International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 5134-5146

https://doi.org/10.70517/ijhsa464355

Research on AI-driven innovation in educational management models and personalized student development

Weina Li[1,*]
[1] Faculty of Education and Liberal Arts, INTI International University, Nilai, Negeri Sembilan, 71800, Malaysia
Corresponding authors: (e-mail: LXY257257@126.com).

Abstract The application of AI technology promotes the innovation of education management mode and provides more possibilities for students' personalized development. Based on AI technology, this study proposes a smart education management model and analyzes the consumption and behavioral data of college students through data mining methods to explore possible paths for their personalized development. The research method mainly includes the construction of student behavioral portrait and the application of K-prototype clustering algorithm. In the data collection phase, 61.84 million consumption records, 31,754 students' performance data and 1,304,736 book borrowing records were analyzed. It was found that based on the K-prototype clustering algorithm, students were categorized into four consumption groups, in which the average monthly consumption of students in group 3 was $291.61 and the average monthly food and beverage consumption was $128.46, while group 4 had the highest average monthly consumption of $961.56. In addition, the analysis revealed the significant correlation between factors such as age and gender and students' consumption behavior. The conclusion shows that the smart education management model and data mining methods provide important support for education management and students' personalized development, especially in the areas of consumption behavior, academic performance and mental health, providing basic data support for personalized education management.

Index Terms Artificial Intelligence, Smart Education, Data Mining, Student Behavioral Portrait, K-prototype Clustering, Personalized Development

I. Introduction

Educational management refers to the educator through the organization and coordination of the educational team, give full play to the role of educational human, financial and material information, the use of a variety of favorable conditions within the education, the process of activities to achieve the goals of educational management with high efficiency [1]. For colleges and universities, education management is a kind of macro-adjustment of the internal resources and the mode of education and teaching development, so that all aspects of the college and university dynamics can be in the right position, and actively play its role in the advantages of the management approach [2]-[4]. However, the traditional education management mode in the face of the impact of wisdom education, reveals many limitations and deficiencies [5]. It is urgent for us to build a new ecology of more open, collaborative and intelligent education management through multifaceted efforts.

With the development and application of communication technology, a huge social change is being set off. Under such a background, the change of intelligent education in the field of education is gradually emerging, which not only profoundly changes the traditional teaching and learning methods, but also puts forward unprecedented challenges and opportunities to the education management mode [6]-[8]. Based on the deep integration of technology and education, realizing the innovation and change of education is the core goal pursued by smart education [9]. Wisdom education provides impetus for the innovation and change of education, and wisdom education drives the innovation and change of education concept, teaching mode, learning concept and learning mode, education system and talent cultivation mode, as well as teaching management and teaching evaluation [10]-[13]. In addition, education administrators can appropriately leverage smart technologies to collect relevant information, improve databases, and use platforms to provide personalized services for students, with a view to promoting education management to keep pace with the times [14]-[16]. However, the in-depth implementation of smart education is not an overnight solution, which requires us to carry out profound changes and innovations at the level of education management to meet the needs of education development in the new era.

In this study, we first start from the basic data of students' consumption behavior and academic performance to construct students' behavioral portrait, and classify students by K-prototype clustering algorithm to identify their different behavioral patterns. Secondly, combining with data mining technology, the intrinsic relationship between students' behaviors and their academic performance, mental health, living habits and other factors is explored through methods such as correlation analysis and multivariate analysis of variance (ANOVA). Through such analysis, targeted education management programs can be provided to universities, especially in the implementation of personalized education, to achieve accurate teaching resource allocation and academic counseling.

II. Innovative design of AI-based smart education management model

This chapter designs a smart education management model based on artificial intelligence and proposes an innovative path for smart education based on visualization technology.

II. A. Construction of a smart education management model

Constructing a smart education management model is the direction of the development of education informatization in colleges and universities, which can promote the development of education and teaching in schools and meet the needs of selecting innovative talents with information technology in the new era. Based on the smart campus data center, fully utilizing the digital campus environment, and using the Internet of Things, cloud computing, big data, ubiquitous network and other technologies, this study builds a smart education management model as shown in Figure 1. This model realizes six kinds of intelligent education business, such as intelligent teaching, intelligent environment, intelligent evaluation, intelligent management, intelligent scientific research, intelligent service, etc. It provides new interactive application scenarios for teachers, students, parents, administrators and other users, and integrates teaching, evaluation, scientific research, management of colleges and universities with a new type of big data application platform, which is used to realize the visibility of information perception, the clarity of application interaction, the flexibility of education means and the timeliness of service response. Flexibility of information perception, clarity of application interaction, flexibility of education means and timeliness of service response. At the same time, the usability of the intelligent education management mode is guaranteed by the visualization security platform, which embodies the integration characteristics and makes the intelligent and advanced classroom teaching environment, ubiquitous network learning platform, transparent and efficient education management system, integrated and innovative network scientific research environment, accurate and reliable teaching evaluation system, and online and offline attentive teaching service platform a reality.

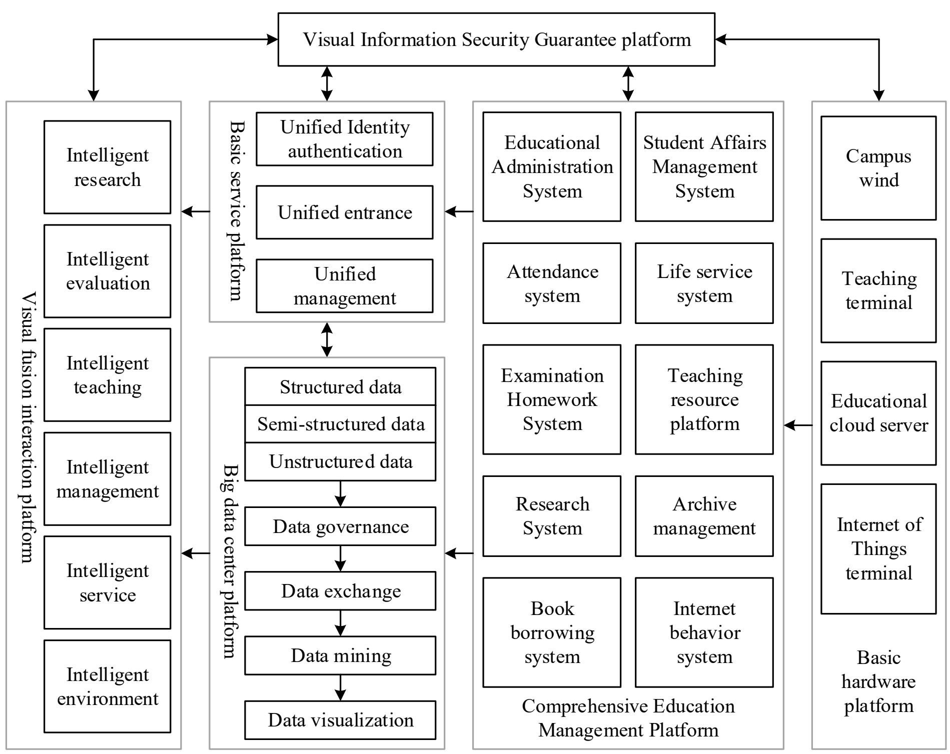


Figure 1: Smart education management mode

II. B. Innovative Path of Smart Education Based on Visualization Technology

Intelligent education is inseparable from visualization technology, and visualization technology greatly promotes the development of intelligent education. Intelligent education has three basic features: basic educational data, visualization of performance content and precision of intelligent algorithms. Visualization technology as an intermediate layer can effectively bridge the gap between data, systems and people, and its role in smart education management is shown in Figure 2.

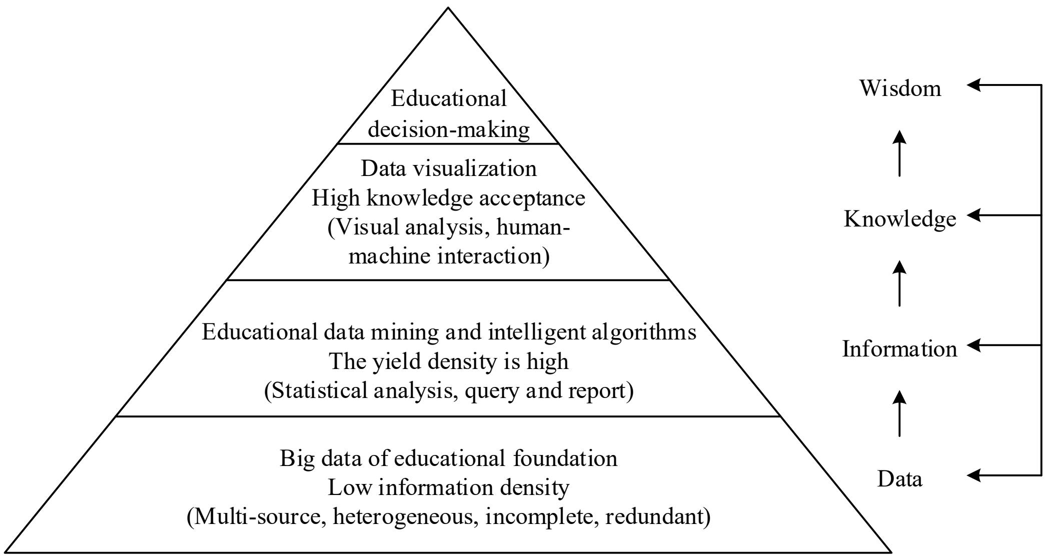


Figure 2: The Role of Visualization Technology in Smart Education Management

As an important means of human-computer interaction and analysis and decision-making, visualization technology presents data in the form of graphs and images suitable for human eyes to observe, establishes image-aware communication between the human brain and boring data, which can provide insight into the situational information implied in the data and assist in educational decision-making. Based on the visualization technology, this study considers the following six main directions of visualization innovation in smart education:

(1) Smart environment visualization innovation

Intelligent environment is the cornerstone of modern educational scenarios, and the visualization factors include spatial structure, environmental layout, combination of desks and chairs, interactive big screen, lighting and sound, computer equipment, mobile terminals, etc. By reconstructing and upgrading the network infrastructure, creating a brand-new learning and living space, and applying the technologies of multi-screen interaction, virtual simulation, wireless projection, etc., we can build a personalized, intelligent, and visualized new type of campus environment, training environment, and classroom environment, which can build a new campus environment, training environment and classroom environment with advanced technology, and can build a new campus environment with advanced technology. and classroom environment, can build a modern “smart campus” environment with advanced concepts, scientific and practical facilities.

(2) Intelligent teaching visualization innovation

The smart teaching platform improves the traditional education and teaching methods, promotes the learning methods of discussion, inquiry, collaboration and autonomy, and strengthens the practical teaching methods of work rules, process and result orientation. Through networked teaching means, it builds a ubiquitous learning platform to support dynamic learning. Networking, virtualization, visualization, visualization of teaching content, so that students sitting in the classroom can perceive the real physical scene, free from the confinement of the traditional classroom, can listen to the views of first-class scholars at home and abroad in the classroom, to further open up the student's horizons, and effectively improve the ease of absorption of knowledge.

(3) Wisdom evaluation visualization innovation

By constructing a data analysis platform for teaching behavior, and taking “classroom behavior can be perceived, teaching effect can be measured, assessment quality can be quantified, and the teaching process can be controlled” as the goal, the smart evaluation visualization uses the classroom camera system with the technology of “Internet of Things + Cloud Computing”, Intelligent Teaching Platform and Campus Life Service System to realize data collection; at the same time, it adopts structured face recognition and behavioral big data analysis technology to accurately analyze and identify the data in the classroom, such as teachers' teaching and students' learning. By constructing a database of teaching behavior, it realizes the transformation of teaching evaluation from academic evaluation to learning ability evaluation, from uni-dimensional evaluation to multi-dimensional evaluation, from ultimate evaluation to process evaluation, and from fuzzy evaluation to precise evaluation.

(4) Wisdom management visualization innovation

The application of visualization technology to wisdom management helps to promote the development of education management towards intelligence, refinement and standardization. By improving the school information management system and constructing visualized and customized service processes, it is possible to streamline, optimize and rebuild educational management protocols, and improve the effectiveness of planning, decision-making, execution and management work. Wisdom management visualization encourages managers to use data to speak, adopt visual management decision-making and coordination and control, and realize the qualitative change from traditional management to wisdom management.

(5) Smart Service Visualization Innovation

Service visualization makes the security and service work of the future school more appropriate, and truly people-oriented. Build an online and offline one-stop service platform with the goal of data service and business process integration, and realize the integration of on-campus services through the online service platform, the physical platform of the Sunshine Service Hall and the open self-service platform, so as to achieve "service visualization", "one-time completion" and even "do it well without running", so that all teachers, students and staff can share efficient, convenient and comfortable services.

(6) Smart Security Visualization Innovation

With the expanding scale of educational network, speeding up of information highway and increasing network applications, educational network security is facing more and more severe tests. Educational security visualization technology builds a good communication bridge between people and security systems, protects the increasingly important educational cyberspace, and provides a reliable and graphic data source for formulating campus network security policies. Security visualization makes the intelligent education business system more solid and secure, and can realize the network security goal of “visible, well managed, prevented and responded to”.

III. Data mining based behavioral analysis method for college students

In order to give full play to the effect of the intelligent education management mode and realize the personalized development of students, this chapter proposes a method of student behavior analysis in colleges and universities based on data mining technology, i.e., personalized education management is realized by constructing student behavioral portraits and performing cluster analysis based on student behaviors.

III. A. Methods for constructing a feature library of student behavioral portraits

III. A. 1) Data acquisition

This study is based on the student behavior records generated by college students at school, including student consumption data at school and school library book borrowing records. The consumption data includes students' dining consumption, supermarket consumption, campus transportation consumption, sports and fitness consumption and other consumption data.

In order to protect the privacy of students, the data studied in this paper have been desensitized, and the raw data contain more than 61.84 million consumption records of all teachers and students from 2021 to 2024, 31,754 students' performance data, and 1,304,736 students' book borrowing records. The original Campus One Card consumption records after desensitization contain 10 attributes: student number, consumption date, merchant account, consumption machine pos number, transaction time, transaction crediting time, transaction amount, department code, identity code, and gender. The original student book borrowing record contains the student number, book borrowing time, and book name. The student grade record contains the student number, semester, grade point average, and grade point.

III. A. 2) Data pre-processing

Pre-processing of raw data before big data mining analysis can ensure that the requirements for the analysis of student behavioral results are met and the quality of data mining results is improved. Data preprocessing techniques mainly include data cleaning, data integration, data normalization, data transformation and other techniques.

(1) Data cleaning

Data cleaning is the process of identifying and correcting incomplete, noisy and inconsistent data. When the data is acquired, the data format and data content are first checked to understand the basic situation of the data. Then use commonly used data cleaning methods to clean the dirty data, such as deleting or filling missing values, eliminating or correcting noisy data.

(2) Data integration

Data integration is the merging of data from multiple isomorphic or heterogeneous data sources, which helps to reduce inconsistencies and redundancies in the dataset in order to improve the accuracy and speed of the subsequent mining process.

(3) Data Transformation

Data transformation is to transform data into a form suitable for mining, making the data mining process more effective and the mined patterns easier to understand. The transformation methods mainly include smoothing, attribute construction, aggregation, normalization, discretization, and conceptual layering from nominal data.

(4) Data generalization

Data generalization is the original data set for statute representation, that is, the use of smaller data sets to represent the original data set, but still almost maintain the integrity of the original data, or statute after the data set can still achieve the experimental effect of the original data. The data statute method strategy has three methods: dimensional normalization, data normalization and data compression.

III. A. 3) Constructing a library of student behavioral portrait features

This paper combines the problems encountered in daily campus management and the needs of smart campus construction to perform statistical analysis, cluster analysis, correlation analysis and association rule analysis on students' consumption data, performance data and book borrowing data, as well as to construct students' social relationship network. Based on these research objectives, relevant indicators are designed to better analyze data mining and portray student profiles.

(1) Indicators for evaluating students' consumption ability

In order to quantify the consumption level of students, analyze the consumption structure of students and the differences in consumption among different students, the following indicators are defined in this paper:

1) Consumption Intensity: Considering the total amount of consumption, considering the number of times of consumption and the average level of consumption of the whole school, this paper establishes the consumption intensity indicators to measure the consumption ability of the students, including the breakfast consumption intensity $BreIntensity$, the lunch consumption intensity $LunIntensity$, the dinner consumption intensity $SupIntensity$, and the average daily Consumption Intensity $DayIntensity$ and Other Consumption Intensity $ElseIntensity$. A lower consumption intensity means a higher likelihood of poor economic conditions, and is defined in the following order:

$$BreIntensity = \frac{BreakfastExp}{AvgBreakfastExp} \tag{1}$$

$$LunIntensity = \frac{LunchExp}{AvgLunchExp} \tag{2}$$

$$SupIntensity = \frac{SupperExp}{AvgSupperExp} \tag{3}$$

$$DayIntensity = \frac{DayExp}{AvgDavExp} \tag{4}$$

$$ElseIntensity = \frac{ElseExp}{AvgElseExp} \tag{5}$$

$$\begin{aligned} Intensity &= BreIntensity + LunIntensity \\ &+ SupIntensity + ElseIntensity + DayIntensity \end{aligned} \tag{6}$$

where Exp denotes the amount of student consumption and Avg denotes the average value.

2) Cafeteria Dining Rate

In this paper, the cafeteria dining rate is designed to better portray students' dining behavior at school and reflect students' consumption ability. Cafeteria dining rate refers to the ratio of the number of times students dine in the school cafeteria to the total number of times they usually dine in school during a certain period of time, including breakfast cafeteria dining rate, lunch dining rate and dinner dining rate, with a lower cafeteria dining rate representing a higher consumption ability and a higher likelihood of it being higher, which is defined as follows:

$$BreakfastRate = \frac{BreakfastCount}{TotalCount} \tag{7}$$

$$LunchRate = \frac{LunchCount}{TotalCount} \tag{8}$$

$$SupperRate = \frac{SupperCount}{TotalCount} \tag{9}$$

where, $Count$ indicates the number of meals taken.

Other basic evaluation indicators of consumption ability, including the total amount of students' consumption in catering, the average monthly consumption amount, the average number of times of consumption per month, the average daily consumption amount, the average consumption amount of the three meals, and the total amount of other consumption.

(2) Evaluation Indicators of Students' Living Habits Regularity

Students' living habits can be reflected by campus consumption data. Students' consumption time at school reflects whether their eating and living habits are regular, specifically the number of times they consume breakfast, lunch and dinner can be taken as one of the indicators of their living habits. The number of breakfast consumption represents whether students have the habit of waking up early, and the number of dinner consumption reflects whether students have the tendency to lose weight. The number of times and time of taking the campus minibus in the campus transportation data also reflects the regularity of students' travel. Meanwhile, according to the school's class schedule, 8:00-12:00 p.m. from Monday to Friday, weekends and holidays can be regarded as the time for students' outings, and the data on sports consumption reflects students' fitness patterns.

(3) Evaluation indicators of students' social relationship network

Data can be analyzed according to the consumption data, based on whether the two people often consume at the same time in the same place to determine whether the two people are friends or not, and then portray the student social relationship network, and through the visualization technology to real-time display of the student social relationship changes. This paper describes the social ability of students through social intensity, i.e., the number of friends of a student, and describes the closeness of the relationship between two people through intimacy.

III. B. K-prototype based clustering models

K-means algorithm [17] is the most common clustering algorithm, which is suitable for dealing with numerical data. K-modes algorithm [18], as an extension of K-means, is suitable for datasets with discrete attributes. In the actual clustering processing, the student characteristic data often contains both categorical data such as gender, age, education and numerical data such as scale data. K-prototype algorithm [19] is a hybrid data clustering algorithm based on the simultaneous consideration of numerical features and dissimilarity coefficients of subtyped features, such algorithms can deal with both subtyped and numerical data, which overcomes the limitation of K-means and K-modes algorithms that only support a single type of data and provides better clustering features.

III. B. 1) Distance metrics

A mixed dataset with n samples $X = \{X_1, X_2, X_3, \cdots, X_n\}$, Each data has a m dimension attribute, i.e., $X_i = \{x_{i1}, x_{i2}, \cdots, x_{ir}, x_{i(r+1)}, \cdots, x_{im}]$, let the first r attributes be numeric and the last $m-r$ attributes be categorical. Let the number of clusters be K and $Q = \{Q_1, Q_2, Q_3, \cdots, Q_k\}$ be a prototype of cluster L, which is the center of cluster L. For the dissimilarity of numerical attributes, the Euclidean distance measure is used:

$$d_{Num}(X_i, Q_i) = \sum_{j=1}^{r} (x_{ij}^N - q_{ij}^N)^2 \tag{10}$$

For the dissimilarity of category-type attributes, the Hamming distance measure was used:

$$d_{Cla}(X_i, Q_i) = \sum_{j=r+1}^{m} \delta(x_{ij}^C, q_{ij}^C) \tag{11}$$

In equation (11), $\delta(x_{ij}^c, q_{ij}^c) = \begin{Bmatrix} 0, x_{ij}^c = q_{ij}^c \\ 1, x_{ij}^c \neq q_{ij}^c \end{Bmatrix}$.

The distance from X_i to the prototype Q_i can be expressed as:

$$d(X_i, Q_i) = d_{Num}(X_i, Q_i) + \lambda_i d_{Cla}(X_i, Q_i) \tag{12}$$

In Eq. (12), λ_i is to regulate the weight between the two attribute data, and when $\lambda_i = 0$ is equivalent to the K-means algorithm, the larger λ_i is, the more weight the category-type attribute accounts for. The variance after standardization of numerical type attributes is 1, so λ_i is set to 0.5.

The criterion for K-prototype clustering is to choose an appropriate loss function to measure the distance of numerical and categorical variables to the prototype, then this loss function can be defined as:

$$E(X,Q)=\sum_{i=1}^{n}\sum_{j=1}^{k}w_{ij}d(X_i,Q_i) \tag{13}$$

In equation (13), $w_{ij}\in[0,1]$, denotes whether the sample X is partitioned into the cluster L or not, with 1 if it is, and 0 if it is not.

III. B. 2) Steps of the K-prototype algorithm

Step1: Randomly select k initial prototypes (clustering centers) Q.

Step2: According to Eq. (12), calculate the distance $d(X_i,Q_i)$ between each sample point and K prototypes, classify it into the class closest to the center, and update the clustering center Q.

Step3: Recalculate the mixing distance $d(X_i,Q_i)$ from the sample point to the current Q and update Q.

Step4: Repeat Step2 and Step3 until no sample changes the category, i.e., the iterative loss function is minimized.

III. B. 3) Evaluation of clustering effects

Using scale data as numerical data and individual student characteristics as categorical data, the average profile coefficient value c was chosen to determine the number of clusters, and the profile coefficient method combines cohesion and separation of clusters for assessing the effectiveness of clustering. Namely:

$$c=\frac{1}{n}\sum_{i=1}^{n}\frac{b_i-a_i}{\max\{a_i,b_i\}} \tag{14}$$

In equation (14), a_i is the average distance between X_i and other samples in the same cluster, reflecting the cohesion of the cluster to which X_i belongs, b_i is the average distance between X_i and the nearest cluster, reflecting the separation of X_i from the rest of the clusters, and n is the sample number. $c\in[-1,1]$, where a larger value implies better performance.

IV. Experimental results and analysis

This chapter provides a practical application of the proposed data mining-based behavioral analysis method for colleges and universities, which lays the foundation for educational management based on students' personalized data and thus promotes students' personalized development.

IV. A. Analysis of Campus Consumption Behavior of Higher Education Students

IV. A. 1) Clustering results of students' campus consumption behavior

Using the K-prototypes clustering algorithm, the campus consumption behavior model of four categories of college students with significant behavioral characteristics is obtained with *k*=4. The number of data contained in the four clusters is shown in Table 1, and the distribution of the student population in each category is relatively uniform, indicating that the overall clustering effect is better. Among them, the number of students in cluster 3 category accounts for the highest proportion, reaching 38.18%, and the number of students in cluster 1 category accounts for the lowest proportion, only 18.35%.

Table 1: Distribution of clustering data on students' campus consumption behaviors

Category number	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Frequency (Percentage)	7628 (18.35%)	9246 (22.24%)	15874 (38.18%)	8829 (21.23%)

Based on the variance of the clusters to derive the variability of each cluster class, the results of the comparison of the means of the four numerical variables of students' consumption behaviors in the four clusters are shown in Table 2, where CB1 and CB2 denote the average monthly amount of daily life consumption and average monthly amount of food and beverage consumption, and CB3 and CB4 denote the frequency of average daily life consumption and average daily food and beverage consumption, respectively. * denotes p<0.05, ** denotes p<0.01, and later the same.

It can be seen that cluster 3 students have the lowest average monthly consumption amount and the lowest average daily consumption frequency, cluster 4 students have the highest average monthly catering consumption amount and average daily catering consumption frequency, cluster 1 students have the highest average monthly daily life consumption amount and average daily life consumption frequency, the difference between the average monthly daily life and catering consumption levels of cluster 2 and cluster 3 students is small, and the average monthly daily life consumption of cluster 1 student group is much higher than catering consumption. Overall, the mean value of Cluster 4, the group with the highest amount of campus consumption of students in this university, is 961.56, and its consumption level is significantly higher than that of Cluster 3, which has a mean value of 291.61.

Table 2: The comparison results of variance analysis differences of consumption behaviors

Consumer behavior index	Cluster 1	Cluster 2	Cluster 3	Cluster 4	F	p
	Mean±SD	Mean±SD	Mean±SD	Mean±SD		
CB1	447.52±148.75	214.84±147.63	128.46±101.57	422.95±213.16	13572.814	0.000**
CB2	184.97±129.64	257.41±165.82	163.15±134.89	538.61±142.85	15872.637	0.000**
CB3	2.78±0.94	3.34±0.75	1.94±0.93	4.49±1.02	17685.941	0.000**
CB4	2.53±0.97	0.84±0.72	0.71±0.64	1.18±0.95	12593.452	0.000**

IV. A. 2) Correlation Analysis of Students' Campus Consumption Behavior

In order to further explore the correlation between students' gender, age, study situation, physical and mental health and students' consumption behavior in this study, correlation and multivariate ANOVA analysis were performed on the related data at macro level by SPSS28.0.

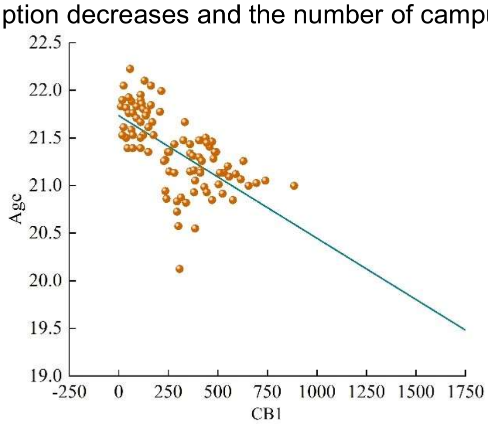
(1) Correlation analysis

Taking age as a variable, the relationship between the four student campus consumption behaviors and age was analyzed using correlation as shown in Table 3. It can be seen that all four data of students' campus consumption behaviors and age present significant levels, and the Pearson correlation [20] coefficients are all negative, that is, there is a significant negative correlation between consumption behaviors and age.

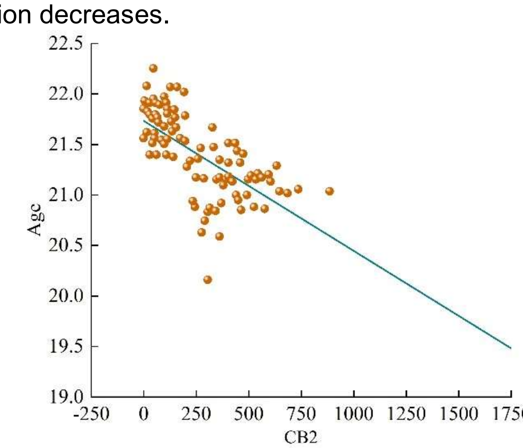
Table 3: Correlation analysis

		CB1	CB2	CB3	CB4
Age	Correlation coefficient	-0.088**	-0.084**	-0.082**	-0.019**
	Significance	0.000	0.000	0.000	0.000

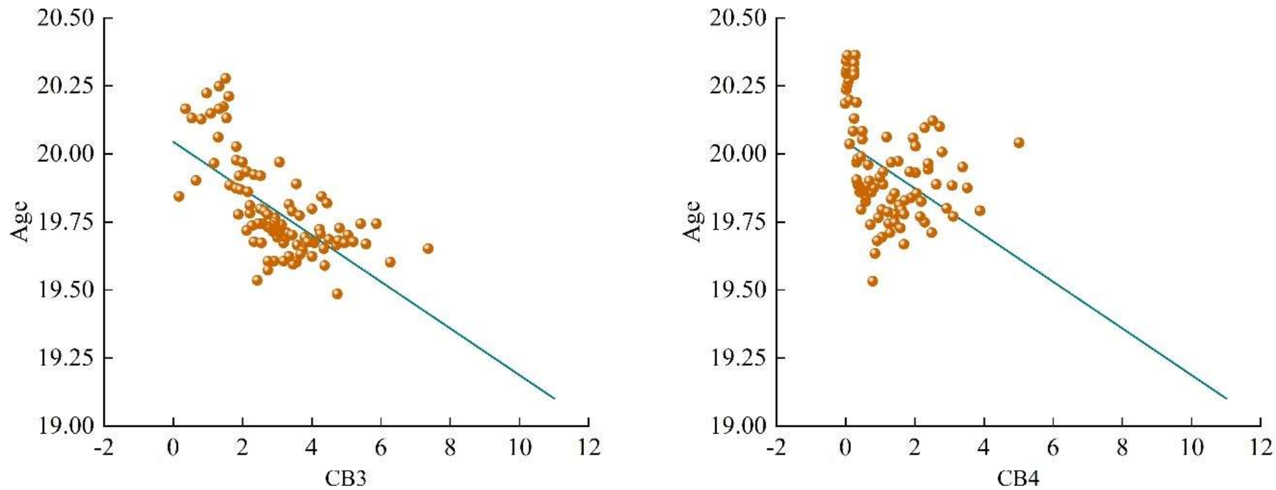
Meanwhile, the scatter plot of the relationship between students' campus consumption behavior and age in colleges and universities is shown in Figure 3, where (a) ~ (d) represent the relationship between CB1, CB2, CB3, CB4 and age, respectively. It can be seen that there is a decreasing linear trend between the two variables of students' campus consumption behavior and age, indicating that as students' age increases, the amount of campus consumption decreases and the number of campus consumption decreases.



(a) CB1



(b) CB2



(c) CB3 (d) CB4

Figure 3: Scatter plot of the relationship between students’ consumption behavior and age

(2) Multivariate ANOVA

Considering the interrelationship between multiple independent variables and multiple dependent variables in the study, the multivariate ANOVA was used to analyze the relevant effects of students' campus consumption behaviors and students' subjects as shown in Table 4, which shows that, in the basic situation of the students, the significance of the effects of gender and age on their campus consumption behaviors is 0.000<0.001, reaching the level of great significance. In the academic profile, the significance of the number of library books borrowed on the number of average daily food and beverage consumption reaches the highly significant level, and for the average monthly amount of food and beverage consumption and the average daily number of daily life consumption are both significant at the 5% level. The p=0.025<0.05 for average academic performance ranking of students on average number of daily life consumption is also significant. Among the physical and mental health conditions, the effect of the number of abnormal body temperatures on the average number of meals consumed per day and the average number of meals consumed per day in daily life was significant, meanwhile, the effect of the mental health conditions only on the average number of meals consumed per day reached the level of significance with the corresponding p=0.006<0.05.

IV. B. Cluster Analysis of Students' Online Behavior

IV. B. 1) Clustering results and basic information on various student groups

The K-prototype algorithm was applied to cluster students' online behaviors, and the final output was the category labels of each student. The number of student groups in each category and the gender ratio are shown in Table 5. Through clustering analysis, 3413 students were divided into 4 categories, with 251, 974, 1816, and 372 students in category 1, category 2, category 3, and category 4, respectively.

IV. B. 2) Characterization of Campus Internet Usage by Various Student Groups

In terms of Internet access hours, the total Internet access hours, daytime hours, nighttime hours, midweek and weekend hours of categories 1, 2 and 4 are close to each other, while category 3 is slightly higher than categories 1, 2 and 4, and there is no significant difference in terms of Internet access hours among the categories in general.

In terms of Internet traffic, the total traffic, daytime and nighttime traffic of category 3 is slightly higher than that of categories 1, 2 and 4, and category 1 has the lowest Internet traffic in all statistical values. In terms of traffic time ratio, the traffic time ratio of category 1 to category 4 increases in order.

Further analyzing the 40-month traffic time ratios of each category of students as shown in Figure 4, it can be found that after entering the second year of college, the traffic time ratios of each category of student groups have increased significantly and have a similar pattern. After entering the third year of college, the flow time ratios of category 4 and category 3 increase significantly, and category 4 is significantly higher than the other three categories.

Further analysis of the traffic time ratios of daytime and nighttime Internet access for the 4 categories of student groups showed similar patterns of traffic time ratios for each category during the daytime during semester months 1-40, while the traffic time ratios for categories 3 and 4 at night increased significantly after entering the junior year, and category 4 was significantly higher than the other 3 categories.

Table 4: The effect test between students' consumption behavior and the students themselves

			df	Mean square	F	Sig.	Partial Eta squared
Basic situation	Gender	CB1	1	16287395.84	424.817	0.000	0.012
		CB2	1	18527436.67	475.682	0.000	0.014
		CB3	1	119.4	72.814	0.000	0.003
		CB4	1	275.649	174.653	0.000	0.005
	Age	CB1	4	3062859.473	92.342	0.000	0.012
		CB2	4	3245196.245	95.849	0.000	0.013
		CB3	4	32.594	19.631	0.000	0.003
		CB4	4	294.573	182.834	0.000	0.024
Learning situation	The number of times books have been borrowed from the library	CB1	4	148627.591	4.544	0.004	0.000
		CB2	4	132744.283	4.285	0.008	0.000
		CB3	4	6.195	3.07	0.012	0.000
		CB4	4	17.483	11.286	0.000	0.002
	Average academic performance ranking	CB1	9	19524.82	0.507	0.914	0.000
		CB2	9	15629.34	0.384	0.969	0.000
		CB3	9	4.627	3.165	0.025	0.002
		CB4	9	3.248	2.423	0.228	0.000
Physical and mental health condition	The frequency of abnormal body temperature	CB1	6	51584.392	2.456	0.257	0.000
		CB2	6	60735.473	2.642	0.183	0.000
		CB3	6	5.462	3.759	0.015	0.002
		CB4	6	8.631	5.614	0.000	0.000
	Mental health situation	CB1	5	42837.6	1.237	0.364	0.000
		CB2	5	42493.52	1.185	0.385	0.000
		CB3	5	2.153	0.714	0.659	0.000
		CB4	5	7.264	3.825	0.006	0.000

Table 5: The number and gender ratio of various student groups

Category	Number of people	The number of male students	Proportion of male students	The number of female students	Proportion of female students
Category 1	251	65	25.90%	186	74.10%
Category 2	974	802	82.34	172	17.66
Category 3	1816	1428	78.63	388	21.37
Category 4	372	264	70.97	108	29.03

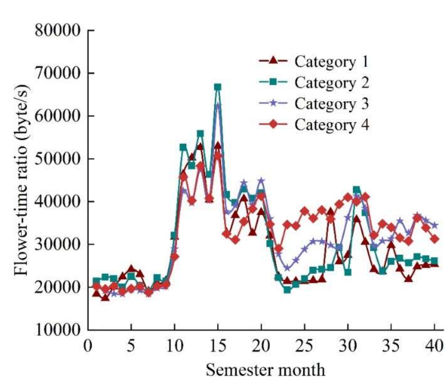


Figure 4: The ratio of online data usage time for various student groups

The single-time Internet traffic of each type of student group in each semester month is shown in Fig. 5. The values and patterns of the average single-login traffic of the 4 types of student groups are similar in the freshman and sophomore years, and after entering the junior year, the traffic-times ratio of category 4 rises significantly, and continues to be at the highest value of the several categories in the junior and senior academic years. In terms of time-to-count ratios, Category 1 students were lower than the other student groups for the first three academic years, and slightly higher than the other categories in the fourth academic year.

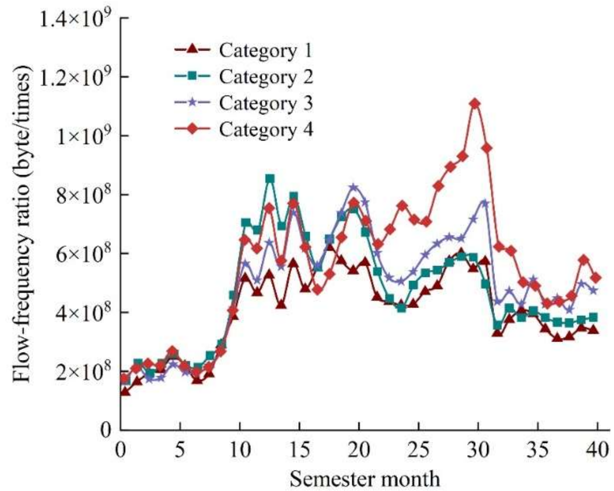


Figure 5: Single Internet traffic for various student groups

IV. B. 3) Analysis of other school attendance data for various groups of students

The network data reflects the characteristics of campus network usage of all types of students, and combined with other school behavior data, it helps to comprehensively understand the characteristics of all types of students, so as to further analyze the impact of different types of campus network usage behaviors on students' performance in school. Other school behavior data involved in this study mainly include achievement data, campus card consumption data, total number of books borrowed, and statistical values of physical fitness test data.

In terms of grades, this study counted the arithmetic mean and weighted mean scores of students in each category, and the results showed that the average grades of student groups in categories 1, 2, 3, and 4 decreased in order. In terms of physical fitness test, the long-distance running scores of categories 1, 2, 3, and 4 decreased category by category, and the comprehensive physical fitness test scores of category 1 students were higher than those of the other three categories, while the scores of categories 2, 3, and 4 students were close to each other.

In terms of breakfast consumption, this study counted the number of times each type of group consumed breakfast in the cafeteria, and the results showed that the number of times category 4 consumed breakfast in the cafeteria was significantly lower than the first 3 categories.

In terms of the number of books borrowed, there was no significant difference between the various groups of students.

IV. C. In-depth Application of Campus Behavioral Portrait of College Students

IV. C. 1) Analysis of study hours and mental health

Through the analysis of college students' campus behaviors, it was found that there is a complex association between study time commitment and mental health status. Usually, the mental health scores of the student group with longer study hours are higher, but there are some students whose high study hours did not bring mental health gains or even declined, suggesting the negative impact of too much study pressure.

The scatter plot of study hours and mental health is shown in Figure 6. The distribution of scatter points in the figure indicates that most students' mental health scores were positively correlated with the number of study hours, which implies that moderate study hours have a potential role in promoting students' mental health. This observation may be attributed to the accumulation of a sense of academic achievement, a key factor in mental health improvement. At the same time, the data also revealed an interesting phenomenon, that is, the existence of a group of students who were able to maintain a high level of mental health within a relatively limited study time. This finding may imply a deeper meaning of the differences in time management efficiency, or it may indicate that this group adopts more diversified and effective mental health maintenance strategies, such as regular physical exercise, active participation in leisure and recreational activities, and other non-study activities, which have been shown to have a significant effect on mental health.

In summary, emphasizing the non-linear relationship between study hours and mental health, and highlighting the role of individual differences, requires educators to pay attention to the management of students' study stress and to advocate diversified ways of mental health maintenance.

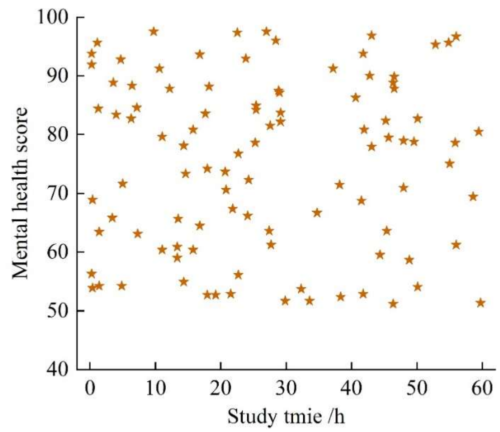


Figure 6: Scatter plot of study duration and mental health analysis

IV. C. 2) Loneliness score analysis

In order to improve students' mental health, achieve mental health management of students in higher education, and promote students' personalized development, this section explores the potential impact of loneliness on individual mental health. The scatter distribution of mental health scores and loneliness scores is shown in Figure 7. Each data point in the figure corresponds to one student, and the horizontal and vertical axes quantify the degree of loneliness and the mental health status score of the students, respectively.

Observation of the data distribution shows that there is a significant correlation between loneliness scores and mental health scores in most student groups, which is manifested in the fact that students with higher loneliness scores generally have lower mental health scores, and loneliness has a negative impact on mental health. Loneliness may induce depression, anxiety and social isolation in students, and these psychological states in turn adversely affect mental health. However, it is worth noting that there are several anomalous data points in the graph, where some of the students have high loneliness scores but their mental health scores remain relatively high.

In addition, another group of students was observed in the graph, who had low loneliness scores but also low mental health scores. This phenomenon suggests that loneliness is not the only factor affecting students' mental health, and that other external stressors such as academic pressure and interpersonal relationship problems may independently or interact with students' mental health, resulting in impaired mental health. Therefore, when exploring students' mental health problems, it is necessary to consider multidimensional influencing factors in order to develop more comprehensive and individualized intervention strategies.

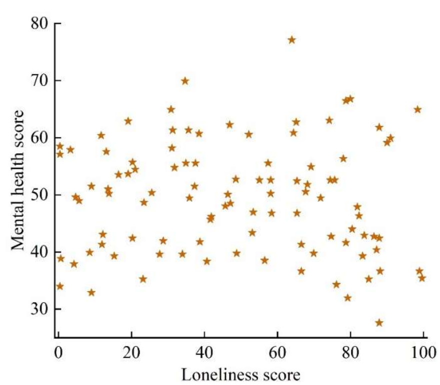


Figure 7: Scatter distribution of mental health scores and loneliness scores

V. Conclusion

Based on data mining technology, this study conducted a detailed analysis of the consumption behavior of students in colleges and universities, using the K-prototype clustering algorithm to classify students into four major consumption groups, in which the average monthly consumption amount of students in category 3 is 291.61 yuan and the average daily consumption frequency is low, while students in category 4 have the highest level of consumption, with an average monthly consumption of 961.56 yuan. The data show that students' consumption behavior is closely related to their age, gender and other factors, especially students' consumption behavior is significantly negatively correlated with their age. In addition, the study also found that students' academic performance, the number of books borrowed and other factors have a certain impact on their consumption behavior, especially students with higher academic performance, their average daily consumption frequency is relatively low.

In terms of mental health, there is a complex relationship between the investment of study time and the state of mental health. Moderate study time helps to improve mental health, but excessive study pressure may have a negative impact on students' mental health. Therefore, educational administrators need to pay more attention to students' individual needs and consider students' multidimensional data comprehensively when making educational decisions. This study provides a theoretical basis and practical guidance for the promotion of smart education and personalized education management for students, and promotes the pace of digital transformation in education.

References

[1] Garanin, M. A., & Krasnova, E. A. (2019, October). Management Model of Innovative University. In Institute of Scientific Communications Conference (pp. 1463-1475). Cham: Springer International Publishing.

[2] Aniskina, N., & Terekhova, E. (2019). Innovative methods for quality management in educational organizations. International Journal of Quality & Reliability Management, 36(2), 217-231.

[3] Aithal, P. S., & Aithal, S. (2015). An innovative education model to realize ideal education system. International Journal of scientific research and management (IJSRM), 3(3), 2464-2469.

[4] Toytok, E. H. (2016). School Leaders' Innovation Managements and Organizational Stress: A Relational Model Study. Universal Journal of Educational Research, 4(n12A), 173-179.

[5] Sciarelli, M., Gheith, M. H., & Tani, M. (2020). The relationship between quality management practices, organizational innovation, and technical innovation in higher education. Quality Assurance in Education, 28(3), 137-150.

[6] Iqbal, H. M., Parra-Saldivar, R., Zavala-Yoe, R., & Ramirez-Mendoza, R. A. (2020). Smart educational tools and learning management systems: supportive framework. International journal on interactive design and manufacturing (IJIDeM), 14(4), 1179-1193.

[7] Yang, X., & Su, W. (2021, July). Research and design of blended teaching mode based on smart learning environment. In 2021 International Conference on Education, Information Management and Service Science (EIMSS) (pp. 136-140). IEEE.

[8] Deng, C., Feng, L., & Ye, Q. (2024). Smart physical education: governance of school physical education in the era of new generation of information technology and knowledge. Journal of the Knowledge Economy, 15(3), 13857-13889.

[9] Díaz-Parra, O., Fuentes-Penna, A., Barrera-Cámara, R. A., Trejo-Macotela, F. R., Ramos-Fernández, J. C., Ruiz-Vanoye, J. A., ... & Rodríguez-Flores, J. (2022). Smart Education and future trends. International journal of combinatorial optimization problems and informatics, 13(1), 65.

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

[11] Mulawarman, W. G. (2022). The Effect of Smart Management and School Efficiency on School Performance in the Digital Era. Eurasian Journal of Educational Research (EJER), (100).

[12] Vinogradova, N. V., Popova, T. N., Chehri, A., & Burenina, V. I. (2020, December). SMART technologies as the innovative way of development and the answer to challenges of modern time. In ITM Web of Conferences (Vol. 35, p. 06010).

[13] Aldahwan, N., & Alsaeed, N. (2020). Use of artificial intelligent in Learning Management System (LMS): a systematic literature review. International Journal of Computer Applications, 175(13), 16-26.

[14] Long, X., & Li, G. (2024, September). Scenario Construction of Smart Education in Universities: Teaching Practice Based on Management Courses. In 2024 14th International Conference on Information Technology in Medicine and Education (ITME) (pp. 363-367). IEEE.

[15] Gupta, A. K., Aggarwal, V., Sharma, V., & Naved, M. (2023). Education 4.0 and Web 3.0 Technologies Application for enhancement of distance learning management Systems in the Post–COVID-19 ERa. In The Role of Sustainability and Artificial Intelligence in Education Improvement (pp. 66-86). Chapman and Hall/CRC.

[16] Abdallah, A. K., & Abdallah, R. K. (2025). Smart Solutions for Smarter Schools: Leveraging Artificial Intelligence to Revolutionize Educational Administration and Leadership. In Encyclopedia of Information Science and Technology, Sixth Edition (pp. 1-14). IGI Global.

[17] Bo Li, Mengle Chen, Qi Zhao, Xu Yang & Xiaofeng An. (2024). Research on student behavior recognition method based on human physiological information perception. Journal of Computational Methods in Sciences and Engineering,24(6),3574-3591.

[18] Yonglin Zhao. (2024). Research on the Path of Internationalization Management in Higher Education Based on K-modes Algorithm. Exploration of Educational Management,2(5).

[19] K. Lakshmi, N. Karthikeyani Visalakshi & S. Shanthi. (2017). Cuckoo Search based K-Prototype Clustering Algorithm. Asian Journal of Research in Social Sciences and Humanities,7(2),300-309.

[20] Sheng Wang & Jiabi Zhao. (2025). Analysis of Student Learning Behavior on Online Education Platforms Based on Deep Learning. Journal of Circuits, Systems and Computers,34(06).