International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-17

https://doi.org/10.70917/ijcisim-2026-0135
Article

Recognition and Enhancement of College Students' Self-Directed Learning Behaviors Based on Multilayer Perceptual Machines

Lanyan Yang *

Computer and Information Engineering College, Guizhou University of Commerce, Guiyang, Guizhou, 550014, China; ylanyang12@sina.com

Abstract: This paper proposes a method for identifying college students' autonomous learning behaviors based on a multi-layer perceptron. The process of identifying and analyzing college students' autonomous learning behaviors is constructed from four aspects: data collection, data preprocessing, data analysis, and online learning outcomes. Based on the characteristics of college students' online autonomous learning behaviors, the multi-layer perceptron method is employed to assess the performance of their autonomous learning behaviors. Learning behavior variables are selected from three dimensions: behavioral engagement, social engagement, and cognitive engagement. The impact of each learning behavior variable on learning outcomes is analyzed. Four key factors are identified: video viewing count, assignment practice count, unit test count, and participation count. Based on the performance levels of these four key factors, different learning modes are proposed, and the self-directed learning behavior performance of students in each learning mode is analyzed. The correlation coefficient between behavioral engagement and online learning outcomes is 0.724, indicating a high correlation. Among the four learning modes, "rhythmic-type" learners maintained stable learning efforts and good unit test scores throughout the course, with stable autonomous learning behavior performance and outstanding autonomous learning behavior capabilities. Enhancing college students' learning behavior can cultivate students to develop "rhythmic-type" learning behavior.

Keywords: multi-layer perceptron; autonomous learning; learning outcomes; learning behavior variables

1. Introduction

Learning behavior serves as a crucial basis for assessing learners' learning status, and promptly identifying abnormal learning behaviors is key to improving learning outcomes [1]. In traditional face-to-face learning settings, experienced teachers can accurately judge abnormal learning behaviors in real-time by observing learners' facial expressions, gestures, and other overt characteristics, and promptly intervene and correct such behaviors to ensure smooth learning progress [2-4]. However, in large-scale online learning environments represented by MOOCs, traditional observation methods are ineffective, and abnormal learning behaviors accumulate covertly, becoming a significant cause of severe learning issues or even learning failure [5]. In a large-scale teaching context, self-directed learning has become a normalized learning method [6]. Self-directed learning is an urgent need for cultivating innovative and practical talent. Self-directed learning behaviors vary from person to person, exhibit diversity, real-time characteristics, and uncertainty, with no clear boundary between "normal" and "abnormal," making it difficult to distinguish them using static, deterministic methods [7].

To address this issue, domestic and international scholars have preliminarily established a research framework comprising four aspects: self-directed learning behavior recognition technology, evaluation indicator system construction, evaluation method research, and application scenario research [8-10]. Literature [11] established a model for identifying students' autonomous learning behavior in a MOOC teaching environment to predict their learning performance, and compared multiple linear regression (MLR), multi-layer perceptron (MLP), and classification and regression trees (CART), achieving good recognition results. Literature [12] used educational data mining techniques to investigate learning behaviors in MOOCs, finding that students could be classified based on their learning styles, which could be predicted from their behaviors, and assessed whether students were suitable for MOOC self-directed learning methods. Literature [13] linked observable learning behaviors of MOOC students with underlying cognitive attributes, finding that interactive learning is associated with self-directed learning outcomes. This study provides insights for improving self-directed learning in MOOCs. Literature [14] analyzed interaction data in MOOC online classrooms using clickstream analysis and unsupervised learning methods to explore learner behavior in MOOC self-directed learning environments, and compared learning behaviors in two courses (Engineering CS101 and Humanities and Statistics Learning). Reference [15] proposes a new framework for assessing learning engagement in large-scale online courses using facial expression recognition and introduces a new domain-adaptive learning behavior method, which has proven its effectiveness and outperforms existing methods. Reference [16] proposes a hybrid neural network model for detecting student learning behaviors in MOOCs, aiming to enhance the efficiency of large-scale online learning by providing personalized recommendations.

This paper analyzes the essence and characteristics of self-directed learning and discusses the advantages and significance of self-directed learning for college students. To further identify college students' self-directed learning behaviors, a method for identifying college students' self-directed learning behaviors based on a multi-layer perceptron is designed in combination with a multi-layer perceptron. Correlation analysis methods are used to analyze the factors influencing online learning outcomes. Learning behavior variables are selected based on three dimensions: cognitive engagement, social engagement, and behavioral engagement. On this basis, the relationships between each learning behavior variable and learning outcomes are explored. Four factors—video viewing count, assignment practice count, unit test count, and discussion post count—are selected for analysis. A learning pattern analysis is conducted on these four factors to investigate how college students' self-directed learning behaviors manifest under different learning patterns.

2. Independent Learning Behavior among College Students

2.1. The Essence and Characteristics of Autonomous Learning

Learning is a conscious act that represents active participation and proactive acquisition. From the perspective of human development, learning is both a necessary means of individual survival and a process that promotes maturity and enhances quality. The content of learning can include knowledge, skills, or experience, and its scope can encompass a single book, an individual, or society as a whole. Such learning is conscious, proactive, and positive; indeed, learning itself is an autonomous process.

At the university level, learning is still primarily focused on textbook knowledge. Textbook knowledge represents previous generations' understanding of objective phenomena. University students learn textbook knowledge by absorbing the experiences of authors and teachers, processing, digesting, and assimilating these experiences through their own efforts. This transforms the knowledge into their own, enabling them to gain individual experience—a process of converting indirect experience into direct experience. From the perspective of knowledge acquisition, the essence of self-directed learning is the process by which students generate, realize, and develop knowledge on their own.

In this process, autonomous learning begins when external knowledge and experiences are selected by students and incorporated into their cognitive domain. In other words, only knowledge that attracts a student's attention can be selected and recognized. Therefore, the conscious attention students pay to the knowledge and information they are learning marks the beginning of autonomous learning, which is the stage of knowledge generation. To promote autonomous learning, the content of "book" knowledge must align with students' intrinsic needs. Only attention to knowledge driven by intrinsic needs leads to selective learning, which is autonomous learning.

The process of knowledge realization involves students learning, understanding, and assimilating the information and knowledge they have selectively noticed, forming and establishing new knowledge. This is a critical stage in the transformation of indirect experience into direct experience. This process is directly constrained by the individual student's learning ability.

The most important and highest level of manifestation of autonomous learning lies in students' ability to transcend existing truthful knowledge and create thinking activities that can guide practice and meet their own needs. In such thinking activities, the memory information repository in students' minds is fully mobilized, information is fully activated, knowledge systems are fully organized, and students' goal values are fully realized. This stage of autonomous learning is referred to as the development stage of knowledge and represents the highest realm of autonomous learning.

Essentially, this process of exploration, selection, construction, and creation mirrors the general process by which students learn and master knowledge. Therefore, it can be said that the essence of self-directed learning is the process of self-generation, realization, and development of knowledge [17-18].

2.2. The Significance of Independent Learning for College Students

First, the mission of education in today's era is to enable students to "learn how to learn," with self-directed learning at its core. In the context of humanistic education emphasizing a "student-centered" approach, self-directed learning among college students has taken on new significance. Broadly speaking, self-directed learning among college students should be oriented toward the development of society and human progress, with these as its goals, enabling students to exercise their autonomy in the learning process. In a narrow sense, university students' autonomous learning refers to the acquisition of additional knowledge through self-directed learning outside of classroom instruction. Autonomous learning serves as an effective means to address disparities among students.

Second, university students possess a broad range of choices in autonomous learning, enabling them to better fulfill the "self-actualization" needs emphasized by Maslow.

Third, the continuous breakthroughs and self-transcendence in learning achieved through autonomous learning contribute to students' personal development and creativity. College students' thinking abilities are relatively mature, having reached the advanced stage of formal thinking development.

Finally, self-directed learning not only cultivates the development of college students' learning abilities but also helps foster their sense of autonomy and independent spirit. Humanism emphasizes the individual's sense of agency, and autonomy is the most fundamental and core aspect of this agency. The immense energy unleashed by college students' sense of autonomy motivates them to actively, enthusiastically, and fully engage in understanding and transforming the objective world.

3. A method for Identifying College Students' Autonomous Learning Behavior Based on Multilayer Perceptrons

This paper proposes a method based on multilayer perceptron to identify college students' autonomous learning behavior. The main steps of this method include feature extraction of college students' autonomous learning behavior, construction and learning of multilayer perceptron model, and application of multilayer perceptron model to identify college students' autonomous learning behavior.

3.1. Basics of Multi-Layer Perceptrons

A multi-layer perceptron (MLP), also known as an artificial neural network, is a mathematical tool constructed using a hierarchical structure similar to that of the human brain's neural network [19]. The units within each layer include neurons, activation functions, and loss functions, among others. We will now introduce each of these components in detail.

A neuron is the basic computational unit of a neural network. It receives inputs from other neurons or external data and then calculates an output. Each input value has a weight, and the size of the weight represents the importance of that input relative to other input values. The output value is determined by the activation function inherent to the neuron.

An activation function is a function added to an artificial neural network to help the network learn complex patterns in the data. Activation functions are generally non-linear, and it is precisely because of these non-linear activation functions that neural networks possess powerful function approximation capabilities, enabling them to fit various curves. Typically, activation functions have the following characteristics:

① They are continuous and differentiable, or only a few points are non-differentiable, allowing numerical optimization methods to be directly used to learn the network's parameters.

② The activation function and its derivative should be as simple as possible. Since the network contains a large number of neurons, complex activation functions can significantly impair the network's performance.

③ The domain of the derivative of the activation function should be within an appropriate range; otherwise, it will affect the efficiency and stability of training. Selecting an appropriate activation function has a significant impact on the accuracy of this experiment.

With the perceptron model, accurate neural network parameters are also required to obtain an accurate model. After the neural network outputs the calculation results, the loss function calculates the difference between the result and the actual value, thereby guiding the next step of training in the correct direction. The purpose of training a multilayer perceptron is to obtain the closest output to the overall input of the training set, rather than minimizing the error for a single sample. Therefore, after one training iteration, the model typically calculates the overall loss function value for all samples to determine whether the network's error has been trained to an acceptable state. x represents a sample, and the loss function $E(y, f(x))$ can be used to measure the degree of inconsistency between the true value y and the predicted value $f(x)$, with smaller values generally being better.

3.2. Calculation Principles of Multi-Layer Perceptron Models

Generally speaking, deep learning training requires data volumes in the tens of thousands or even millions, with data dimensions that are more diverse. When using limited learning behavior data to extract actual data samples and applying deep neural networks with strong feature extraction capabilities, overfitting may occur. Therefore, this paper selects a multi-layer perceptron network to establish a shield construction excavation parameter prediction model. Among them, x_1, x_2, x_3 and x_4 are the raw data input into the multi-layer perceptron.

The backpropagation (BP) algorithm is a learning algorithm suitable for multilayer perceptrons. In the structure of the backpropagation algorithm, the first layer is the input layer, containing two neurons i_1, i_2 and a bias b_1. The second layer is the hidden layer, containing two neurons h_1, h_2 and a bias b_2. The third layer is the output layer, containing two neurons y_1 and y_2, and $w_i (i = 1, 2, \cdots, 8)$ are the weights between layers, with the rectified linear unit (ReLu) as the activation function.

(1) Forward propagation of data flow. The weighted sum of inputs to neuron h_1 is s_{h1}:

$$s_{h1} = w_1 i_1 + w_2 i_2 + b_1 \tag{1}$$

The output of neuron h_1 is:

$$o_{h1} = ReLu(s_{h1}) \tag{2}$$

The calculation for other neurons in the hidden layer and neurons in the output layer is similar.

(2) Backward propagation of error information. The total output error is:

$$E = \sum_{i=1}^{n} E_i \tag{3}$$

$$E_i = \sum_{i=1}^{n} \frac{1}{2} (\hat{y}_i - y_i)^2 \tag{4}$$

Among them, E_i is the prediction error of the i th output unit. $\hat{y}_i$ is the predicted value of the i th output unit, and y_i is the i th output unit. n is the number of neurons in the output layer.

Taking w_5 as an example, the weights from the hidden layer to the output layer are updated. The impact of w_5 on the overall error is:

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial y_1} \cdot \frac{\partial y_1}{\partial s_{y1}} \cdot \frac{\partial s_{y1}}{\partial w_5} \tag{5}$$

Among them, s_{y1} is the weighted sum of the inputs of y_i. The weight of w_5 is updated as follows:

$$w_5^+ = w_5 - \eta \frac{\partial E}{\partial w_5} \tag{6}$$

Among them, η is the learning rate.

Taking w_1 as an example, update the weights between the input layer and the hidden layer. That is:

$$\frac{\partial E_1}{\partial o_{h1}} = \frac{\partial E_1}{\partial o_{y1}} \cdot \frac{\partial o_{y1}}{\partial s_{y1}} \cdot \frac{\partial s_{y1}}{\partial o_{h1}} = (y_1 - \hat{y}_1) w_5 \tag{7}$$

$$
\begin{aligned}
\frac{\partial E}{\partial w_1} &= \frac{\partial E}{\partial o_{h1}} \cdot \frac{\partial o_{h1}}{\partial s_{h1}} \cdot \frac{\partial s_{h1}}{\partial w_1} \\
&= (\frac{\partial E_1}{\partial o_{h1}} + \frac{\partial E_2}{\partial o_{h1}}) \cdot \frac{\partial o_{h1}}{\partial s_{h1}} \cdot \frac{\partial s_{h1}}{\partial w_1} \\
&= (y_1 - \hat{y}_1) w_5 i_1 + (y_2 - \hat{y}_2) w_7 i_1
\end{aligned} \tag{8}
$$

$$
w_1^+ = w_1 - \eta \frac{\partial E}{\partial w_1} \tag{9}
$$

The above is the computational approach of the backpropagation algorithm. After the weights are updated, they are recalculated using the new weights, and the process is iterated until the error decreases to within the acceptable range.

3.3. Extracting Characteristics of College Students' Independent Learning Behavior

This paper primarily employs four processes: data collection, data processing, process mining, and knowledge representation, to mine and analyze learners' online self-directed learning behavior patterns. The framework for mining online self-directed learning behavior patterns is shown in Figure 1.

(1) Data collection: The network can capture learners' implicit learning behaviors, reflecting the logical relationships between their actual thoughts and actions. The collection of online learning behavior data is the foundation of this study, which is conducted on the Moodle platform. The Moodle platform automatically records and stores learners' online learning logs, including timestamps, usernames, event names, and other data, which are essential for obtaining information about learning behavior patterns.

(2) Data Processing: The learner log data recorded in the Moodle platform is detailed but redundant. This study requires cleaning and filtering these data according to its research needs. Suspected erroneous or abnormal samples are removed, and duplicate or invalid behaviors are filtered out. The processed data facilitates more targeted mining of online learning behavior patterns and the discovery of learners' online learning patterns.

(3) Process Mining: There are numerous algorithms for process mining. This paper employs Microsoft's Sequential Analysis and Clustering Analysis algorithm, a hybrid algorithm combining clustering and Markov chains. A key feature of this algorithm is its use of sequential data, enabling it to identify similar sequential patterns (e.g., clickstream analysis of learners browsing a website) through clustering analysis and determine the probability of a data point belonging to a specific category using probabilistic methods.

(4) Knowledge Representation: The identified self-directed learning behavior patterns are presented to users in a visual and intuitive manner, facilitating further in-depth analysis.

3.4. Methods for Identifying Independent Learning Behaviors among College Students

Based on the characteristics of college students' online self-directed learning behavior extracted above, a method for identifying college students' self-directed learning behavior using a multi-layer perceptron (MLP) model is employed to assess the self-directed learning performance of individual college students, thereby determining their self-directed learning ability (as reflected in test scores in this study).

The network structure of the multi-layer perceptron model used is shown in Figure 2. This multi-layer perceptron model consists of four layers. The first layer is the input layer, where the input is the feature vector extracted from the online learning behavior data, i.e., the weights of the extracted characteristics of each college student's self-directed learning behavior. The second and third layers are hidden layers, used to extract higher-level feature information from the basic features of the input layer. The fourth layer is the output layer, which outputs the probability of whether each college student's learning behavior data constitutes autonomous learning behavior. In this model, the output of each layer's neural network nodes is a function of the previous layer's neural network nodes.

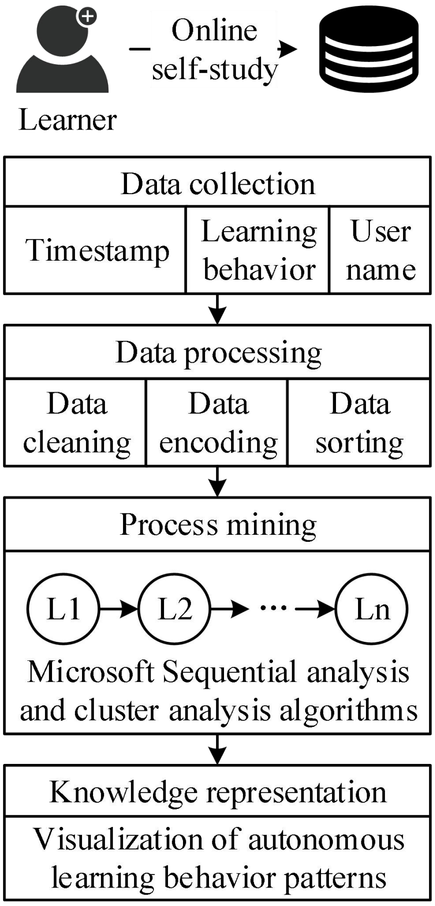


Figure 1. Online self-learning behavior pattern mining framework.

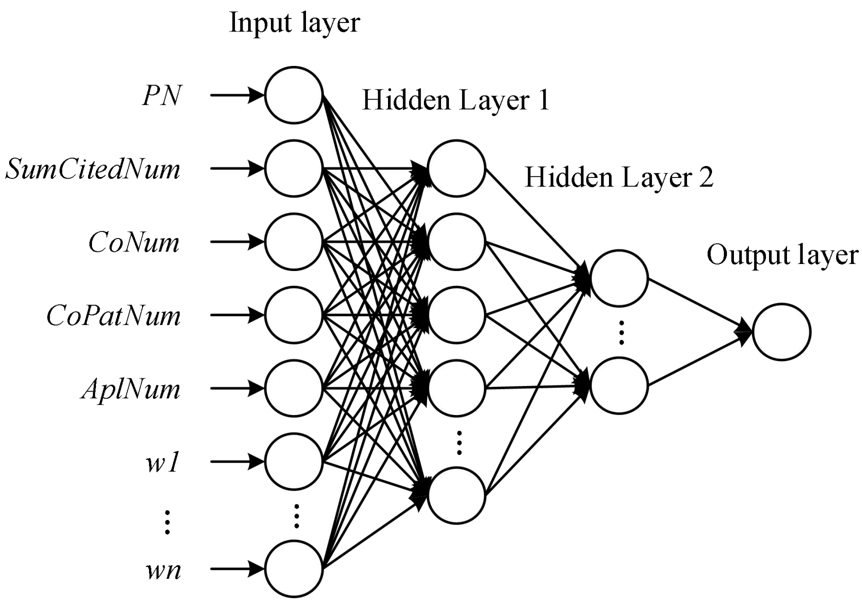


Figure 2. Multi-layer perceptual mode network structure.

For the i th neuron node in the k th layer, its output depends on the activation function $f^{(k)}(\theta_i^{(k)})$ of that layer, where $\theta_i^{(k)}$ is the weighted sum of the inputs to that node, defined as:

$$\theta_i^{(k)} = \sum_{j=1}^{N_{k-1}} w_{ij}^{(k)} Z_j^{(k-1)} \tag{10}$$

In the equation, $Z_j^{(k-1)}$ is the output value of the j th node in the $k-1$ th layer. $w_{ij}^{(k)}$ is the connection weight from the j th node in the $k-1$th layer to the i th node in the k th layer. N_{k-1} is the number of nodes in the $k-1$th layer.

For the first to third layers of the multi-layer perceptron model, the rectified linear unit (ReLU) is selected as the activation function for that layer, i.e., $f(z) = \max\{0, z\}$;

For the fourth layer of the multi-layer perceptron model, since the output is the probability of an inventor being a technological innovation talent, i.e., the output needs to be mapped to the range (0, 1), the sigmoid function is used as the activation function for this layer, i.e.:

$$f(z) = \frac{1}{1+\exp(-z)} \tag{11}$$

Model parameter learning: In order to discover autonomous learning behavior in online learning behavior data based on the constructed multi-layer perceptron model, it is first necessary to learn the multi-layer perceptron model. The learning of the multi-layer perceptron model mainly involves learning the connection weights between nodes in each layer. The learning process is typically implemented using the backpropagation (BP) algorithm based on the principle of stochastic gradient descent. Specifically, a small random number is first set as the initial weight of the multi-layer perceptron. Then, the training dataset is input into the MLP network, and the backpropagation (BP) algorithm based on the principle of stochastic gradient descent is used to train the network, adjusting the network parameters so that the actual output values obtained after processing by the MLP model are as close as possible to the expected output values.

In this paper, this means making the probability values of the identified college students' self-directed learning behaviors as close as possible to the labels indicating whether the college students engage in self-directed learning behaviors. The mean squared error is used as the loss function during the training process, and this loss function $J(w)$ can be expressed as:

$$J(w) = \frac{1}{m} \sum_{i=1}^{m} (h_w(x^{(i)}) - y_i)^2 \tag{12}$$

In the equation, $h_w(x^{(i)})$ is the probability of the i th college student engaging in self-directed learning behavior, calculated using the multi-layer perceptron model described in the paper. y_i is the pre-labeled tag for the i th self-directed learning behavior, indicating whether the college student is a self-directed learning talent (if $y_i = 1$, the college student is a self-directed learning talent; if $y_i = 0$, the college student is not a self-directed learning talent). m is the total number of self-directed learning behaviors among college students in the patent collection.

4. Analysis of College Students' Learning Behavior Data and Identification of Autonomous Learning Behavior

4.1. Analysis of Learning Behavior Based on Online Learning Behavior Data

This paper constructs a process for identifying and analyzing college students' autonomous learning behaviors from four aspects: data collection, data preprocessing, data analysis, and online learning effectiveness. This process is used to guide stakeholders in monitoring learners' learning patterns and states and making precise decisions, thereby reducing the quality crisis in online learning.

Taking a four-month "University English Translation" course offered by a certain university on the ChaoXing platform from September 1, 2024, to December 30, 2024, as an example, learning behavior data generated by 2,416 learners on the platform was collected and categorized into four types and ten items. The description of online learning behavior variables is shown in Table 1.

① Behavioral engagement: number of tasks completed, number of course videos watched, video viewing duration, and number of chapter studies.

② Cognitive engagement: chapter quiz scores and regular exam scores.

③ Social engagement: total number of discussions, number of discussions posted, and number of discussions replied to.

④ Online learning outcomes: overall score.

The collected online learning behavior data was preprocessed, primarily involving the following three steps:

① Data cleaning: removing unnecessary fields, eliminating missing values, and outliers.

② Data integration: using Excel to integrate tables from different data sources into a single table.

③ Data standardization: using the z-score standardization method to convert online learning behavior data of different scales into uniformly measured z-score scores.

Table 1. The online learning behavior variable description.

Dimensional division	Learning behavior variable	Learning behavior variables
Behavior input	The number of section tasks	Complete the total number of tasks
	Video viewing number	Watching the total number of video
	Video viewing duration	The total length of video
Cognitive input	The number of chapters learned	The total number of chapters (the number of visits)
	Chapter test results	The score of the chapter test is calculated according to the test
	Normal test scores	Score for all exams
Social investment	Total discussion number	The number of discussions and the total sum of the discussion
	The number of discussions published	The total number of discussions is published
	Reply discussion number	Reply to the total number of discussions
Online learning	Total score	$\sum Learning\ behavior\ variable * Weighting\ ratio$

The study analyzed online learning behavior data using various statistical methods as follows:

① Analysis of factors influencing online learning outcomes: Correlation analysis and regression analysis were used to validate the relationships between behavioral engagement, cognitive engagement, social engagement, and online learning outcomes.

② Prediction of online learning outcomes: Spearman's correlation analysis was used to preliminarily screen learning behavior variables influencing online learning outcomes, followed by multiple linear regression analysis to identify key variables. The key variables were used as samples for multiple predictive models, and the optimal predictive model was selected based on accuracy and comprehensive evaluation indicators to accurately predict online learning outcomes and identify potential risk individuals.

③ Learner clustering: Hierarchical clustering was used to further divide identified potential risk individuals and non-risk individuals into distinct groups with similar behaviors. Based on the data analysis results, effective interventions were implemented for learner groups with similar behaviors.

4.1.1. Analysis of Factors Affecting Online Learning Effectiveness

The relevant analysis is shown in Table 2. A correlation analysis was conducted between behavioral engagement, cognitive engagement, social engagement, and online learning effectiveness to preliminarily verify the relationships between the various influencing factors. The results in the table indicate that there is a significant positive correlation between behavioral engagement, cognitive engagement, social engagement, and online learning effectiveness ($p < 0.05$, $r > 0$).

Table 2. Correlation analysis.

Dimensional division	Behavior input	Cognitive input	Social investment	Learning effect
Behavior input	1	0.458*	0.091*	0.724*
Cognitive input	0.458*	1	0.103*	0.682*
Social investment	0.091*	0.103*	1	0.275*
Learning effect	0.724*	0.682*	0.275*	1

Further validation of how behavioral engagement, cognitive engagement, and social engagement influence online learning outcomes was conducted using multiple linear regression analysis. The results of the multiple linear regression analysis are shown in Table 3.

The results indicate that behavioral engagement, cognitive engagement, and social engagement all positively influence online learning outcomes ($p < 0.05$, $r > 0$), meaning that the higher the learner's behavioral engagement, cognitive engagement, and social engagement, the better the learner's online learning outcomes. Among these, social commitment has the most significant impact on online learning outcomes, followed by cognitive commitment, with behavioral commitment having the least impact.

In summary, behavioral commitment, cognitive commitment, and social commitment are key factors influencing learners' ability to achieve good online learning outcomes, and there is an interdependent relationship among these forms of commitment.

Table 3. Multivariate linear regression analysis.

Model	Unnormalized coefficient		Normalization factor	t	p	VIF
	B	Standard error	B			
Constants	35.261	0.551		61.859	0.006*	
Behavior input	0.075	0.000	0.664	83.046	0.000*	1.228
Cognitive input	0.372	0.007	0.608	84.692	0.002*	1.453
Social investment	0.689	0.153	0.052	6.559	0.001*	1.094

4.1.2. Various Learning Behavior Variables and Learning Outcomes

The correlation analysis between various learning behavior variables and learning outcomes is shown in Table 4. A Spearman correlation analysis was conducted between the nine learning behavior variables and online learning outcomes (total scores) to preliminarily identify the variables influencing online learning outcomes.

The results in the table indicate that nine variables are significantly positively correlated with online learning outcomes ($p < 0.05$, $r > 0$): number of homework exercises, number of course videos watched, video viewing duration, number of chapter studies, number of unit quizzes, regular exam scores, total number of discussions, number of discussions posted, and number of discussions replied to. Among these, the number of homework exercises, number of course videos watched, number of discussions posted, and number of unit quizzes rank higher in terms of correlation.

Table 4. The correlation analysis of learning behavior variables and learning effects.

Serial number	Variable	Correlation (r)
1	Homework practice	0.727*
2	Video viewing number	0.704*
3	Video viewing duration	0.513*
4	The number of chapters learned	0.585*
5	Unit test number	0.601*
6	Normal test scores	0.524*
7	Total discussion number	0.157*
8	The number of discussions published	0.612*
9	Reply discussion number	0.194*

4.2. Analysis of Learning Patterns and Exploration of Independent Learning Behavior

This chapter analyzes the patterns of effort and evolution of learners in a dataset of 11 real online courses. The four learning modes proposed here—stagnant, focused, rushed, and rhythmic—are analyzed and discussed. The primary focus is on the patterns of effort and outcomes of learners across the four modes over time in different types of courses, with the aim of identifying college students' autonomous learning behaviors.

Specifically, for each learning mode, this paper calculates weekly averages for each course, including average activity duration, activity count, and average quiz scores, to analyze the dynamic changes in course activity effort (activity duration and count) and outcomes (quiz scores) across different learning modes. In-depth analysis was conducted on the datasets of all 11 courses, and the observed patterns were found to be largely consistent. This paper selects “Artificial Intelligence” (representing science-related courses) and “University English Translation” (representing non-science-related courses) for demonstration and analysis.

4.2.1. Average Duration of Activities

Exploring the dynamic changes in the weekly learning activities (watching videos, completing assignments) of learners across four learning modes.

To better present the analysis results, the actual learning duration of the learner group is normalized by dividing it by the maximum learning duration, and the normalized values are used to represent the average activity duration of each learning group.

Since the evolutionary patterns of learning activity duration for the “Artificial Intelligence” course are consistent with those of the “University English Translation” course, this section only presents the evolutionary pattern analysis for the “University English Translation” course.

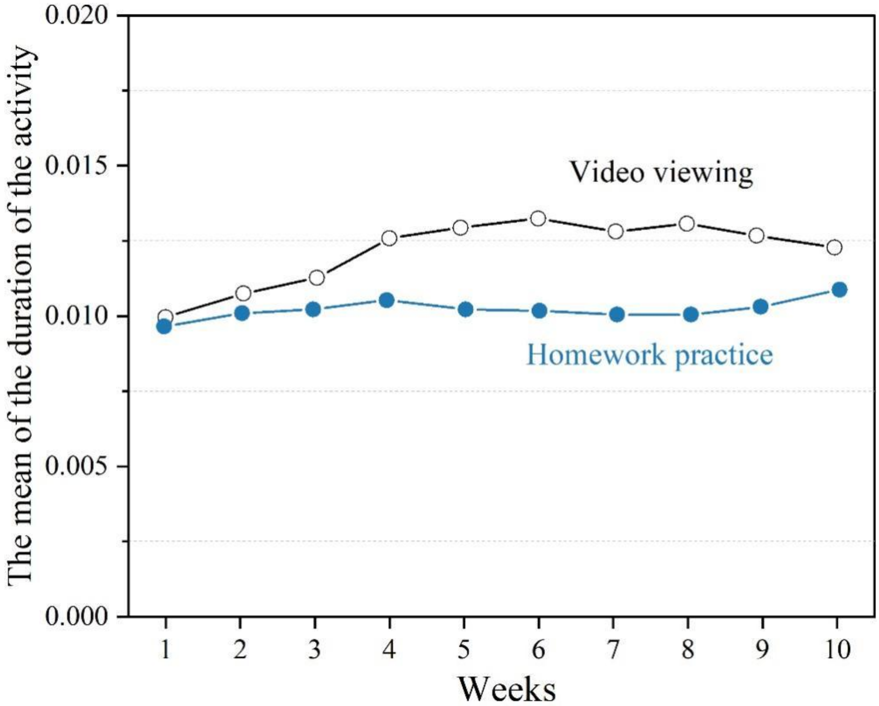
The video viewing and homework practice durations for the “University English Translation” course are shown in Figure 3.

Figure (a) represents stagnant learners, whose average video viewing and homework practice durations remain between 0.010 and 0.015.

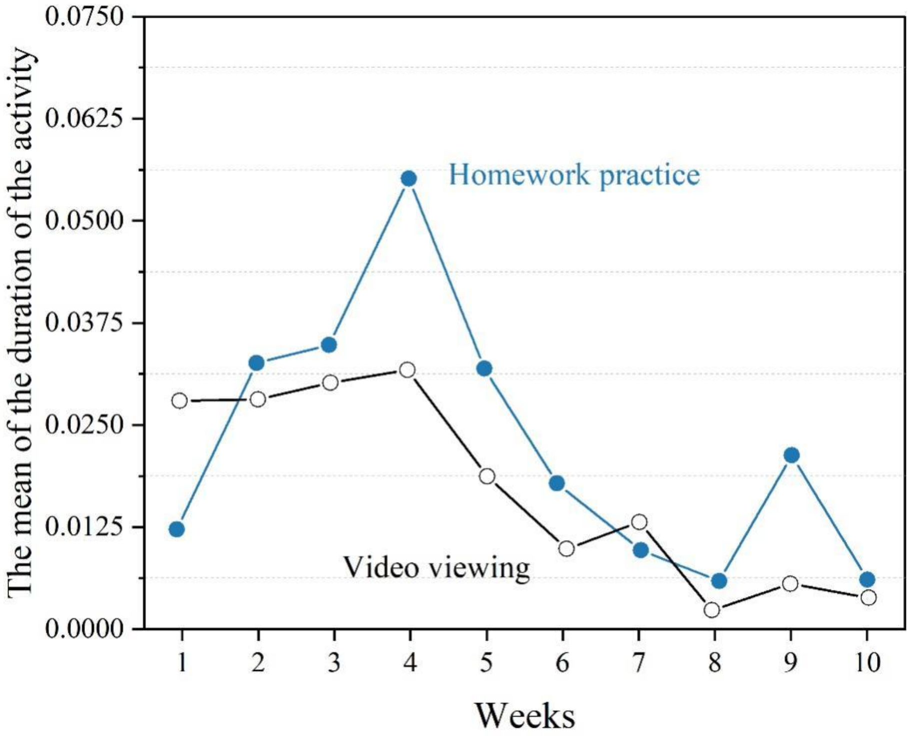
In Figure (b), the homework practice duration reaches its maximum value of 0.05625 in the 4th week, indicating focused learners.

Figure (c) represents the rush-type learner, with the average duration of homework practice reaching 0.100 in the 7th week.

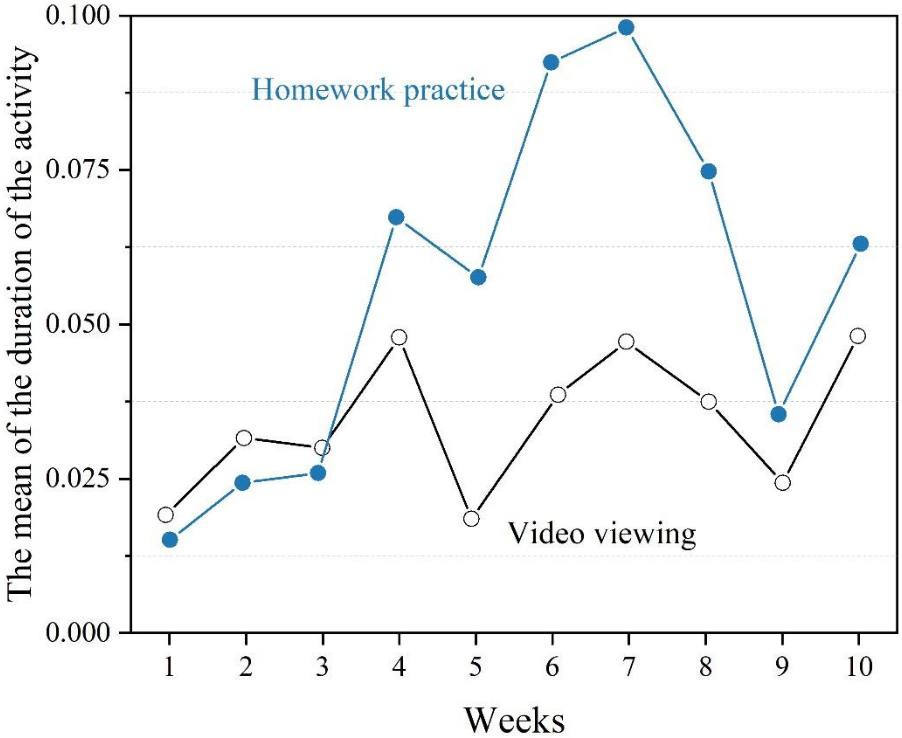
Figure (d) represents the rhythmic-type learner, with the average duration of homework practice reaching its peak at 0.125 in the 6th and 8th weeks.



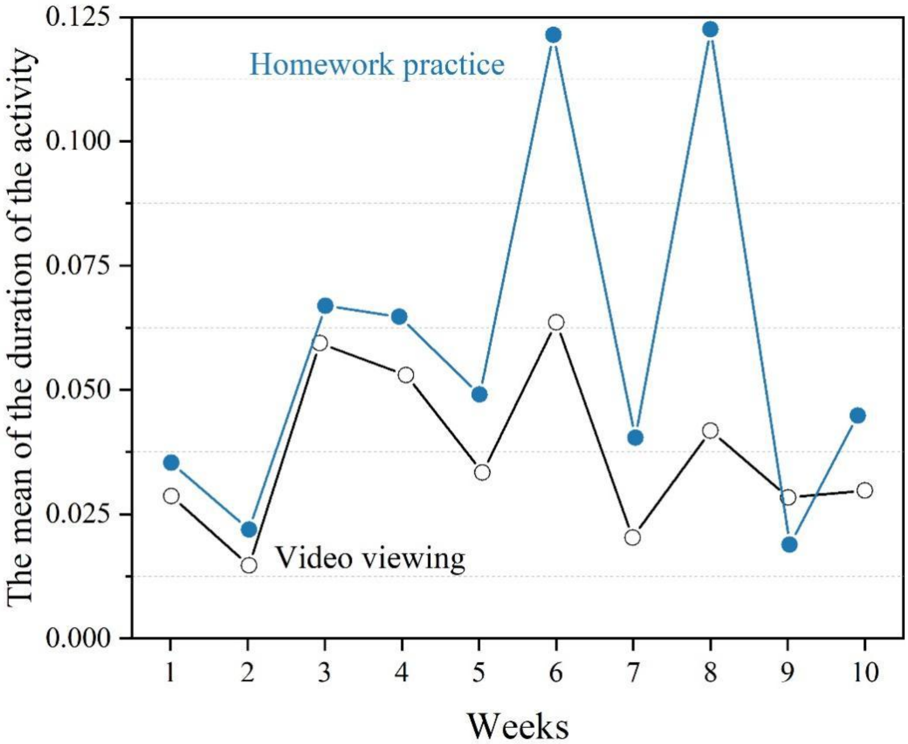
(a) Stagnant learner



(b) Focus on learners



(c) Worker learner



(d) Rhythmic learner

Figure 3. "College English translation" video viewing and homework practice hours.

4.2.2. Number of Learning Activities

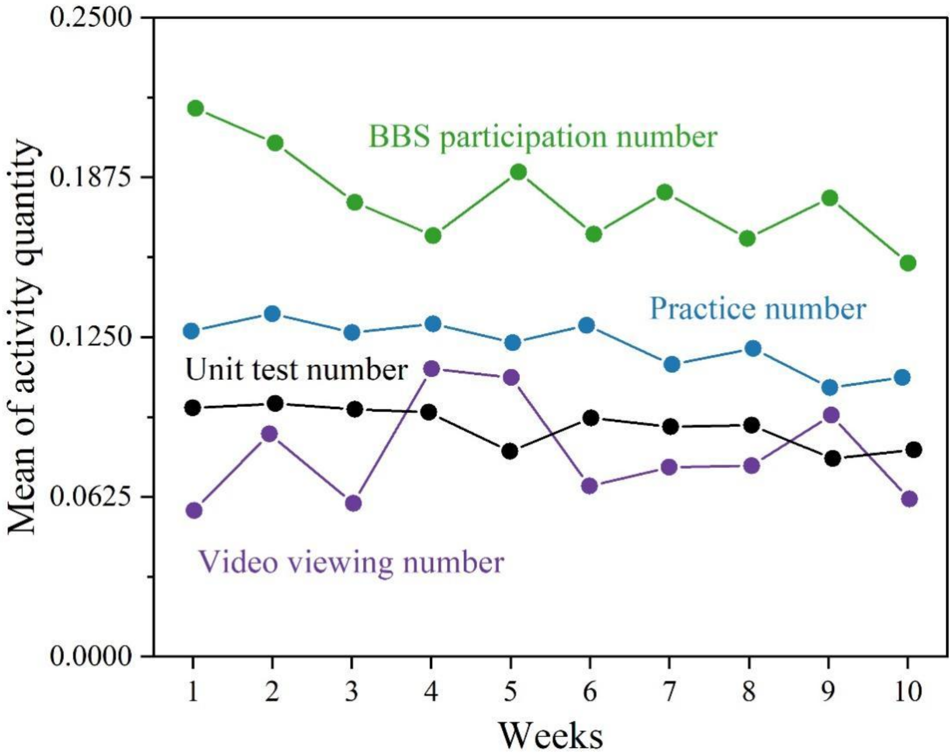
The dynamic evolution of the number of “University English Translation” learning activities is shown in Figure 4. Learning groups with different learning modes exhibit distinct characteristics and trends in their learning efforts and behaviors.

Throughout the entire course learning phase, “stagnant-type” learners made minimal video viewing efforts but almost no participation in homework exercises.

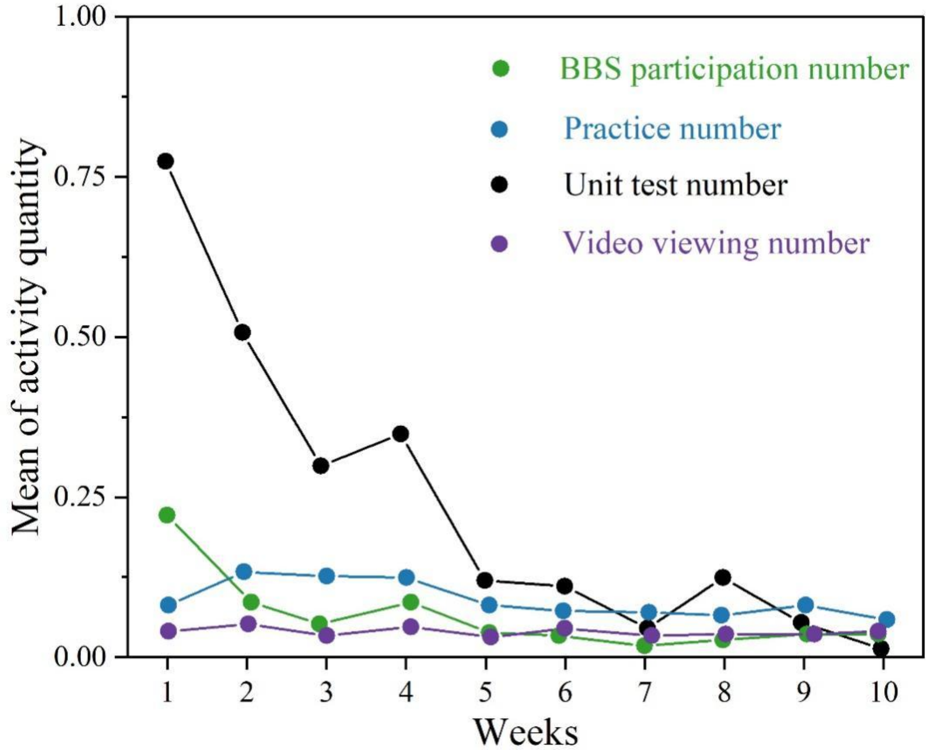
“Focused-type” learners demonstrated higher initial learning efforts, with gradually increasing homework exercise and video viewing efforts, which then tended to decline.

“Rush-type” learners consistently invested significant effort in homework exercises and video viewing throughout the course, with the highest homework exercise effort occurring during the mid-course phase.

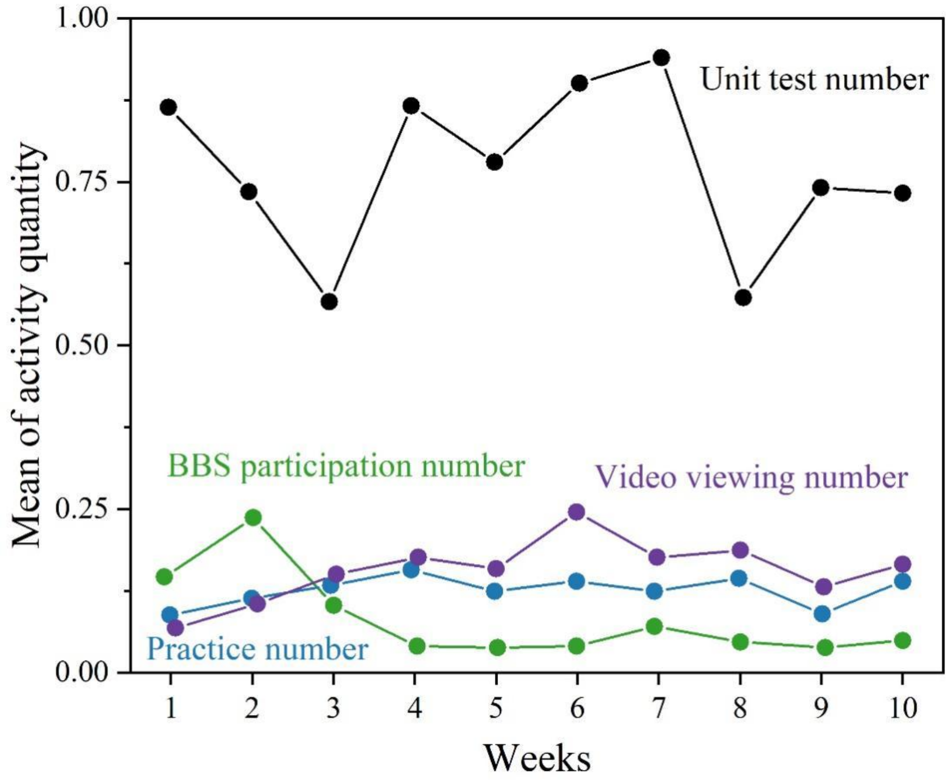
“Paced” learners have a longer learning investment duration, and their changes in video viewing and homework exercise investment are consistent throughout the entire course learning phase.



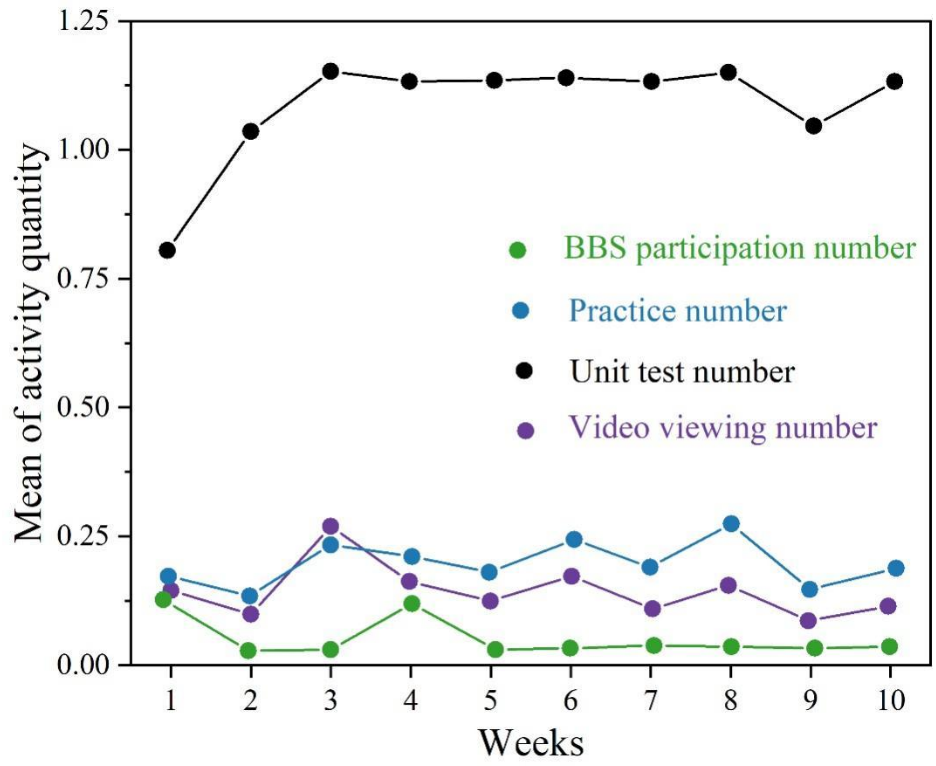
(a) Stagnant learner



(b) Focus on learners



(c) Worker learner



(d) Rhythmic learner

Figure 4. The dynamic evolution of "college English translation" learning activity.

4.2.3. Changes in Test Scores

This paper uses unit test scores to indicate learners' learning outcomes.

The evolution of learners' unit test scores is shown in Figure 5, which illustrates the dynamic changes in unit test scores for learners of different learning modes in the “University English Translation” and “Artificial Intelligence” courses.

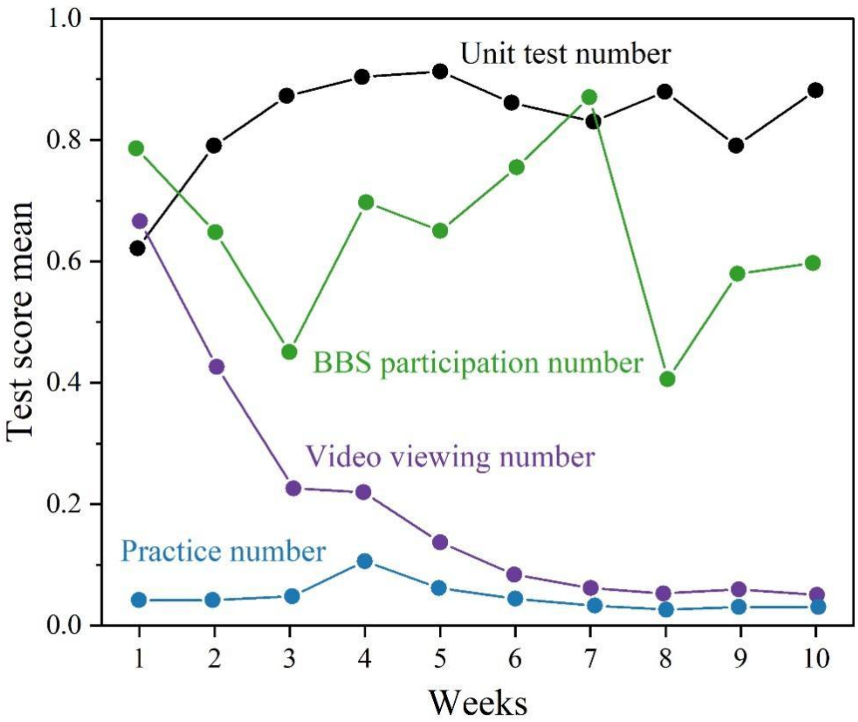
It can be observed that unit test scores (outcomes) are closely related to learners' course activity participation (i.e., effort).

Specifically: “Stagnant-type” learners generally achieve lower unit test scores, as they only took a few unit tests toward the end of the course, with self-directed learning behaviors emerging only in the later stages of the course.

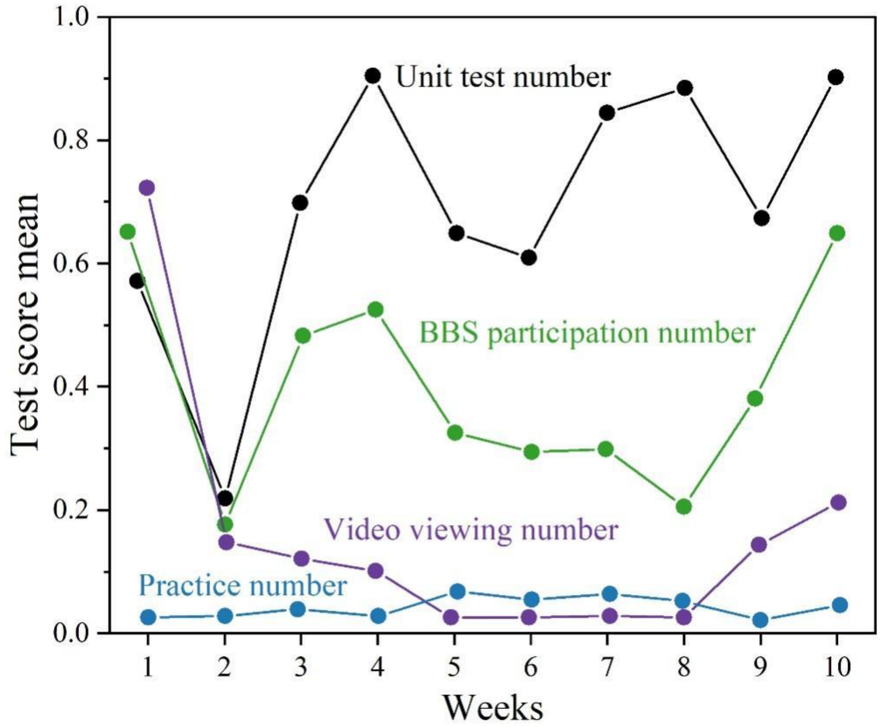
“Focused” learners achieved good unit test scores in the first four weeks due to higher learning investment during this period. However, as learning investment decreased in the middle and later stages of the course, their unit test scores also gradually declined. Self-directed learning behavior showed a decreasing trend and did not appear consistently.

“Rhythmic” learners maintained stable learning effort and good unit test scores (outcomes) throughout the course, with consistent self-directed learning behavior.

“Rush-type” learners made less learning effort in the early stages of the course, resulting in poor unit test scores. However, as their learning effort increased in the later stages, their unit test scores also improved. Self-directed learning behaviors were less frequent in the early stages of the course and more frequent in the later stages. Self-directed learning behaviors showed a trend of increasing from less to more.



(a) College English translation



(b) Artificial intelligence

Figure 5. The evolution of the learners' unit test scores.

5. Conclusion

This paper proposes a method for identifying college students' autonomous learning behavior based on the principles of the multi-layer perceptron (MLP) model. After processing online learning behavior data, the study analyzes the number of video views, homework practice sessions, unit tests, and participation in discussions as indicators of autonomous learning behavior, based on factors influencing online learning outcomes.

Among the factors influencing online learning outcomes, behavioral engagement shows a positive correlation with online learning outcomes, with a correlation coefficient of 0.724, higher than cognitive engagement (0.682) and social engagement (0.275). Among the various learning behavior variables, the number of video views, the number of homework exercises, the number of unit tests, and the number of participation in discussions are closely related to learning outcomes. Based on the above four learning behavior variables, four learning modes are proposed, and the characteristics of each learning mode are analyzed to explore the manifestation of self-directed learning behavior. The analysis of self-directed learning behavior shows that "rhythmic-type" learners exhibit stable self-directed learning behavior and maintain good test scores. In contrast, "stagnant-type" learners have fewer video viewings and homework practice sessions, exhibit less self-directed learning behavior, and have relatively lower test scores.

Funding

The 2024 Guizhou Provincial Higher Education Institution Undergraduate Teaching Content and Curriculum System Reform Project: "Research on Teaching Reform of E-Commerce Security and Payment Courses in the Context of Interdisciplinary Integration" (Project No. JG2024245). Guizhou Provincial First-Class Course Construction Project "Electronic Payment and Business Security" (Project No. SXXH2022232).

References

1. Chen, X., Breslow, L., & DeBoer, J. (2018). Analyzing productive learning behaviors for students using immediate corrective feedback in a blended learning environment. Computers & Education, 117, 59-74.
2. Zhang, J. (2023). Cognitive Status Analysis for Recognizing and Managing Students' Learning Behaviors. International Journal of Emerging Technologies in Learning, 18(16).

3. Li, Y., Qi, X., Saudagar, A. K. J., Badshah, A. M., Muhammad, K., & Liu, S. (2023). Student behavior recognition for interaction detection in the classroom environment. Image and Vision Computing, 136, 104726.
4. Wang, S. (2021). Online Learning Behavior Analysis Based on Image Emotion Recognition. Traitement du Signal, 38(3).
5. Bánhegyi, M., & Fajt, B. (2022). University Students' Autonomous Learning Behaviors in Three Different Modes of ICT-based Instruction in the COVID-19 Era: A Case Study of Lockdown Learning. Studies in Self-Access Learning Journal, 13(1).
6. Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. Computers & education, 104, 18-33.
7. Hao, X., & Gu, X. (2024). Analysis of key roles in large-scale online learning: Interactive participation characteristics and knowledge construction behaviour patterns. British Journal of Educational Technology, 55(3), 910-932.
8. Zhang, J. H., Zhang, Y. X., Zou, Q., & Huang, S. (2018). What learning analytics tells us: Group behavior analysis and individual learning diagnosis based on long-term and large-scale data. Journal of Educational Technology & Society, 21(2), 245-258.
9. Hu, K., Jin, J., Zheng, F., Weng, L., & Ding, Y. (2023). Overview of behavior recognition based on deep learning. Artificial intelligence review, 56(3), 1833-1865.
10. Fu, R., Wu, T., Luo, Z., Duan, F., Qiao, X., & Guo, P. (2019, December). Learning behavior analysis in classroom based on deep learning. In 2019 Tenth International Conference on Intelligent Control and Information Processing (ICICIP) (pp. 206-212). IEEE.
11. Wang, X., & Zhang, W. (2022). Improvement of students' autonomous learning behavior by optimizing foreign language blended learning mode. Sage Open, 12(1), 21582440211071108.
12. Bai, X., Wang, X., Wang, J., Tian, J., & Ding, Q. (2020, August). College students' autonomous learning behavior in blended learning: Learning motivation, self-efficacy, and learning anxiety. In 2020 International Symposium on Educational Technology (ISET) (pp. 155-158). IEEE.
13. Huan, S., & Yang, C. (2022). Learners' Autonomous Learning Behavior in Distance Reading Based on Big Data. International Journal of Emerging Technologies in Learning (iJET), 17(9), 273-287.
14. Tong, Y., & Zhan, Z. (2023). An evaluation model based on procedural behaviors for predicting MOOC learning performance: Students' online learning behavior analytics and algorithms construction. Interactive Technology and Smart Education, 20(3), 291-312.
15. Zhong, S. H., Li, Y., Liu, Y., & Wang, Z. (2017). A computational investigation of learning behaviors in MOOCs. Computer Applications in Engineering Education, 25(5), 693-705.
16. Li, H., Kim, M. K., & Xiong, Y. (2020). Individual learning vs. interactive learning: A cognitive diagnostic analysis of MOOC students' learning behaviors. American Journal of Distance Education, 34(2), 121-136.
17. Doğan, G., Sunar, A. S., Duru, İ., & White, S. (2018). Who is the English as a second language speaker in this MOOC. International Journal of Information and Education Technology, 8(3), 213-217.
18. Shen, J., Yang, H., Li, J., & Cheng, Z. (2022). Assessing learning engagement based on facial expression recognition in MOOC's scenario. Multimedia Systems, 28(2), 469-478.
19. Li, C., & Zhou, H. (2018). Enhancing the efficiency of massive online learning by integrating intelligent analysis into MOOCs with an application to education of sustainability. Sustainability, 10(2), 468.
20. Mingyue Qi. (2025). Study on the Effect of Autonomous Learning of College Students in the Online Education Model. Journal of Innovation and Development,10(2),9-11.
21. Ng & Tsun Sing. (2025). Cultivation of Students' Autonomous Learning Ability in the Context of Artificial Intelligence. Journal of Artificial Intelligence Practice,8(2).
22. Yan Yu Xie,Xiao Dong Li,Xiu Ying Liu,Jing Xin Yu & Jun Fei Wang. (2025). Integrating molecular simulations with multilayer perceptron neural networks to predict the CH4/H2 adsorption and separation of MTV-MOFs. International Journal of Hydrogen Energy,132,155-165.