College Student Management System Based on K-means Clustering Algorithm

Chunmei Yu, Yun Wang*

Sichuan Vocational and Technical College, Suining, China
snwy_1020@163.com
*Corresponding author

Abstract: Student management is a problem that every university needs to solve. The purpose of this paper is to design a college student management system based on K-mean clustering algorithm. The purpose of this campus project analysis and research is to help teachers fully understand, master and adapt to students' skills, achieve the purpose of teaching, and write and develop optimized K-means algorithms. Firstly, through the analysis of the student management data, the system function module is added to the student management system. Then the number of K clusters is selected to set the total range, and the optimal value of the number of K clusters is selected by calculating the ratio of the inner and outer distances. Use the K-means method to analyse the optimization of the algorithm to analyse the performance on student management. The optimized K-mean algorithm has completed the classification of students well, and the average score of Pmoral Learning is 17.54 for the first category. College student management is a very important innovation, which has played a certain role in promoting the development of student management.

Keywords: K-means Clustering; College Students; Student Management; System Design

1. Introduction

With the rapid development of information technology in colleges and universities in my country, the rapid development of information technology and the continuous development of college students' informatization have brought many technical problems to the way of information transmission and storage [1]. The departmental functions of each school provide various types of information required by each department [2]. No public website stores and shares this information, and no one is responsible for updating it [3]. In this way, even if a small amount of news media is available, the information is unreliable [4].

Various functional departments of colleges and universities repeatedly store a large amount of data, and the relevant data cannot be displayed in a timely and efficient macroscopic manner, which also exposes the drawbacks of the overall statistical status of data information [5]. Oguguo B identified the impact of learning management systems (LMS) on student performance in educational measurement and assessment courses. The study employed a non-equivalent group quasi-experimental research design using the entire undergraduate population of Imo State University, Nigeria. Data collection was performed using an instrument called the Measurement and Assessment Achievement Test (MEAT). The results of the study showed that students taught using LMS (Moodle) performed better than students using the CAI4ME package [6]. Aziz E designed and developed an Interactive Student Activities Management System (iSAMS) as a one-stop hub to help UiTMCM Student Affairs Department facilitate the management of student activities. This web-based system is capable of tracking, displaying and storing all event information from the request for proposal stage to closure, including tracking attendance, event progress, collecting feedback, reporting completed events, storing supporting documents and generating reports for administrative purposes [7]]. It can be seen that the further development of college student management system plays a vital role in the healthy development of colleges and students.

The program development of this paper aims to meet the current needs of private college students' information management, and to create a cohesive system to deal with student information management as much as possible. The college student management software program provides students with a series of program activities such as basic student information, attendance information, and scholarship management. Using this program, colleges and universities can effectively organize student

information and use it to make informed decisions that create a connection between a student and his or her students. On the other hand, students can also easily obtain the information they need through the management system and leave their views on the school's current curriculum and education.

2. Research on the Design of College Student Management System Based on K-means Clustering Algorithm

2.1 B/S System

B/S system (Browser/Server, browser/server) is also called B/S system [8]. Due to the advancement of network technology, it is a major extension of C/S services. In this B/S system, the user interface is implemented through browsers [9]. Therefore, from this feature, we can see that the biggest advantage of B/S service is simple operation and maintenance, and users can access and process data from different types of networks [10].

The biggest feature of B/S service is that users can access a large amount of data and multimedia information on the network through standard browsers, and these information are based on Web servers [11]. Each web server has a physical connection to a data server, so it is indeed possible to store a large amount of data on one data server [12].

2.2 K-means Clustering Algorithm

The k-mean algorithm, also known as k-mean or k-mean, is a classic compound algorithm used for many computing tasks. It is the distance of the database as a function of quality optimization from the original database. The algorithm performs iterations, finds the end of the function, and divides the dataset into categories such as a reduced J-score index, creating a similar space between categories.

The search of the k-mean compiled algorithm is limited to a small fraction of the total possible space. If the sample size between each class is small, the k-mean algorithm tends to achieve better results. However, if the sample similarity between classes is high, further clustering may occur. Therefore, due to the ensemble of algorithms, it is possible to obtain the evaluation function of the smallest solution for a region than the whole world.

The K-algorithm method is used to measure specific points in the sample. When encountering compilation problems, the algorithm is more efficient at the end of the task. The algorithm is highly scalable and efficient even in the case of large databases. Because its complexity is O (n k t), where n is the number of all objects, the k value is user-defined, and t is the number of iterations.

When the k-mean clustering algorithm is used, the difference between clusters is obvious, and the clustering effect is better.

The main disadvantage of the K-Means algorithm is that setting different k values often leads to completely different results. You can use this algorithm to analyze data distributions, such as centers, control groups, and densities, and choose an appropriate value for k until the center of the cluster stops.

2.3 User Situation of Student Management System

The roles of system users can be divided into three categories, namely: system administrator, teacher and student. If the users of the entire student management system are counted as a basic class, then the student is the first-level generalization of the users of the basic class, and this role has the lowest power in the system. System administrators and teachers are another generalization of ordinary users. Among them, the system administrator user has the highest system authority, can have authority over all aspects of the entire system, and can delegate authority to other users.

(1) Students

Students are the users with the least power and the most important type of users in the system. The purpose of this plan is to allow schools to better organize and manage students, and to make it easier for students to understand their school capabilities and learning. Therefore, whether students can achieve flexible and controllable goals is an important indicator to measure.

(2) Teachers

Teachers are the most interactive roles with students during school. Teachers not only impart

knowledge to students, but also guide students' life and psychology. Teachers are the main body who directly communicate with students, understand and evaluate students.

(3) System administrator

The system administrator has the highest authority of the student management system, and is responsible for the assignment of user authority and the responsibilities of user management in the system.

3. Investigation and Research on the Design of College Student Management System Based on K-means Clustering Algorithm

3.1 Design of Student Management System

The student management system adopts the B/S three-tier mode structure to develop the system. The three-layer architecture is divided into three different layers, namely the WEB display layer, the business logic layer and the data access layer. Each layer is responsible for completing different tasks, which improves the efficiency of system development and facilitates system developers. Perform system testing and maintenance.

3.2 Database System Code Design

The database system development must use the corresponding development environment, the database system used in the design is the Oracle system, so this article chooses PL/SQL Developer as the development environment. PL/SQL Developer is a complete development environment dedicated to developing program units stored in an Oracle database. PL/SQL Developer has many intelligent functions such as syntax enhancement, SQL language and PL/SQL language help content, detailed object description, code development assistance, intelligent message editing, etc., which can meet the needs of high-performance users.

3.3 Optimized K-means Algorithm

In this paper, a selection method for solving the initial value K is proposed. The method is to limit the scope of clustering according to the actual situation. The algorithm will run nm times and select an optimal number of clusters from these clusters as the optimal number of clusters according to the following formula. The calculation of the V value is shown in Equation 1:

$$V = \frac{d_{\text{Inside distance}}}{d_{\text{Outer distance}}}, (d_{\text{Inside distance}} = \sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - c_i)^2, d_{\text{Outer distance}} = \frac{1}{k(k-1)}\sum_{i=1}^{K}\sum_{j=1}^{K}(c_i - c_j)^2) \quad (1)$$

Where k is the number of centers. The meaning of this type is that when the ratio of coherence to external distance is the smallest, it means that the coherence within a cluster is high and the coupling between clusters is low, that is, the number of clusters is the optimal number of clusters.

Regarding the selection of the starting point, the method adopted in this paper is: the starting point must meet the following conditions: the distance between these center points is as large as possible. The points around these initial centers should be dense. Regarding the reciprocal distance of the center point, as shown in formula 2:

$$d = \frac{1}{k(k-1)}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(c_i - c_j)^2 \quad (2)$$

This is the average of the sum of all center point distances and better reflects the total size of the distances between complex centers.

4. Analysis and Research on the Design of College Student Management System Based on K-means Clustering Algorithm

4.1 Implementation of System Function Modules

The achievement and employment analysis subsystem, as an important part of the college student management system, concentrates the main business logic of the system, so this article will take the achievement and employment subsystem as an example to introduce the implementation of the system.

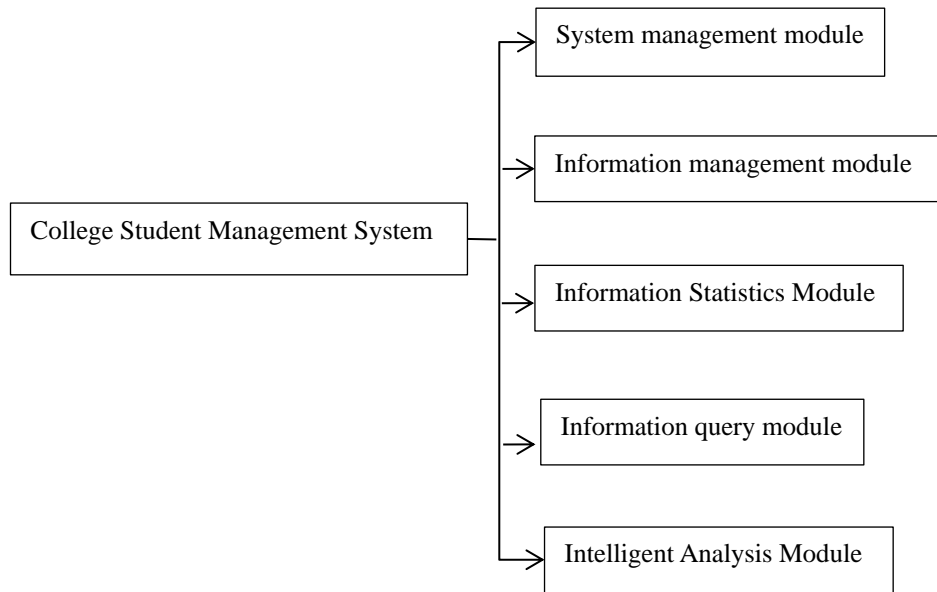The structure diagram of the college student management system is shown in Figure 1:

Figure 1: System block diagram

The modules of this program are the main parts of system control module, information control module, statistical information module and intelligent analysis module. System management: mainly focuses on user access and user control; Information management: mainly has the functions of adding, deleting, modifying and querying teacher information, student information, course information, grade information, class information, etc.; Statistics of poor students and statistics of students and information topics are mainly used to find information-rich control information management documents; intelligent analysis: analyze the relationship between courses, the relationship between courses and employment, and divide them into intelligent analysis subsections according to different users. Module (for the functions of student job selection, course warning, course performance analysis, employment analysis, promotion analysis, etc.); Teacher intelligence analysis sub-module (for teacher course analysis, student job ability analysis, course and post correlation analysis, course effect analysis, student Employment forecasting and other functions); education and management intelligent analysis sub-module (enrollment forecasting, employment forecasting analysis, student learning early warning, employment early warning and other functions).

4.2 The Results of the Optimization Algorithm Applied to the Student Management Analysis System

The main function of student analysis management is to perform cluster analysis on the student campus activity data in the system, divide students into several characteristic categories and point out the main characteristics of each category. After a period of operation, the classification results will be displayed in the form of a histogram. After analysis, the previous analysis interface will also give students' classification information, category information (category information is no longer -1) and average results of various aspects, etc. According to the default method, this system optimizes K-means classification the result is shown in Figure 2:
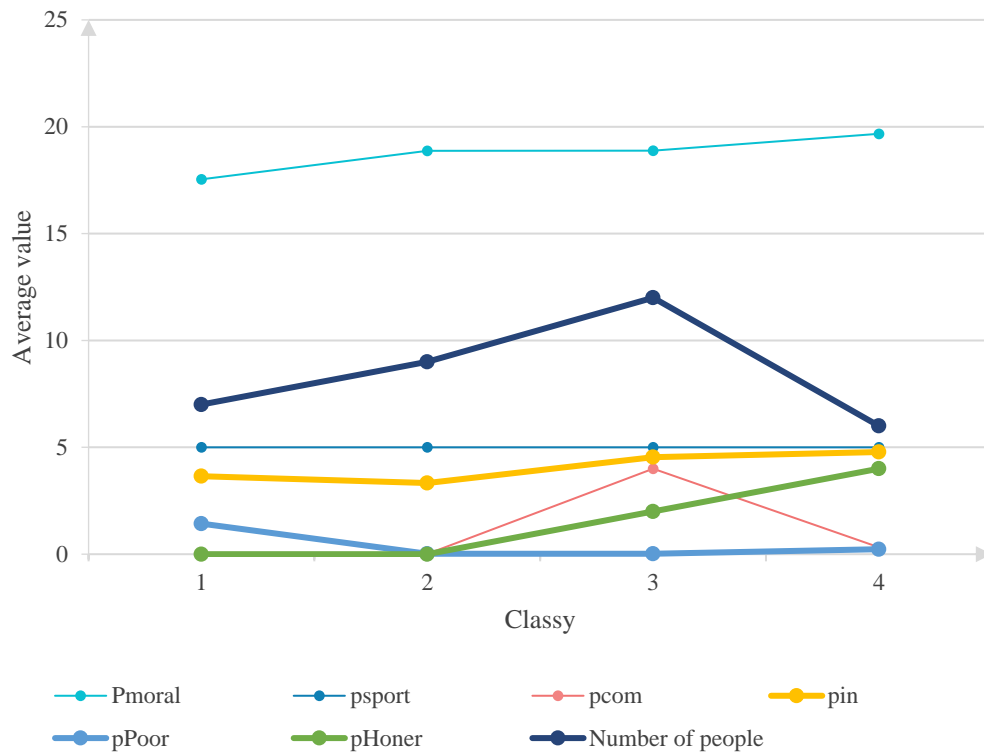
Figure 2. Display of the average value within the student category

Table 1. Average within student category

Pmoral	psport	pcom	pin	pPoor	pHoner	Classy	Number of people
17.54	5	0.03	3.65	1.43	0	1	7
18.87	5	0.01	3.33	0.02	0	2	9
18.88	5	4	4.54	0.02	2	3	12
19.67	5	0.32	4.78	0.23	4	4	6

Among them: Pmoral is the average score of the moral education score within the class, psport is the average score of the sports score within the class, pcom is the average score of the competition score within the class, pin is the average score of the intellectual education score within the class, and pPoor is the average score of the poverty level within the class, pHoner is the average score of the scholarship grade within the class. Classy is the class number, and the number of people is the number of people in the class. The optimized K-means algorithm divides all students into four categories, each of which has more prominent characteristics, as shown in Table 1. The characteristics of the first category are that they have average academic performance, most of them are poor students, who have not won scholarships and have no competition points. The second category is characterized by lower academic performance, no scholarship, no poor students, and no competition. The characteristics of the third category are that the competition results are very outstanding, the scholarship is very high, and the grades are in the middle and upper reaches, but they are not poor students. The characteristics of the fourth category are that they have the best academic performance and good scholarships, and most of them are poor students with average performance in competitions. It can be concluded that the optimized K-means algorithm has completed a relatively good classification of the students, and there is a relatively large gap between the classes. And the characteristics of the class are very obvious, which is more suitable for the instructor to manage the classification of students.

5. Conclusion

The research of college student management system is a difficult subject with theoretical and practical significance. This paper discusses the implementation process of college student management system, which has certain application value. Student management system can provide database for principals to customize. They work through data processing. The modular design of the system is closely related to students' professional life and social practice in school, and can timely reflect the

basic data of students' daily performance, providing data support and decision-making basis for students in school.

References

[1] Yin X H. Construction of Student Information Management System Based on Data Mining and Clustering Algorithm [J]. Complexity, 2021, 2021(2):1-11.

[2] Gay A E. Using a Content Management System for Student Digital Humanities Projects [J]. Online, 2019, 43(1):32-35.

[3] Djuraevna A Z. Formation of Student Management Activities in the Higher Education Management System [J]. Psychology (Savannah, Ga.), 2021, 58(2):1494-1499.

[4] Jasmis J, Aziz A, Jono M, et al. An Analysis Model for an Integrated Student Activities Management System for Higher Education during RMO/CMCO/PASCA COVID-19 Period in Malaysia [J]. Procedia Computer Science, 2021, 179(2):798-803.

[5] Alkhateeb M A, Abdalla R A. Factors Influencing Student Satisfaction towards Using Learning Management System Moodle [J]. International journal of information and communication technology education: an official publication of the Information Resources Management Association, 2021, 17(1):138-153.

[6] Oguguo B, Nannim F A, Agah J, et al. Effect of learning management system on Student's performance in educational measurement and evaluation [J]. Education and Information Technologies, 2021, 26(2):1-13.

[7] Aziz E. Design Architecture of an Integrated Student Activities Management System for Higher Education [J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(5):1676-1683.

[8] Chandra T. Evaluation of Student Satisfaction in Using the Learning Management System for Online Learning at XYZ University [J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(6):2810-2816.

[9] Martellucci L, Krishna K K. Analysis of Air-Cooling Battery Thermal Management System for Formula Student Car [J]. Journal of Transportation Technologies, 2021, 11(3):436-454.

[10] Jamaluddin J , Bundu P , Anshari A , et al. Course and Training Models Online Application Services (LAO-KURSUS) and Learning Management System (LMS) to Improve Student Competencies[J]. Universal Journal of Educational Research, 2021, 9(6):1258-1273.

[11] Surendran K. Student Academic Management System Using Blockchain Technology [J]. Journal of Advanced Research in Dynamical and Control Systems, 2020, 12(SP3):1410-1415.

[12] Songsom N , Nilsook P , Wannapiroon P , et al. System Design of a Student Relationship Management System Using the Internet of Things to Collect the Digital Footprint[J]. International Journal of Information and Education Technology, 2020, 10(3):222-226.