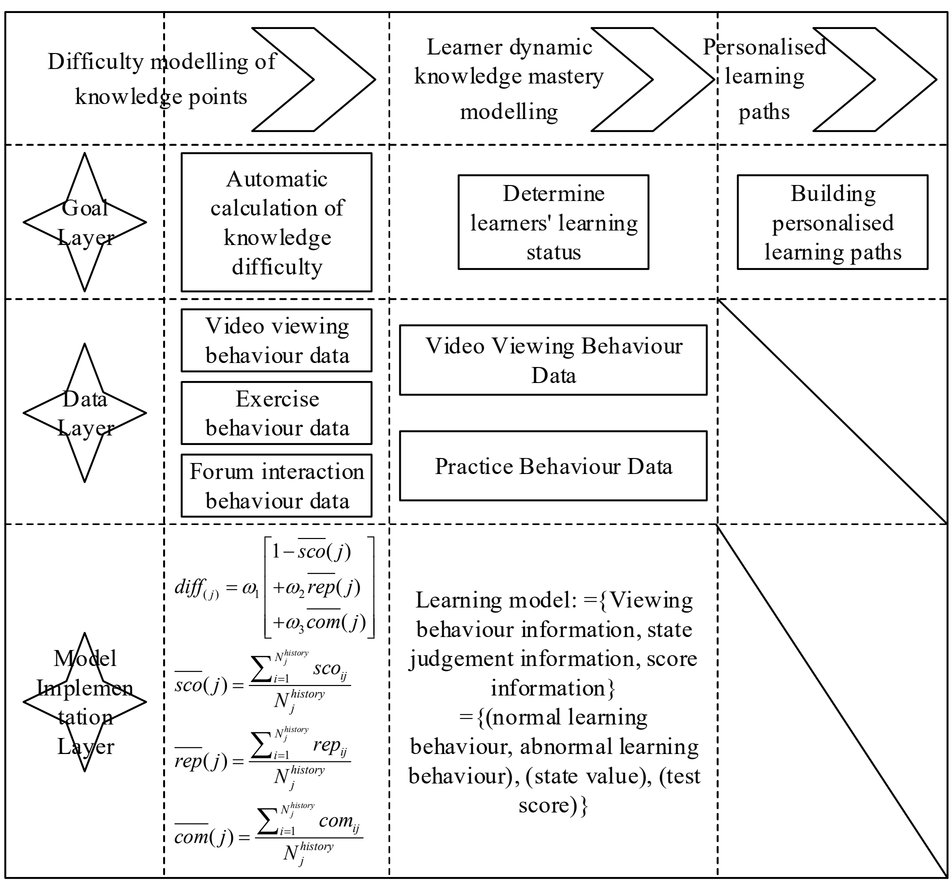

Figure 2: Personalized learning path research framework

III. Algorithms for analyzing large-scale educational data

III. A. Educational Data Classification and Characterization

Online education platform for students to meet the students can watch teaching videos, browse the teaching content and other requirements at any time and any place, for teachers, but also allows teachers to build the teaching content at any time and any place, scheduling the course time. The data generated on the online education platform is huge and rich. Reasonable classification of these data is also an important task, and the general classification is shown in Figure 3.

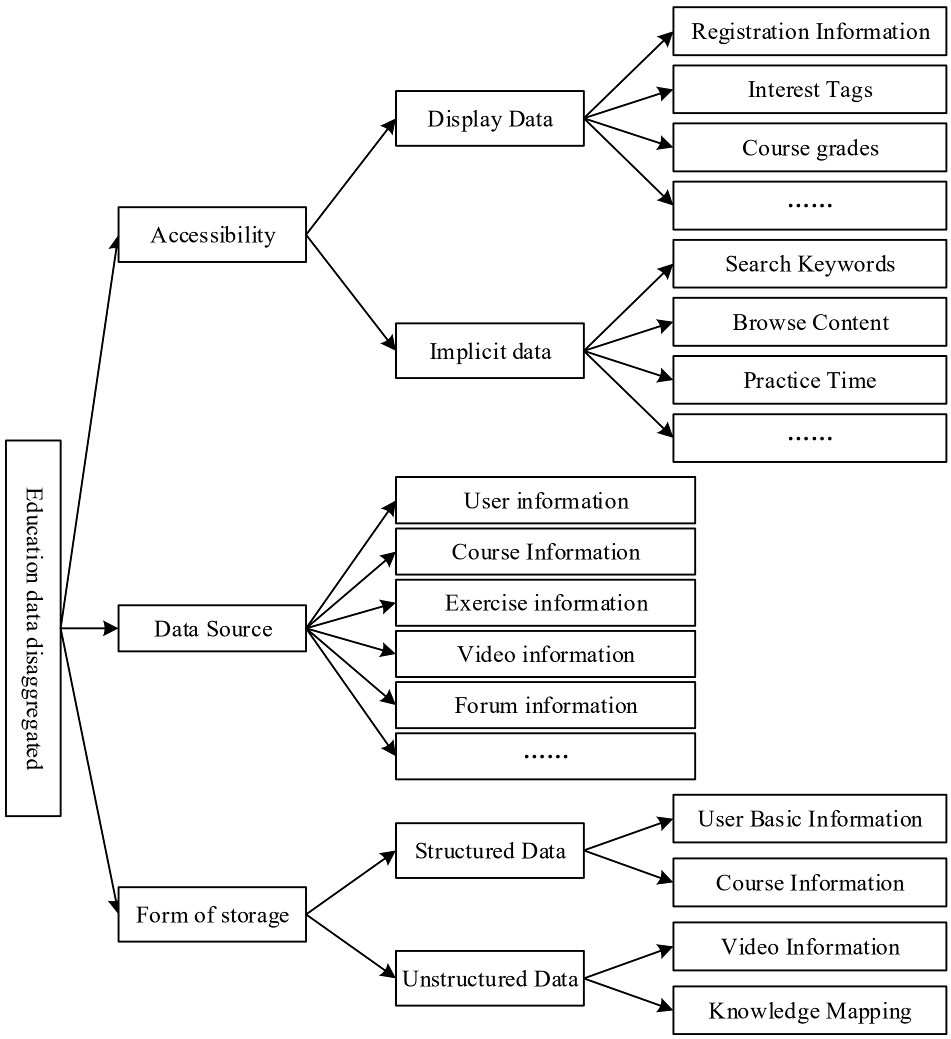

Figure 3: education data classification

(1) Classification according to the way of acquisition

Educational data can be categorized into display data and implicit data according to the way of obtaining educational data. Information that can be directly obtained is display information, such as user registration information, favorite tags, selected courses, subject grades, forum evaluations and so on. Information that cannot be seen directly is implicit information, such as search traces, history labels, browsing time, etc.

(2) Categorization by source

Data can be classified into information such as personal information, course information, image information, forum interaction information, behavioral information and knowledge graph through data sources. Personal information specifically refers to the user's basic information, subject grades and study habits. Course information specifically refers to the teaching content, course structure and other information. Video information consists of audio, video, etc. Forum interaction refers to learning postings, communication between posters; behavioral information includes browsing behavior, watching video records. Knowledge graph, on the other hand, is the relationship graph between knowledge points. Most of the knowledge mapping is done by teachers first, and then combined with manual methods to automatically complete the update and maintenance.

(3) Classification by storage form

Data classification from the perspective of the storage form of data can be divided into structured and unstructured. The so-called structured refers to the two-dimensional table structure, the data that can be stored in the form of logical structure is called structured data.

III. B. Algorithms for analyzing educational data

III. B. 1) Cognitive Network Analysis (CNA)

Inspired by cognitive games, cognitive network analysis has been extracted, which focuses on the different performances of individuals in virtual environments to determine their cognitive abilities. Currently, educational games in MOOCs are exactly this type of analysis, which is evolved from the social network analysis (SNA) method, where diverse relationships in interactive games are extracted through characterization and quantification into computer-acceptable data. Learners' cognitive models are constructed from their respective personalities, values, learning abilities, and cognitive theories, and learners present different performances in different teaching and learning environments, and individual cognitive models can be constructed and refined by accumulating data through continuous iteration [28].

III. B. 2) Social network analysis (SNA)

In sociology there is a concept of social relations modeling (SRM), which reveals the relationships that members of a society have with each other and the effects of individual-to-individual connections. This method of analysis is based on the mathematical idea of graph theory, which describes the interrelationships between individuals in the form of a topological map of a network. Scholars in the field of education integrate multiple disciplines to create a social structure in the field of education by mathematically exploring learner performance, social influence, and relationships between learners, and then analyzing the influence of groups and structures within this social structure to create a network of related relationships. Usually social network analysis requires certain computer software for analysis and visualization. Visual social network analysis tools are Tulip, Gephi, NodeXL, UCLINET, GISMO, SNAPP, Pajek. The most powerful and open source of these learning analytics tools is UCLINET, which allows for cluster analysis, centrality and location analysis, and closeness analysis between members of a learning network community on data generated during the learning process [29].

III. B. 3) Content Analysis Method (CA)

The content analysis method focuses on the content of the learning through objective description and systematic quantitative analysis of the content related to this learning. The descriptive and quantitative analysis analyzes the language and characteristics of the content and examines its impact on the dissemination of information in order to draw conclusions about the context and the environment in which the content is used. This method of analysis, through qualitative and quantitative analysis of students' learning process and learning content and learning effects, ultimately predicts the characteristics of students' learning behaviors and forecasts their learning [30].

III. B. 4) Key Steps in Learning Path Planning

The three main steps of personalized path construction are divided into the following steps.

(1) Construct a knowledge point difficulty model and automatically calculate the difficulty of knowledge points. The detailed method is to take learners' historical learning behaviors, such as learners' posting and replying behaviors in forums and learners' practice test scores, as the main input parameters for constructing the knowledge point model.

(2) Construct the learner state model. The model is mainly based on the normalized results of learners' online learning behaviors and practice test scores to determine learners' learning status.

IV. Educational Data Analysis and Pathway Planning Effectiveness Analysis

IV. A. Data Collection and Cleaning

The data used in this paper comes from an open-source dataset from a third-party platform, which is mainly used for public research projects on topics such as personal development of students and supervision and management of schools. The content of this dataset includes the behavioral data of students using campus one-card swipe cards and the grade ranking data in the teaching management system in two academic years of a university. Specifically, it includes book borrowing data, one-card data, dormitory access data, library access data, and student achievement data. The student information (e.g., student number) in the dataset has been anonymized through data desensitization. The information of each part of the data fields is shown in Table 1.

Table 1: The various points are based on field information

Student library borrowing data		
Field number	Field name	Field description
1	id	School number
2	borrow_time	Lending time
3	book	title
4	isbn	Book number
Student 1 cartoon system data		
1	id	School number
2	card_type	Consumer class
3	card_place	Consumption location
4	card_way	Consumption mode
5	card_time	Consumption time
6	card_amount	Consumption amount
7	card_balance	Residual amount
Student dormitory threshold data		
1	id	School number
2	dorm_time	Access time
3	dorm_direction	In and out the direction (0 to 1)
Student library threshold data		
1	id	School number
2	library_time	Access time
3	library_time	Entrance number
Student scores ranking data		
1	id	School number
2	college	College number
3	rank	Ranking

Among them, the consumption category field specifically includes POS consumption, card loss and card replacement. Consumption method fields mainly include cafeteria, supermarket, boiling water, bath, library, laundry, copy center, academic affairs office, school bus and school hospital. Data cleansing is generally application-specific. Specifically, data cleansing is the process of streamlining a database to remove duplicate records and converting the remainder to an acceptable standard format. To better protect student privacy, grades are converted to rankings and normalized, and data mining algorithms are utilized for learning. The consumption method field in the one-card data is the key data to be analyzed, but when there is a missing value in the raw data, the row where this data is located should be deleted by the operation. When the borrowing time, consumption time, in and out time fields appear twice exactly the same situation, the same will be the data where the line for deletion operations, repeated records will affect the data analysis, and may even lead to the existence of errors. If more than half of the traditional features extracted in a line appear to be 0, it should be interpreted as the student's behavioral data in school due to personal reasons or other reasons are missing, and the operation of deleting this line is carried out. Through data transformation and integration, data from multiple data sources are combined and then merged and transformed into a form suitable for use in data mining, so as to carry out data generalization and delete the attributes in the original attribute set that are not relevant to the mining task of this campus card, so as to achieve a reduction in the number of dimensions.

IV. B. Learning data analysis

In order to better interpret the behavioral data and have a more comprehensive understanding of the data in advance, this paper first classifies the students' performance ranking into three categories, i.e., good, medium and poor, based on normal distribution. The labels were set as “1”, accounting for 19.73% of the total number of students, “2”, accounting for 60.11% of the total number of students, and “3”, accounting for 20.16% of the total number of students, respectively. 20.16%. The campus behavior and grade ranking data were then analyzed to convert the raw behavioral data into behavioral characteristics related to academic performance. This paper compares the differences in the number of trips to the library, borrowing books, and going to the Document Printing Center between the three types of students during exam and non-exam periods. The study data was analyzed as shown in Figure 4, where library borrowing data was counted for students from all colleges over a period of two academic years. The results show that students of category I visited the library 52 times, students of category II visited the library 46 times and students of category III visited the library 40 times. The behavior of studying and borrowing books from the library decreases in order with the category of students. The first category of students are most equipped with good study and borrowing habits, while the third category of students have relatively fewer of the above three behaviors and lack practical action in their daily campus routine.

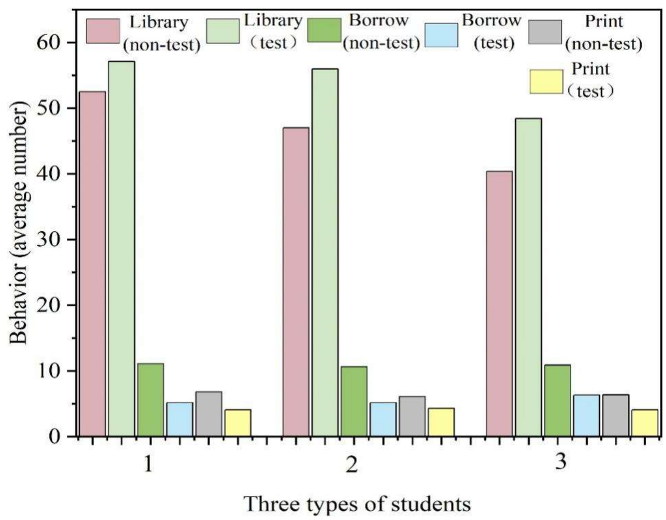

Figure 4: Behavior comparison

Due to the non-uniformity of teaching arrangements and daily management of the colleges, the students' work and rest patterns will also differ, in order to avoid large data errors, this paper selected the No. 1 college students' behavioral data to be analyzed independently as shown in Fig. 5, and the statistics of the three types of students' three-meal diets in the No. 1 college, the results show that the number of lunches and dinners is much larger than the number of breakfasts, which indicates that most of the students have the phenomenon of irregular diets. . Students belonging to the first category have the highest number of breakfasts, indicating that in comparison the students in this category have more good eating habits. Whether it is the utilization of books, the contribution to club activities, or the consumption of food, all of them show the importance of self-discipline to a college student, which is also determined by the characteristics of college students themselves.

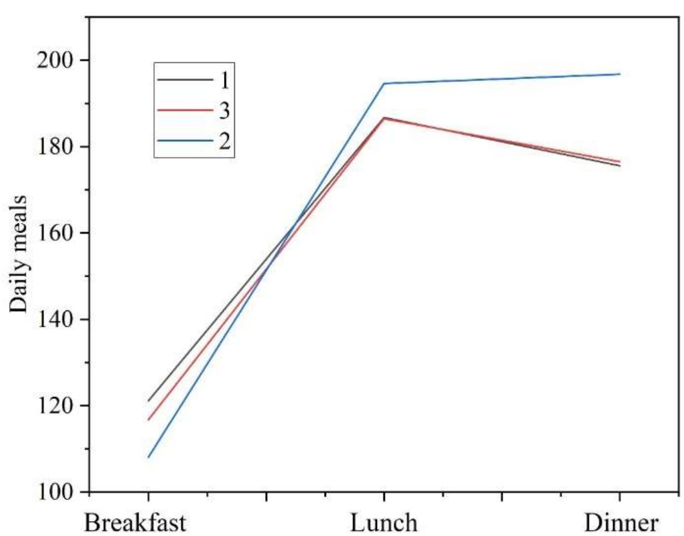

Figure 5: The three meals statistics of three students of the college

In order to prove the correlation between students' behavioral characteristics and academic performance, this paper further uses the statistical methods of factor analysis and principal component analysis to explain the correlation. Firstly, factor analysis was used to extract the common factors, and the factors affecting academic performance were rationalized by rotating the component matrix to explore the weight of each factor and common factor on the impact of performance. The principle of factor analysis is a statistical analysis method that belongs to the process of dimensionality reduction by studying the correlation coefficient matrix between variables and reducing the intricate relationship between these variables to a few comprehensive factors with minimal loss of information and classifying the variables accordingly. Therefore, the basic structure of the observed data was explored by examining the internal dependencies of students' behavioral characteristics, and several hypothetical variables were used to represent their basic data structure. Before factor analysis, the selected features were tested by KMO and Bartlett's test, and the KMO test was used to check the correlation and partial correlation between the variables. The closer the KMO statistic is to 1, the stronger the correlation between the variables, and the weaker the partial correlation is, the better the effect of factor analysis is. The results of the analysis are shown in Table 2, and the KMO statistic is 0.722 (>0.6), which is suitable for factor analysis.

Bartlett's spherical test judged that if the correlation array is a unit array, the independent factor analysis method for each variable is invalid. The results showed that the approximate chi-square value of Bartlett's test was 48490.940 and the concomitant probability value of Sig. < 0.05 reached the significance level. It indicates that the correlation coefficient matrix of the factors is not a unit matrix and there is correlation among the variables. Therefore, based on the results of the above analysis, the original variable (behavioral characteristics of students) is suitable for factor analysis.

Table 2: Kmo and bartlett tests

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	0.722	
Bartlett's Test of Sphericity	Approx. Chi-Square	48490.940
	df	154
	Sig.	0.001

Principal component analysis is to find out the independent composite indicators that reflect multiple variables and how the internal structure between multiple variables can be revealed by several principal components. Using principal component analysis to examine the correlation between multiple variables and the predictive function, 18 factors affecting the ranking of students' performance were analyzed as shown in Table 3, and the first seven factors were able to explain 69.945% of the overall variance, which effectively reflected the overall information and had a significant relationship with academic performance.

Table 3: explains the total variance

	Initial eigenvalue			Extract the sum of squares and load			Rotate the squares and load		
Constituent	Tot	V %	C%	Tot	V %	C%	Tot	V %	C%
1	4.312	23.930	23.930	4.312	23.930	23.930	3.415	18.959	18.959
2	1.905	10.574	34.504	1.905	10.574	34.504	1.784	9.884	28.843
3	1.714	9.463	43.967	1.714	9.463	43.967	1.763	9.735	38.578
4	1.425	7.714	51.681	1.425	7.714	51.681	1.650	9.135	47.708
5	1.178	6.489	58.17	1.178	6.489	58.17	1.553	8.635	56.321
6	1.114	6.152	64.322	1.114	6.152	64.322	1.310	7.256	63.563
7	1.013	5.623	69.945	1.013	5.623	69.945	1.145	6.382	69.945
8	0.906	5.028	74.973						
9	0.887	4.897	79.87						
10	0.750	4.160	84.03						
11	0.714	3.938	87.968						
12	0.482	2.642	90.61						
13	0.432	2.391	93.001						
14	0.368	2.036	95.037						
15	0.304	1.682	96.719						
16	0.284	1.534	98.253						
17	0.194	1.051	99.304						
18	0.130	0.696	100						

Campus card swiping behavior belongs to students' spontaneous behavior and is an important factor reflecting students' academic performance. Due to the individual differences of each student and the diversity and complexity of environmental factors, it is not possible to generalize and explain the model by extracting the campus card swiping behavior characteristics alone. Behavioral indicators affecting academic ranking can be discussed in the context of student performance classification, and data mining algorithms can be used to explore the practical value of these behavioral data in management, teaching and learning.

IV. C. Analysis of path planning effects

In order to verify the superiority of the above methods, this paper compares the results derived from learning style-based CAN with those derived from SNA and CA methods in terms of user satisfaction and the degree of matching between the paths and users' learning styles. This experiment has been tested on all four experimental users, and it can be clearly seen that all three methods continuously derive learning paths with higher user satisfaction as the number of iterations increases, and gradually stabilize after reaching a certain height. It can be seen that for users with different learning styles, all three methods are able to plan online learning paths for users. However, from the experimental results of all four users, it can be seen that through 50 iterations, the results of the CAN method proposed in this paper are better than the other two, that is to say, the learning path planning method proposed in this paper is able to obtain a higher degree of user satisfaction. Secondly, the learning paths derived from the three methods were compared to assess the degree of matching between the learning path and the user's learning style, taking User 2 as an example. The results are shown in Table 4, the similarity between the learning path and the user's learning style is the average of the similarity of 10 nodes, which shows that the learning path planned by the CAN method is more compatible with the learning style of User 2.

Table 4: learning path and user 2 learning style similarity

Method		CAN	SNA	CA
Learning path	Note1	0.73	0.80	0.74
	Note2	0.74	0.64	0.62
	Note3	0.80	0.68	0.75
	Note4	0.84	0.58	0.86
	Note5	0.82	0.76	0.80
	Note6	0.90	0.76	0.76
	Note7	0.70	0.80	0.68
	Note8	0.73	0.86	0.69
	Note9	0.75	0.82	0.58
	Note10	0.63	0.64	0.68
	Note11	0.78	0.73	0.70

In this section, the learning path planning method in the OMO teaching model proposed in this paper based on the data analysis algorithm is validated by crawling the data from the OMO teaching model of a large-scale online learning platform. The feasibility and effectiveness of the three algorithms CAN, SNA and CA are compared by simulating experimental subjects with four different learning styles. The results show that the learning paths planned by the CAN algorithm are more in line with the learning styles of the users and can bring higher satisfaction to the users.

V. Conclusion

Analysis of the data showed that there were significant differences in the behavioral patterns of students with different academic achievements on campus. Comparison of library behaviors showed that the number of library visits made by students in category I (high achievers) was 52, which was significantly higher than that of students in category II (46) and category III (40). Analysis of student behavioral characteristics by KMO and Bartlett's spherical test yielded a KMO statistic of 0.722 (>0.6) and a Bartlett's test significance level of $p=0.001$, indicating that student behavioral characteristics are suitable for factor analysis. The results of principal component analysis showed that the first seven factors explained 69.945% of the overall variance, effectively reflecting the correlation between students' behavior and academic achievement.

Comparison experiments of personalized learning path planning methods proved that the CAN method performed well in user satisfaction and learning style matching. Taking User 2 as an example, the highest similarity between the learning path planned by the CAN method and the user's learning style reaches 0.90, which is higher than the SNA method (0.86) and the CA method (0.86). The experimental results of users with four different learning styles consistently show that the learning paths planned by the CAN method can obtain higher user satisfaction.

Personalized learning path planning based on large-scale educational data analysis provides important support for the OMO teaching model. By constructing a knowledge point difficulty model and a learner state model, it can recommend personalized learning resources based on learners' online behavioral characteristics and improve learning efficiency. Educational data analysis also provides a scientific basis for teaching decisions, helping teachers identify learning differences and adjust teaching strategies. Future research can further explore the fusion analysis method of multi-source heterogeneous educational data, optimize the learning path planning algorithm, and promote the application and innovation of OMO teaching mode in a wider range.

Funding

1. Jiangxi Provincial Education Science”14th Five-Year Plan” Planning Project (22YB265).

2. Research Project on Teaching Reform in Jiangxi Provincial Higher Education Institutions (JXJG-23-24-16).

3. Science and Technology Project of Jiangxi Provincial Department of Education (GJJ2402521).

4. 2021 Provincial First-Class Undergraduate Professional Construction Project (Computer Science and Technology).

References

[1] Kong, L., & Han, D. (2025). Research on optimization strategy of dynamic planning method in resource balancing in online and offline integrated teaching in higher education. J. COMBIN. MATH. COMBIN. COMPUT, 125, 9337-9352.

[2] Guo, W., & Niu, Y. (2021). Interactive teaching system of basketball action in college sports based on online to offline mixed teaching mode. Mobile Information Systems, 2021(1), 9994050.

[3] Peng, X. (2023). University mixed teaching modes: Personalized online and offline integration. Adv. Educ. Technol. Psychol, 7, 134-140.

[4] Yang, N. (2020, November). Research on the Application of BOPPPS teaching mode in the interactive integration of online and offline teaching in colleges and universities during post-epidemic period. In 2020 Conference on Education, Language and Inter-cultural Communication (ELIC 2020) (pp. 265-268). Atlantis Press.

[5] Li, F. (2022). Mixed Teaching Mode for Online and Offline under COVID-19 Epidemic Situation. Frontiers in educational Research, 5(1), 16-24.

[6] Hong, G., Jie, X., & Qi, S. (2025). Research on the Mode Reform of Online and Offline Interactive Feedback in Internal Medicine Nursing Teaching Based on SPOC and Flipped Classroom. Frontiers in Educational Research, 8(1).

[7] Yu, S. (2024). Utilization of online and offline mixed education model in ESP teaching—taking financial English teaching as an example. Adv Educational Technol Psychol, 8(1), 267515584.

[8] Veeraiyan, D. N., Varghese, S. S., Rajasekar, A., Karobari, M. I., Thangavelu, L., Marya, A., ... & Scardina, G. A. (2022). Comparison of interactive teaching in online and offline platforms among dental undergraduates. International journal of environmental research and public health, 19(6), 3170.

[9] Yiming, L., Liru, B., & Zhen, L. (2021, April). The Exploration of the Reform Path of Integrating Teaching and Competition in Financial Market Course Based on the Mixed Online-offline Teaching Mode. In 2021 2nd International Conference on Big Data and Informatization Education (ICBDIE) (pp. 479-482). IEEE.

[10] Bian, L. (2021). Integration of “Offline+ Online” Teaching Method of College English Based on Web Search Technology. Journal of web engineering, 20(4), 1207-1218.

[11] Wu, Y., & Wang, J. (2022). Three-stage blended Chinese teaching online and offline for international students: a case study on Chinese teaching for international students in S university. Journal of Higher Education Research, 3(2), 207-211.

[12] Yu, H., Wang, S., Li, J., Shi, G., & Yang, J. (2022). Influence of online merging offline method on university students’ active learning through learning satisfaction. Frontiers in psychology, 13, 842322.

[13] Hazrati Nejad, E., Yigit-Sert, S., & Emrah Amrahov, Ş. (2024). An effective global path planning algorithm with teaching-learning-based optimization. Kybernetika, 60(3), 293-316.

[14] Shou, Z., Lu, X., Wu, Z., Yuan, H., Zhang, H., & Lai, J. (2020). On learning path planning algorithm based on collaborative analysis of learning behavior. IEEE Access, 8, 119863-119879.

[15] Xiao, Q., Zhang, Y. W., Xin, X. Q., & Cai, L. W. (2024). Sustainable Personalized E-Learning through Integrated Cross-Course Learning Path Planning. Sustainability, 16(20), 8867.

[16] Shi, Z., Wu, Z., Zhang, Z., Chen, Y., & Liu, X. (2022). Learning path planning algorithm based on career goals and artificial intelligence. International Journal of Emerging Technologies in Learning (iJET), 17(10), 256-272.

[17] Li, X. (2022). Path Planning for the Fragmented Learning of College Students Based on Artificial Intelligence. International Journal of Emerging Technologies in Learning, 17(19).

[18] Ji, L., Zhang, X., & Zhang, L. (2020, June). Research on the algorithm of education data mining based on big data. In 2020 IEEE 2nd International Conference on Computer Science and Educational Informatization (CSEI) (pp. 344-350). IEEE.

[19] Zhang, W., & Jiang, L. (2018). Algorithm analysis for big data in education based on depth learning. Wireless personal communications, 102(4), 3111-3119.

[20] Dai, Q. Y., Zhang, C. P., & Wu, H. (2016). Research of decision tree classification algorithm in data mining. International journal of database theory and application, 9(5), 1-8.

[21] Mughal, M. J. H. (2018). Data mining: Web data mining techniques, tools and algorithms: An overview. International Journal of Advanced Computer Science and Applications, 9(6).

[22] Liu, X. H., Zhang, D. G., Yan, H. R., Cui, Y. Y., & Chen, L. (2019). A new algorithm of the best path selection based on machine learning. IEEE Access, 7, 126913-126928.

[23] Lei, X., Zhang, Z., & Dong, P. (2018). Dynamic path planning of unknown environment based on deep reinforcement learning. Journal of Robotics, 2018(1), 5781591.

[24] Fan, Y. (2024). Integrating online and offline teaching to promote creativity for STEM learners. Educational Technology Quarterly, 2024(3), 241-254.

[25] Jin Wu. (2025) .Research on the Construction and Use of OMO+PBL Teaching Mode of College English with the Empowerment of Digital Intelligence. Exploration of Educational Management,3(1).

[26] Min Yu,Wenxin Xia,Linlin Zhang,Junnan Lv & Zhengqi Xin. (2024) .Research on the construction and effect evaluation of OMO mode of professional practical training teaching of landscape engineering technology based on virtual simulation technology. Applied Mathematics and Nonlinear Sciences,9(1).

[27] Changzhi Sun,Shijie Huang,Banghui Sun & Shiwei Chu. (2025) .Personalized learning path planning for higher education based on deep generative models and quantum machine learning: a multimodal learning analysis method integrating transformer, adversarial training and quantum state classification. Discover Artificial Intelligence,5(1),29-29.

[28] Peng Zhang & Qing Guo. (2025) .Security-reliability analysis in uplink cognitive satellite-terrestrial networks with LEO relaying. Computer Networks,264,111272-111272.

[29] Cristian Ramos Vera,Iraklis Grigoropoulos,Antonio Serpa Barrientos,Dennis Calle,Royer Olivera Cercado,Luis Hualparuca Olivera & Dritjon Gruda. (2025) .A multinational study of social attitudes, moral beliefs, and personality traits: A network analysis approach. Personality and Individual Differences,242,113216-113216.

[30] Yuyang Shen,Nicole Sankofa & Rachel U. Mun. (2025) .Enhancing Equity in Gifted Education: A Critical Content Analysis of Identifying Emergent Bilinguals in North Texas. Journal of Advanced Academics,36(2),359-395.