International Journal for Housing Science and Its Applications
Publish August 4, 2025. Volume 46, Issue 3 Pages 3056-3068

https://doi.org/10.70517/ijhsa463256

# Model Innovation Path and Practice of Student Management Work in Colleges and Universities Driven by Digital Technology

**Mu Mu**[1,*]
[1] School of Digital Arts, Suzhou Industrial Park Institute of Services Outsourcing College, Suzhou, Jiangsu, 215123, China
Corresponding authors: (e-mail: 18651110126@163.com).

**Abstract** Driven by digital technology, this study systematically constructs an innovative model for student management in colleges and universities, and proposes a full chain management methodology through demand analysis, functional modeling and dynamic behavior mining. Based on the actual needs of colleges and universities, a system architecture covering authority control, multi-role interaction and dynamic configuration capability is designed. The improved K-means clustering algorithm is used to classify student learning habits into clusters, and lagged sequence analysis (LSA) is used to reveal the temporal correlation of learning behaviors. The empirical analysis relies on the behavioral logs and questionnaire data of MOOC platform, and finds that the sequence of learning behaviors is significantly correlated with course grades, e.g., the correlation coefficient of participating in the assessment after reviewing is 0.244, and the correlation coefficient of negatively affecting the grades by overlearning the new content is -0.298. The characteristics of campus network use classify the students into academic focus type 384, recreational and social type 1,032, balanced multitasking type 1807, and light use type 1,000. 1807 and light-use type 169, and their traffic distribution showed significant differences. During the semester, there are key turning points in behavioral patterns (week 5 and week 12), and students' behaviors show stage-by-stage evolution, e.g., group C dynamically adjusts from “diligent mode” to “result-oriented”. Data-driven behavioral modeling and time-series analysis can provide a scientific basis for personalized resource recommendation, precise intervention and dynamic management strategy optimization, and help university student management transform to intelligence and refinement.

**Index Terms** student management in higher education, use case modeling, study habit modeling, K-means clustering, lagged sequence analysis

## I. Introduction

With the continuous development and application of digital technology, student management work in colleges and universities also ushered in new opportunities and challenges [1], [2]. The innovation of student management work in colleges and universities supported by digital technology can not only improve the management efficiency and optimize the management process, but also realize personalized service and promote the development of students [3], [4].

Traditional student management work often relies on manual operation and paper files, which is inefficient and prone to errors [5], [6]. Digital technology, however, can realize the rapid collection, organization and analysis of information, incorporate various types of student information into a unified management system, and realize the intelligent processing and query of information [7]-[9]. Digital technology can be used to centralize the management of students' personal information, academic information, employment information, etc., and data analysis through intelligent algorithms to provide customized services and support for the school and enhance the efficiency and accuracy of student management [10]-[13]. While in the traditional student management work, a large number of manual operations and process approvals are required, which are prone to the problems of information loss and process delays, through the use of digital technology, it is possible to establish the whole process of information management and process operation, and to realize the automation and standardization of the student management process [14]-[17]. Digital technology can be used to realize the online processing of student leave, course selection, graduation application and other processes, students can complete the relevant operations through cell phones or computers, real-time interaction and real-time updating of information, greatly improving the efficiency and accuracy of the process [18]-[21]. In addition, digital technology can realize the personalized service of student management in colleges and universities, as the development needs and personality characteristics of each student are different, the traditional student management mode often fails to meet the individualized needs [22]-[24]. And digital technology can analyze and mine students' multi-dimensional data to understand students' academic level, interests and strengths, development potential and other aspects of information, the growth process to find their own interests and potential, and support their personalized development [25]-[28]. Thus, providing students with personalized development direction and support measures [29].

This study focuses on the construction and practice of student management system driven by digital technology, and proposes a set of full chain methodology covering demand analysis, functional modeling, learning behavior mining and behavior sequence analysis. Based on the actual needs of student management system in colleges and universities, the system architecture is constructed from the perspectives of authority control, module division and dynamic configuration capability, clarifying the system's multi-role user requirements and functional module division, and emphasizing the core roles of authority management, data security and dynamic configuration. Secondly, we analyze the business process and interaction logic of each module by analyzing the three major use case models of students' daily management, quality management and system management. On this basis, improved K-means clustering is introduced to model learners' learning habits and divide them into groups, providing data support for personalized resource recommendation. Finally, combined with the lagged sequence analysis (LSA) method, the temporal correlation of students' learning behaviors is mined to reveal the potential behavioral patterns. The GSEQ tool is used to analyze the coding sequences of learning behaviors, and the residual values are calculated to reveal the significant behavioral transition patterns.

## II. Digital technology-driven innovation of student management model in colleges and universities

### II. A. General Requirements for Student Work Management System in Higher Education Institutions

The Student Work Management Platform is an on-campus system for approving and managing various types of student information. The main target users of the platform are on-campus students, faculty and staff as well as program developers, and the platform is planned to have five modules: notification and announcement, student management, financial aid management, accurate financial aid and system management.

Since the system involves confidential data information such as departmental and student data information, business application data and business report data, it has high requirements for authority division, distributed transaction processing and system security. System administrators can use the student work management platform to enable personal information modification applications and applications for the recognition of students with difficulties, as well as to make multi-dimensional data reports on various types of data in the system, and can also create, classify, modify, and delete the permissions of various user groups, roles, and roles in the system management module. The relevant approving personnel of each business line, such as student information modification and application for recognition of students in difficulty, can enter the corresponding module to approve the audit tasks flowing under their names, and at the same time, they can also view the statistics of the audit information under their current privileges.

For the student management module and financial aid management module, which support the main functions of the platform at the current stage, when the approval requirements change in the new academic year, the management personnel of the corresponding business lines can flexibly configure the approval settings to adapt to the changes in the application requirements in the new academic year, and at the same time, the project administrators can also configure the relevant notices, announcements and questionnaires corresponding to the project.

### II. B. Use case model analysis

On the basis of clear system functional requirements and architectural design, it is necessary to further analyze the user interaction logic in specific business scenarios through the use case model, so as to refine the implementation path of each module.

#### II. B. 1) Model analysis of daily student management

The activities of the student daily management use case model mainly include student daily hygiene management, student daily discipline management, student daily homework management, and student classroom management, which are maintained by the administrators, and the student daily management use case is shown in Figure 1.

#### II. B. 2) Analysis of student quality management models

The main activities performed by the student quality management use case model include student quality improvement management, student feedback information management, student daily assessment management, etc., which are mainly performed by students, as shown in Figure 2 of the student quality management use case.

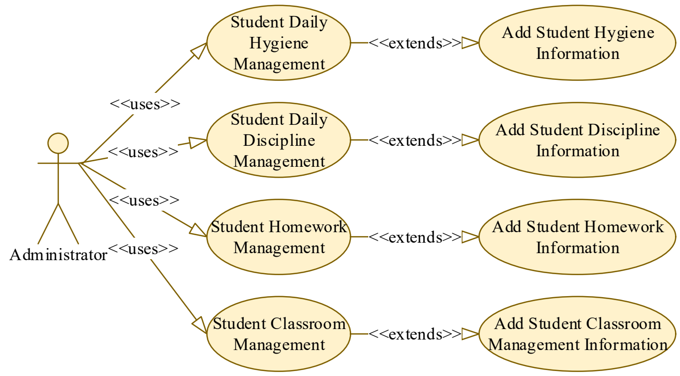


Figure 1: Student daily management use cases

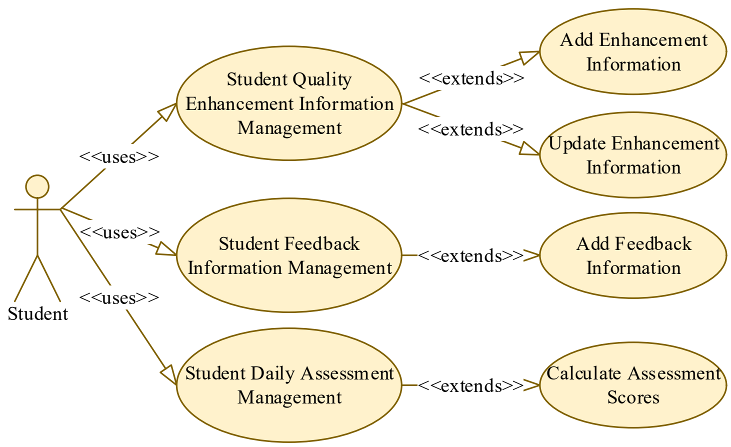


Figure 2: Student quality improvement management use case

**II. B. 3) System management model analysis**

The activities of the system management use case model mainly consist of administrator management and student feedback management, which are maintained by the administrators, and the system management use case is shown in Figure 3:

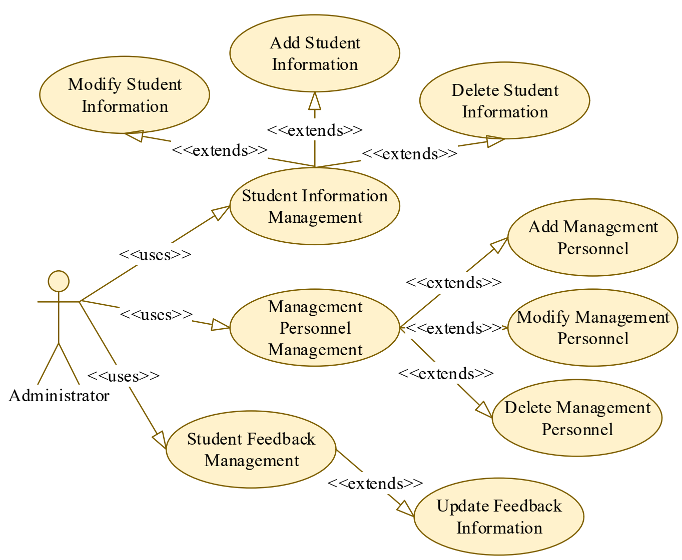


Figure 3: System management use cases

### *II. C.Learner Learning Habits Mining*

After completing the design of functional modules, how to effectively utilize the student behavior data accumulated by the system becomes critical. For this reason, this section focuses on the modeling of learners' learning habits and the division of clusters to provide data support for subsequent personalized services.

#### II. C. 1) Modeling Learner Learning Habits

There are several categories of learners' study habits with several values under each category, so in this paper, we will use vector space model learners' study habit model. The VSM definition of learning habit is shown in equation (1):

$$\begin{cases} V = \left\{ \overrightarrow{P_1}, \overrightarrow{P_2}, \cdots, \overrightarrow{P_k} \right\} \\ \overrightarrow{P_j} = \left\{ c_{j1}, c_{j2}, \cdots, c_{jn} \right\} \\ c_{ji} \in C_i, i \in [1, n] \end{cases} \tag{1}$$

where $V$ is the vector space of learning habits for all learners, $\overrightarrow{P_j}$ denotes the vector of learning habits for the $j$ rd learner, $n$ is the number of learning habit categories, $C_i$ is the value domain of the $i$ th learning habit category, and $c_{ji}$ is the $i$ th specific learning habit of the $j$ th learner.

Based on the results of processing XuetangXDataset, this paper obtains the following learning habit categories, in which the landing location uses simulated data, and all other learning habit categories are obtained by analyzing and mining the dataset.

#### II. C. 2) Learner Cohort Segmentation Based on Learning Habits

Learning habit grouping, i.e., according to learners' learning habits, learners with the same or similar learning habits will be divided into the same groups, and the results of grouping will be used in the learning resources recommendation session.

(1) Learning Habits Grouping Problem Description

The main task of this part is to divide learners into groups with similar hobbies according to their learning habits, so as to carry out the work of recommending friends, sharing resources and so on. Learning habit grouping problem can be abstracted as a clustering problem, clustering problem belongs to "unsupervised learning", the specific implementation method adopts the improved $K$ -mean algorithm.

$K$ mean, define the set of learners $S$ , the set of learners' habit features $P$ , the learners' learning habit feature vectors $\vec{P} = (p_1, p_2, p_3, \cdots, p_n)$ , where $p_i \in P$ . All the learners' habit features as a set of vectors are clustered using the K-means algorithm in order to minimize its cost function. The cost function is as in equation (2):

$$J(c, \mu) = \sum_{i=1}^{m} \left\| x^{(i)} - \mu_{c^{(i)}} \right\|^2 \tag{2}$$

where $m$ - the number of sample points, $x^{(i)}$ - the learning habit feature vector of the $i$ rd sample, $\mu_{c^{(i)}}$ - the plurality of the cluster to which the $i$ th sample belongs.

The more similar the learning habits of the learners in each cluster, the smaller the value of the sum of squared errors.

(2) Learning habit cluster division algorithm

In this paper, the habitual attributes of the learners are all discrete values, and the distance between the values has no practical significance, but can only talk about "similar" or "not similar". Therefore, when calculating the distance between two students, some changes to the original calculation method are needed. The algorithm for calculating the similarity distance is as follows:

The main idea of comparing similarity distances is to compare how many attributes are the same between two points. The more identical attributes, the more similar they are. Here the center of the cluster $c$ is not necessarily the real learner, because the center is chosen as the number of attributes in the current cluster, and it may be a cobbled together "virtual learner".

The idea of K-means is utilized to cluster learners according to their learning habits in the following way:

The overall process is similar to K-means, using similar distance instead of the original Euclidean distance, and calculating the center from taking the average to taking the plurality.

### *II. D.Lagged series analysis process*

The static characterization of learning habits lays the foundation for resource recommendation, while the mining of dynamic behavioral patterns requires more in-depth time series analysis. The application of lagged sequence analysis methods can reveal the potential correlation of learning behaviors from the time dimension and further improve the accuracy of management strategies.

Lagged Sequence Analysis (LSA) is a method proposed by Sackett for assessing the probability of occurrence of sequential behaviors over time, which is able to examine the significance of the resulting learning behaviors, analyze in depth the significance of the probability of occurrence of one behavior after the generation of another, and thus mine the learner's learning behavior patterns. It is generally carried out by the five steps shown in Figure 4 below.

The lag sequence analysis method mainly utilizes the GSEQ tool for analysis. In the GSEQ analysis process, first of all, we need to strictly follow the software format requirements for the coding of behavior, and then enter these codes, and then the system compiled to generate the MDS format file, to check whether the code is consistent, and then conduct the analysis of behavioral sequences, to get the coefficient table of behavioral sequences converted, which usually includes the frequency table and the residual value table, and then finally based on the size of the residual value to filter the behavioral sequences which have significant significance. Finally, the behavioral sequences with significant significance are filtered according to the size of the residual values, and the behavioral conversion diagram is drawn.

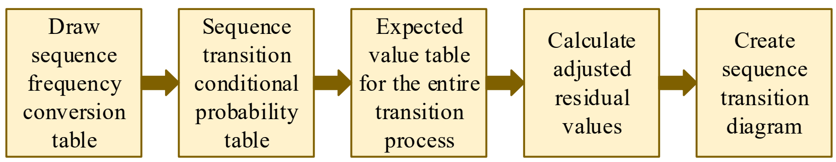


Figure 4: Lag sequence analysis process

In the process of lagging sequence analysis, the following two points need to be focused on and implemented: first, before the lag sequence analysis, it is necessary to code the behavior of students interacting with the learning platform, encode the students' behavior according to the determined coding rules, record the coding results of each student according to the chronological order, and generate the coding sequence. Second, after running the lag sequence analysis, it is necessary to calculate the standard residual value as shown in the following equation 1 through the running results, i.e., the standard residual value (Z-Score, $Z$ score), where $\sigma$ is the residual, the $Z$ score is used to indicate the significance of the transformation between the two behaviors, when the $Z$ score (standard residual value) is greater than 1.96, it means that the behavior transformation relationship is statistically significant, and the larger the $Z$ score, the stronger the transformation relationship between the two behaviors, i.e When the score of $Z$ is greater than 1.96, the two behaviors show continuity.

$$z - score = \frac{\sigma - \bar{\sigma}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\sigma i - \bar{\sigma})}} \tag{3}$$

## III. Data-driven empirical analysis of student behavior and optimization of management strategies in higher education based on data

The theoretical construction of digital technology-driven university student management system architecture design, use case model and learning behavior mining in Chapter 2 lays a methodological foundation for the subsequent empirical analysis. This chapter will verify the feasibility of the theoretical model from multiple dimensions based on actual data, and reveal the characteristics and laws of students' behaviors through data mining techniques to provide empirical support for the precise optimization of management strategies.

### *III. A. Data processing*

#### III. A. 1) Overview of data

The data for this study came from the open MOOC course “Basic Education Reform” in Chinese universities, and the data were categorized into five categories. These include registered student data, course resource data, course learning data, student performance related and student behavior data.

**III. A. 2) Data cleansing**

This study systematically analyzes and cleans the above five types of data, removes unreasonable data, adds values to the vacant data, and packages and organizes the recovered data according to the type of data analysis for later analysis.

For the data related to learner information, this study eliminates the learner IDs and registration data of teachers and teaching assistants for the purpose of checking the effect of course release and maintaining the normal operation of the team, and eliminates the duplicate IDs of learners who have “selected courses, withdrawn from courses, and selected courses” in order to prevent duplicated calculations at a later stage. The calculation is repeated.

For behavioral log data, this study proposes duplicate data provided by the MOOC platform, and takes all rounding for behavioral data not generated during the learning process. Rounding is also taken for the data that are very far from the relevance of this study. In order to prevent affecting the effect of the later study.

Data cleaning is built on the basis of the understanding of the MOOC course, and in the process of eliminating data, behavioral data that can not be judged are asked by multiple parties, including the teacher who runs the course, the participating students, and the technical maintenance personnel, etc., in order to ensure that the data are true and accurate, and the study is rigorous and reliable.

***III. B. Questionnaire Analysis of College Students' Study Time Commitment***

After completing the data cleaning and preprocessing, in order to further explore the characteristics of students' online learning behaviors, this study combines the questionnaire data to quantitatively analyze college students' study time commitment and reveal the potential association between it and course completion rate.

This study discusses the content structure of college students' online learning habits on the basis of theoretical analysis, literature research and open-ended questionnaire survey, and designs the online learning habits questionnaire from multiple dimensions, which mainly focuses on the dimension of college students' online learning time commitment.

The object of this survey is to study college students, the questionnaire takes the form of online random distribution, a total of 200 questionnaires released, 184 valid questionnaires, invalid questionnaires 16.

In the learning process of MOOC courses 67.54% of college students have given up the courses they studied, of which 32.50% gave up because they were too busy with things in the middle of the course and did not have time to continue the course. Then setting the length of the weekly study course is an important issue.

The weekly study time of college students is shown in Figure 5. The proportion of college students who devote less than half an hour or more than two hours to study each time is the smallest among the total number of respondents, accounting for 15.20% and 9.31% respectively, while the proportion of college students who devote between half an hour and one hour to study is the highest at 41.18%, and the proportion of college students who devote between one hour and two hours to study is relatively large at 34.31%.

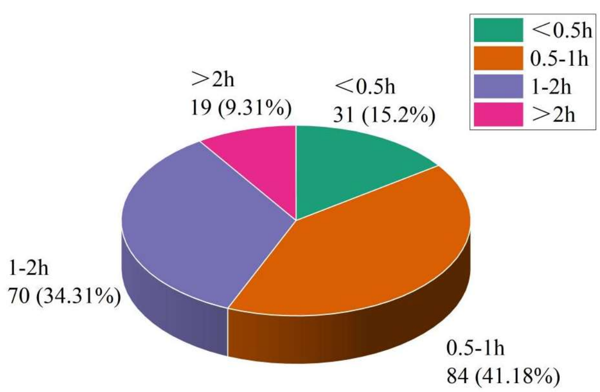


Figure 5: College students devote time to their studies every week

The data of college students' weekly online learning shows that 34.31% of college students are accustomed to investing 1 to 2 hours per week to study online courses, in other words, college students are accustomed to studying four lessons per week.MOOC course development platform can set the course as one lesson per day, and the length of each lesson is set at 20 minutes to 30 minutes which is more suitable.

### *III. C. Data mining based analysis*

Based on the preliminary conclusions of the questionnaire survey, this section further relies on the behavioral log data, through the frequency statistics and correlation analysis of behavioral sequences, to dig deeper into the dynamic relationship between learning behaviors and course grades, and to provide data support for the resource recommendation and intervention strategies.

#### III. C. 1) Descriptive analysis of behavioral sequences

Another commonly used data mining technique in behavioral analysis research is behavioral sequence analysis, which looks for important ways to show learners' learning paths through students' online behavior logs. Behavior sequence analysis will reflect the sequence of learners' behaviors, and when the amount of data is large and sufficient, it will be possible to mine the characteristics of group online behavior. Behavior sequence provides an important perspective for analyzing online learning.

This section of the study is based on the behavioral data that was previously data cleaned and labeled with behavioral labels, following a two-step behavioral sequence study approach. Firstly, the frequency table of behavioral sequence is drawn as shown in Table 1. Where the vertical column indicates the first step of the behavior and the horizontal row indicates the second step of the behavior. Each behavior is represented as B1. learning new content, B2. reviewing old content, B3. browsing to answer questions, B4. browsing and participating in forums, B5. participating in course assessment, B6. understanding course content, B7. asking for help, B8. interrupting, wandering and B9. behavioral noise. For example, a learner logs into the platform and first checks the information about the course. Then starts to learn to study the new chapter, then this behavior sequence is represented as (B6B1): a learner is working on the new course in all 3 analysis units, then this behavior sequence is represented as (B1B1). Corresponding to the behavioral frequency table, it indicates the frequency of the successive occurrence of the behavior, for example, the frequency of the occurrence of the behavior in which a learner moves from behavior label 2 to behavior label 1 (B2B1), i.e., the behavior of reviewing the old knowledge before proceeding to the new content, is 27010 times.

Table 1: Frequency of behavior sequence

| | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 |
|---|---|---|---|---|---|---|---|---|---|
| B1 | 83093 | 24188 | 9622 | 796 | 2106 | 7741 | 507 | 2298 | 1358 |
| B2 | 27010 | 34450 | 6336 | 921 | 9050 | 9144 | 501 | 1778 | 1697 |
| B3 | 6191 | 8401 | 7788 | 751 | 542 | 1995 | 102 | 307 | 722 |
| B4 | 582 | 1670 | 254 | 15418 | 944 | 1603 | 2169 | 174 | 441 |
| B5 | 1210 | 10843 | 213 | 2013 | 46787 | 13974 | 87 | 1660 | 1030 |
| B6 | 15038 | 16550 | 2153 | 1715 | 20607 | 20051 | 599 | 1068 | 1027 |
| B7 | 606 | 820 | 166 | 890 | 167 | 469 | 365 | 62 | 182 |
| B8 | 2015 | 1354 | 391 | 196 | 1524 | 1385 | 45 | 257 | 113 |
| B9 | 2128 | 1892 | 531 | 431 | 915 | 923 | 152 | 86 | - |

#### III. C. 2) Behavioral serial correlation analysis

After the basic analysis of the behavioral sequence table, the characteristics of learners' MOOC learning behavior can be found intuitively. However, the relationship between "learning behavior-learning habit-learning performance" in the learning process cannot be ignored. Learning behavior is the direct performance of learners in the learning process, observable, monitorable behavior, which is the external expression of the learner's internal habits, while the learning performance is the test of the learner's learning effect, in the MOOC course will be reflected in the course results of the learner whether to achieve the learning effect. Achievement is a problem avoided by many researchers, and people always say that "Chinese education" is a score-oriented education, which ignores the change of learners' internal learning process in order to improve the achievement. However, this problem has been improved in the learning process of MOOC courses, in which the grading criteria are set into three categories: unit test, class discussion, and final exam, and the final grades are converted into final grades according to different weights. Therefore, here the course grade is understood as learning performance, not only the score of the exam, but also covers the learners' participation behavior, taking both formative and outcome evaluations into account.

In this study, the correlation and significance of learning behavior sequences with course grades were calculated. The correlation coefficients between the behavioral sequences and grades are shown in Table 2.

Table 2: The correlation coefficient between behavioral sequence and performance

| | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 |
|---|---|---|---|---|---|---|---|---|---|
| B1 | -0.242*** | 0.153*** | 0.033 | 0.023 | -0.155*** | -0.298*** | -0.174*** | 0.049 | 0.038 |
| B2 | -0.203*** | 0.009 | 0.430*** | 0.043 | 0.244*** | 0.149*** | 0.052 | 0.059 | 0.027 |
| B3 | 0.056 | 0.425*** | 0.451*** | 0.013 | 0.033 | 0.004 | 0.056 | 0.038 | 0.026 |
| B4 | 0.068 | 0.010 | 0.201*** | 0.188*** | 0.015 | 0.023 | 0.048 | 0.038 | 0.224*** |
| B5 | 0.032 | 0.272*** | 0.131*** | 0.038 | 0.259*** | 0.035 | 0.034 | 0.042 | 0.056 |
| B6 | -0.313*** | 0.018 | 0.050 | 0.004 | 0.067 | -0.296*** | 0.046 | 0.145*** | 0.062 |
| B7 | -0.185*** | 0.045 | 0.031 | 0.035 | 0.017 | 0.012 | -0.126*** | 0.059 | 0.026 |
| B8 | 0.009 | 0.018 | 0.023 | 0.060 | 0.144*** | 0.009 | 0.074 | 0.068 | 0.015 |
| B9 | 0.070 | 0.018 | 0.159*** | 0.041 | 0.049 | 0.059 | 0.048 | 0.067 | 0.007 |

The serial relationships in the table are also vertical and then horizontal, for example, the correlation coefficient between the proportion of the behavioral sequence “5-2” in all the behavioral sequences of the learner and the learner's performance in the course is 0.244, which is positively correlated and significant at the Sig=0.001 level.

Overall, among the 24 sets of significantly correlated (Sig=0.001) data, 14 sets showed positive correlation and 10 sets were negatively correlated, indicating that the influence of learning behavior sequences on grades is bi-directional. Specifically, some of the behavioral sequences are significantly and positively correlated with grades, for example, the correlation coefficient of B2 (reviewing old content) → B5 (participating in the course assessment) is 0.244, the correlation coefficient of B5 → B5 (participating in the assessment continuously) is 0.259, and the correlation coefficient of B3 (skimming and answering the questions) → B3 (answering the questions continuously) is 0.451.These results suggest that participating in the assessment after reviewing old knowledge, behaviors such as repeated participation in appraisal or continuous problem solving may be effective in enhancing learning performance.

On the other hand, some of the behavioral sequences were significantly negatively correlated with achievement, for example, the correlation coefficient of B1 (learning new content) → B6 (understanding course content) was -0.298, the correlation coefficient of B1 → B1 (continuous learning of new content) was -0.242, and the correlation coefficient of B6 → B6 (repeating understanding of course content) was -0.296.This suggests that excessive concentration on new content learning or repeating understanding of the course framework may distract learners' effective energy and lead to lower performance instead. In addition, the correlation coefficient of B4 (participation in the forum) → B9 (behavioral noise) was 0.224, indicating that there is a potential correlation between irrelevant behaviors of forum participation, which may affect learning concentration.

It is worth noting that the correlation coefficient between B8 (interruption, distraction) and B5 (participation in the assessment) is 0.144, which is positive but less significant, probably reflecting that some learners can still recover their learning effect through the assessment after a short interruption. Overall, the correlation between behavioral sequences and performance reveals the optimization direction of learning path design, for example, students should be guided to form a closed loop between review and assessment to avoid ineffective and repetitive behaviors, so as to enhance the precision of management strategies.

### *III. D. Characterization of Campus Internet Usage by Various Student Groups*

In addition to learning behavior, campus network usage pattern is an important dimension of students' daily management. In this section, students are classified into different groups through cluster analysis, and combined with indicators such as flow rate and duration, to analyze the cyclical characteristics of their network behaviors, which provides a basis for differentiated management.

#### III. D. 1) Data processing

In order to objectively and accurately describe the characteristics of students' online behavior, combined with the periodicity of the actual campus study and life scenes, this study constructs the campus online behavior indicator system through mathematical and statistical methods, which is constructed from a total of three dimensions, namely, duration, traffic, and number of times. Duration is determined by on-line date, off-line date, and on-line time and off-line time; traffic is determined by incoming traffic and outgoing traffic; and number of times is determined by on-line date, on-line time, incoming traffic and outgoing traffic.

Based on the index system, this study obtained a total of 3392 students' statistical value data of multiple indicators, and after data cleaning, blank values and invalid data were eliminated, and 3940 valid data were retained for the next step of data mining analysis. Other data in school include statistical values of achievement data, campus card consumption data, book borrowing data, and physical fitness test data.

#### III. D. 2) Results of cluster analysis and basic information of various student groups

K-means method is chosen for clustering.K-means clustering first determines the optimal number of clusters.In this study, the optimal number of clusters is determined to be 4 classes by Hubert's index and cross-validation method, and the implementation process is mainly based on the mature algorithms of the Nbclust package and the factoextra package in R language. After determining the optimal number of clusters, the K-means algorithm package in R language was utilized to cluster analyze the principal component coefficient vectors of the functional type data of multiple research indicators of 3392 students, and to obtain the results of cluster grouping for each student.

Through cluster analysis, the 3,392 students were categorized into 4 groups.

Academically focused (384): high frequency use of academic resources, low entertainment traffic, concentrated in the daytime;

Entertainment-social type (1032 people): high traffic, focusing on entertainment and social applications, more use at night;

Balanced multitasking type (1807 people): medium traffic, mixed use of academic and entertainment, multiple devices and multiple time periods;

Light use type (169 people): low frequency, low traffic, basic apps mainly.

#### III. D. 3) Characterization of Campus Internet Usage by Various Student Groups

Examples of the amount of time spent on Internet traffic in a day for various student groups are shown in Figure 6.

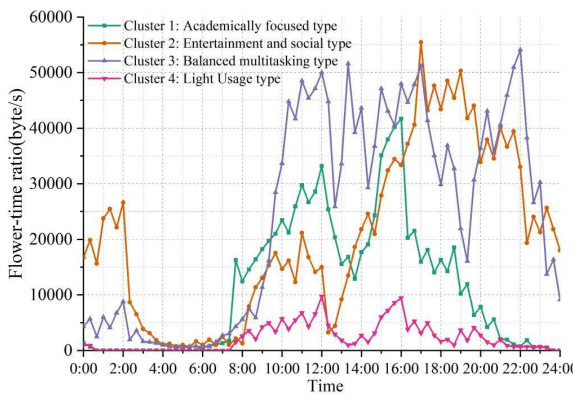


Figure 6: The ratio of online data usage time within a day for various student groups

In terms of the length of time spent on the Internet, academically focused students show a daytime efficiency pattern. This type of student group presents a traffic distribution pattern that is highly synchronized with the course schedule. Traffic peaks from 8:00 a.m. to 12:00 p.m. and from 2:00 p.m. to 4:00 p.m., accounting for 65%-75% of the total traffic throughout the day, and is mainly used for downloading courseware, accessing academic databases, or engaging in online courses. Midday (12:00-14:00) traffic plummets to 5-10%, with devices mostly dormant. At night (after 20:00), the utilization rate is extremely low, only occasionally accessing the academic system around 22:00 to check the schedule. The overall performance of the “sunrise and sunset” regularity, traffic fluctuations of more than 60%, reflecting a clear learning-oriented goals.

Cluster 2 Entertainment and Social Students are the late-night active type. The traffic of this group shows the typical characteristics of “night owls”, and the entertainment time from 12:00 to 2:00 a.m. contributes 80%-90% of the whole day's traffic, focusing on high-definition videos, multiplayer online games and social media interactions. Especially on weekends, the peak hour may extend to 3am, and the traffic volume in a single day can be 2-3 times of that on weekdays. Daytime (8:00-12:00), on the other hand, only maintains social software background message synchronization and occasional course study needs, with less than 20% of traffic, creating a peak-to-valley difference of over 85%. This type of usage pattern is often associated with lower course pressure or loose time management.

Cluster 3 students were mixed multitasking, showing a multimodal balanced type. As the most complex group of network users, its traffic distribution presents a unique "multimodal humpback curve": the learning peak is formed from 10:00 to 12:00 during the day, and literature retrieval, online learning and instant messaging are carried out simultaneously; From 13:00 to 15:00 in the afternoon and from 20:00 to 23:00 in the evening, switch to the entertainment peak to focus on video, games, and social content. Even after 1 a.m., there are about 5% of idle downloads or nighttime social activities. This all-weather, multi-scenario adaptive usage mode reflects about 50% peak-to-valley difference and strong task switching ability.

Cluster category 4 students are light-users, and the online behavior of these users is characterized by short, low-frequency pulses. Traffic is dispersed throughout the day in two micro-peaks between 8:00 am-12:00 pm and 14:00 pm-17:30 pm required for course study. At night (after 21:00), there is almost no network activity, and the peak-to-valley difference throughout the day is less than 30%, reflecting the lowest degree of network dependence, which mainly meets the basic campus learning information acquisition needs.

### *III. E. Characterization of the temporal evolution of student behavior*

To further capture the dynamic evolution pattern of student behavior, this section introduces a time series turning point identification algorithm to delineate the behavioral patterns at different stages of the semester and compare the strategic adjustments of each group at the critical nodes, which provides a scientific reference for stage-specific interventions.

#### III. E. 1) Identification of Turning Points in Learning Behavior Patterns

The analysis of student behavior in the MOOC database continues. In this paper, the coding sequences are analyzed using a time series turning point identification algorithm to identify turning points in the behavioral patterns. This paper compares turning points identified by total frequency sequences. By identifying the turning week for each student, the number of people who experienced a pattern turn in each week can be calculated, and the results of the number of times each week was identified as a turning point are shown in Table 3.

Table 3: The number of times each week is identified as a turning point

| Identification through the coding sequence | | Identification through the total frequency series | |
|---|---|---|---|
| Week | The number of times identified as turning points | Week | The number of times identified as turning points |
| Week1 | 0 | Week1 | 0 |
| Week2 | 1393 | Week2 | 1521 |
| Week3 | 2025 | Week3 | 1834 |
| Week4 | 1119 | Week4 | 2180 |
| Week5 | 3428 | Week5 | 2481 |
| Week6 | 2244 | Week6 | 2227 |
| Week7 | 660 | Week7 | 2169 |
| Week8 | 1538 | Week8 | 2351 |
| Week9 | 3418 | Week9 | 2322 |
| Week10 | 1204 | Week10 | 1961 |
| Week11 | 2816 | Week11 | 1242 |
| Week12 | 4091 | Week12 | 2662 |
| Week13 | 2287 | Week13 | 2506 |
| Week14 | 1460 | Week14 | 1968 |
| Week15 | 596 | Week15 | 855 |
| Week16 | 0 | Week16 | 0 |

Week 12 was identified as a turning point in both the coding sequence and the total frequency sequence. According to the time series turning point algorithm, week 12 was identified the most times and was the first turning point. Week 5 was chosen as the second turning point because week 5 was identified more often in both the coding sequence and the total frequency sequence, and because weeks 5 and 12 allow for a more even division of the semester compared to weeks 9 or 13. Therefore, in this case study, weeks 5 and 12 were used as the final pattern turning points. The entire semester was divided into three phases with different behavioral patterns:(1) weeks 1-5; (2) weeks 6-12; and (3) weeks 13-16.

#### III. E. 2) Types of Learning Behavior Patterns by Stage

In this paper, a time series clustering algorithm is used to cluster and analyze the coding sequences. The students in this case were divided into a total of 3 clusters, and each cluster of students behaved differently in each phase. Figures 7, 8, and 9 demonstrate the behavioral patterns of each cluster after clustering. For ease of interpretation, each cluster is visualized using a sequence of five behaviors rather than a coded sequence. The five behaviors are resource access, teacher-student interaction, assessment, self-directed learning, and grade verification. The average values of the five behavioral time series for each phase of the three clusters are shown in Figures 7, 8, and 9.

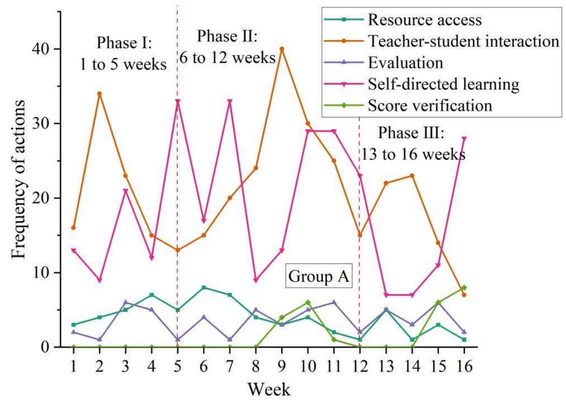


Figure 7: The values of the time series of five behaviors in each stage of Group A

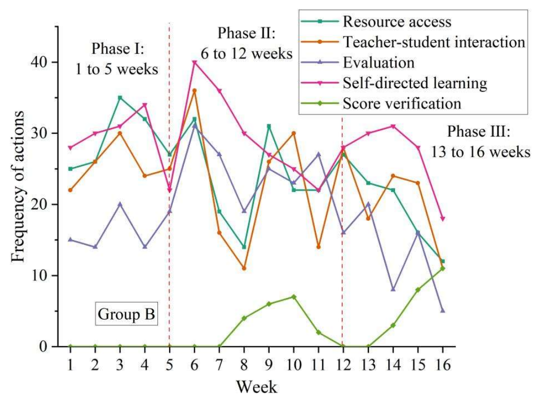


Figure 8: The values of the time series of five behaviors in each stage of Group B

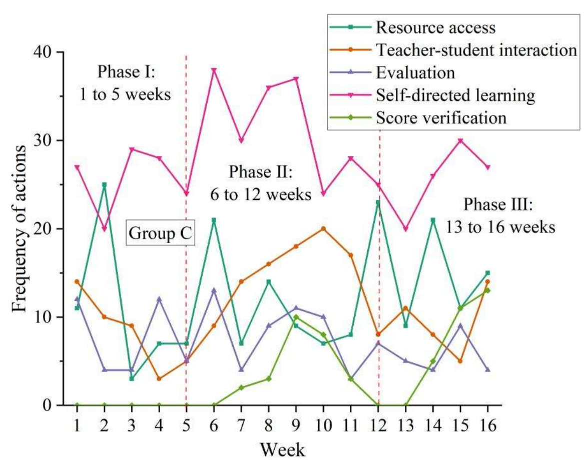


Figure 9: The values of the time series of five behaviors in each stage of Group C

Group A: Gradual decline in engagement. In Weeks 1-5, Group A was defined as an “interaction mode”, with primary behaviors focused on student-teacher interactions (21 peaks), but with significantly lower resource access (5 peaks) and assessment participation (5 peaks). The frequency of interactions continued to decline after peaking at week 2, indicating an initial reliance on social learning but a lack of persistence. After weeks 6-12, the group shifted to a “passive mode,” with overall low frequency of behavior (e.g., 3.6 weekly visits to resources), with only brief spikes in weeks 9 and 11 (e.g., 40 teacher-student interactions and 33 independent studies), reflecting a passive response to course milestones (e.g., exams or deadlines for assignments). By weeks 13-16, behaviors further weakened to a “low engagement pattern” with no more than 6 peaks in all behaviors (e.g., 28 independent study sessions only during the final rush), and the frequency of grade checks rose to 8, indicating that they only temporarily responded to final stress through ad hoc efforts.

Group B: Stable and balanced. Cohort B demonstrated a “balanced pattern” in weeks 1-5, with peaks of 33, 29, and 26 for resource visits, interactions, and assessments, respectively, with a synchronized trend (reaching a peak in week 2 and then declining steadily). This equilibrium continued through weeks 6-12, with all three remaining highly consistent (20 peaks for assessment and 15 peaks for interaction), despite a slight decrease in the frequency of behaviors (e.g., 25 peaks for resource access). At the end of the period (weeks 13-16), its pattern is further consolidated, with resource access (peak 22) and assessment (peak 18) remaining central, the frequency of self-directed learning increasing to 12 (weekly average of 3), and verification of grades rising to 8, demonstrating a systematic focus on learning outcomes. The stability of Cohort B indicates mature self-management skills and the ability to adapt to the needs of different stages of the course.

Cohort C: Dynamically adjusted. In weeks 1-5, Group C was in “diligent mode”, with high frequency of participation at the beginning (28 resource visits and 24 assessments), but the frequency of behavior dropped rapidly after the second week, reflecting its characteristic of “high start, low finish”. In weeks 6-12, the mode shifted to “interaction mode”, with teacher-student interaction becoming the core (peak 26 times), but resource access (10 times) and assessment (12 times) decreased significantly, indicating a shift to socialization-driven strategies in the midterm. At the end of the period (13-16 weeks), the group further evolved into a “low interaction mode”, with interaction frequency plummeting to 5, resource visits (18) and assessments (16) becoming dominant, and verification of grades increasing to 9 (the highest for any of the three groups). This shift from “fully engaged” to “socially dependent” to “results-oriented” shows the flexibility to adjust strategies according to the stage of the course, but also reveals a tendency to isolation at a later stage.

## IV. Conclusion

This study reveals the key patterns of student management in universities driven by digital technology through data mining and empirical analysis.

Behavioral sequence analysis showed that participation in assessment after reviewing old content (B2→B5) was significantly positively correlated with course grade ($r=0.244$, $p<0.001$), whereas sustained learning of new content (B1→B1) was significantly negatively correlated with grade ($r=-0.242$, $p<0.001$), suggesting that over-focusing on the new content may distract from the learning efficiency.

Cluster analysis divides students into four categories, in which the recreational and social group (1,032, 30.4%) accounts for 80%-90% of nighttime traffic, and the academically focused group (384, 11.3%) accounts for 65%-75% of daytime academic traffic, providing a direct basis for differentiated resource scheduling.

Time-series turning point identification indicated that weeks 5 and 12 were critical points in the semester, and that Cohort C (dynamically adjusted) increased the frequency of independent study to a weekly average of 18 during the final phase, and the verification of grades increased to 9, highlighting the need for staged interventions (e.g., intensive revision guidance in week 5, and enhanced appraisal support in week 12).

## References

[1] Yan, H. (2024). Strengthening of Student Management System on School Spirit Building Based on the Construction of Digital Education Resources. The Frontiers of Society, Science and Technology, 6(1).

[2] Koh, J. H. L., & Kan, R. Y. P. (2021). Students’ use of learning management systems and desired e-learning experiences: Are they ready for next generation digital learning environments?. Higher Education Research & Development, 40(5), 995-1010.

[3] Li, W. (2021). Design of smart campus management system based on internet of things technology. Journal of Intelligent & Fuzzy Systems, 40(2), 3159-3168.

[4] Xiang, X. (2022). Construct a new pattern of digital education management in colleges and universities in the era of media integration. International Journal of New Developments in Education, 4(09), 91-95.

[5] Zhang, Y. (2022). Challenges and Strategies of Student Management in Universities in the Context of Big Data. Mobile Information Systems, 2022(1), 3537468.

[6] Rahayu, I. A., Hardina, D., & Ninggrum, T. S. (2024). Self management efforts for working students. European Journal of Psychological Research, 11(3), 72-81.

[7] Yanti, D., & Syahrani, S. (2022). Student management STAI rakha amuntai student tasks based on library research and public field research. Indonesian Journal of Education (INJOE), 2(3), 252-256.

[8] Li, B. (2022). Research on information construction of college student management under the background of internet. In MATEC Web of Conferences (Vol. 365, p. 01012). EDP Sciences.

[9] Purnamasari, R., Suchyadi, Y., Karmila, N., Nurlela, N., Mirawati, M., Handayani, R., ... & Kurnia, D. (2020). Student center based class management assistance through the implementation of digital learning models and media. Journal of Community Engagement (JCE), 2(2), 67-70.

[10] Shen, Y., & Zheng, X. (2022, August). Research on the application of computer Internet digital technology in the student work smart management systems. In 2022 International Conference on Computers, Information Processing and Advanced Education (CIPAE) (pp. 53-61). IEEE.
[11] Songsom, N., Nilsook, P., Wannapiroon, P., Fung, C. C., & Wong, K. W. (2019). System design of a student relationship management system using the internet of things to collect the digital footprint. International Journel of Emerging Technologies in Learning, 14(23), 125-140.
[12] Ji, D., Wang, X., & Zhang, T. (2022). Application of New Media in Student Management from the Perspective of Deep Learning and Evaluation and Analysis of Practical Effects. Computational and Mathematical Methods in Medicine, 2022(1), 1765448.
[13] Qing, L. (2024). Research on the Innovation of Student Education and Management in Colleges and Universities under Network Environment. International Journal of New Developments in Education, 6(5).
[14] Mosteanu, N. R. (2020). Digital university campus–change the education system approach to meet the 21st century needs. European Journal of Human Resource Management Studies, 4(4).
[15] Wu, R., & Wang, J. (2023). Digital Management of Teaching Cases in Colleges and Universities Based on Cluster Analysis. International Journal of Emerging Technologies in Learning, 18(10).
[16] Yang, L., Geng, C., & Zhang, Y. (2023). Construction of Intelligent Platform for College Student Management under the Background of Digital Transformation. Journal of Intelligence and Knowledge Engineering (ISSN: 2959-0620), 1(1), 55.
[17] Fang, Y., Keyuraphan, L., Sutheeniran, N., & Dechhome, P. (2024). The Purpose Model for Improving Digital Management in Universities in Guangxi. World Journal of Education, 14(1), 1-11.
[18] Ekasari, S., Manullang, S. O., Syakhrani, A. W., & Amin, H. (2021). Understanding Islamic Education Management in Digital Era: What Experts Say. Nidhomul Haq: Jurnal Manajemen Pendidikan Islam, 6(1), 127-143.
[19] Galustyan, O. V. (2015). Digital Campus as Electronic Image of the University. Rupkatha Journal on Interdisciplinary Studies in Humanities, 7(3), 263-270.
[20] Hazaparu, V., Paraschivescu, A. O., & Bontas, O. (2023). Benefits of Implementing Digital Technologies in Higher Education Management. Economy Transdisciplinarity Cognition, 26(2), 34-39.
[21] Kowch, E. (2021). Leading transformation with digital innovations in schools and universities: Beyond adoption. Digital transformation of learning organizations, 145-168.
[22] Ali, M., Siddique, A., Aftab, A., Abid, M. K., & Fuzail, M. (2024). AI-Powered customized learning paths: transforming data administration for students on digital platforms. Journal of Computing & Biomedical Informatics, 6(02), 195-204.
[23] Wu, J. (2024). Research on the Optimization of University Management Service Systems Based on Student Needs. Journal of Modern Education and Culture, 1(1).
[24] Chang, Y. C., Li, J. W., & Huang, D. Y. (2022). A personalized learning service compatible with moodle e-learning management system. Applied Sciences, 12(7), 3562.
[25] Maintz, J. (2018). Smart Learning Environments: Integrating User Consent for a Responsible Data Management when Offering Personalized Learner Services. CSR und Hochschulmanagement: Sustainable Education als neues Paradigma in Forschung und Lehre, 107-113.
[26] Yang, Q. (2019). Research on Active Service Model with Personalized Education Resources. Asian Journal of Contemporary Education, 3(1), 95-104.
[27] Ellikkal, A., & Rajamohan, S. (2024). AI-enabled personalized learning: empowering management students for improving engagement and academic performance. Vilakshan-XIMB Journal of Management.
[28] Syed, T. A., Palade, V., Iqbal, R., & Nair, S. S. K. (2017, November). A Personalized Learning Recommendation System Architecture for Learning Management System. In KDIR (pp. 275-282).
[29] Paz-Albo, J. (2017). Is personalized learning the future of school?. Childhood Education, 93(4), 295-299.