

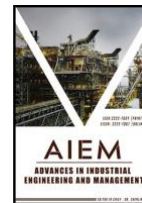


ZIBELINE INTERNATIONAL™

ISSN: 2222-7059 (Print)

EISSN: 2222-7067 (Online)

# Advances in Industrial Engineering and Management (AIEM)

DOI: <http://doi.org/10.7508/aiem.04.2024.282.285>

## RESEARCH ARTICLE

# DESIGN AND IMPLEMENTATION OF SUNSHINE SPORTS SYSTEM FOR COLLEGES AND UNIVERSITIES BASED ON C4.5 ALGORITHM

Yi Shao, Xuefeng Zhang\*

Shanghai Customs College, Shanghai 201204, China

\*Corresponding Author E-mail: [zhangxf2015@163.com](mailto:zhangxf2015@163.com)

This is an open access article distributed under the Creative Commons Attribution License CC BY 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ARTICLE DETAILS

### Article History:

Received 24 April 2024

Accepted 19 July 2024

Available online 30 August 2024

## ABSTRACT

Along with the expansion of the number of college students in China, the number of college students has risen sharply, which leads to a significant increase in anthropometric data involving students, mainly student attendance information and physical fitness test data, etc. Therefore, better data management is a need for physical education in colleges and universities at this stage. Starting from sunshine sports, this study applies the C4.5 algorithm to data mining and establishes an efficient database management system based on students' physical fitness and health testing data, thus guiding colleges and universities to adjust their physical education teaching and management methods in a timely manner in order to comprehensively enhance the physical fitness of college students.

## KEYWORDS

C4.5 algorithm, sports system, database management

## 1. INTRODUCTION

With the increasing expansion of Sunshine Sports activities, the data generated by the Sunshine Sports system is also increasing, including student attendance information and student physical health assessment data. At this time, how to do a good job in data management is an inevitable need at present. In addition, there are often errors in the manual management of these massive information, which is easy to discourage students' enthusiasm for participating in sunshine sports, and cannot achieve the effect of promoting students' fitness. At the same time, it also brings many adverse effects to the development of sunshine sports in Hebei University of technology (Zhang, et al., 2020). Therefore, how to sort out sunshine sports data and analyze and manage students' physical health information conveniently, quickly, with low error rate and high efficiency is an urgent problem to be solved. In addition, with the continuous development of computer technology and the country's attention to the development of computer networks, it will be easy to carry out information management of sunshine sports.

The development of sunshine sports system in colleges and universities is not only able to carry out information management of sunshine sports, but also able to better arrange for students to carry out sports, as well as being conducive to the school's understanding of the physical exertion status of the students and analyses of their physical health status,

so as to make timely adjustments. At the same time, the application of data mining in the physical health assessment data of colleges and universities provides an effective way for schools to obtain valuable and influential important information from a large number of students' physical health assessment data. In the face of different students, we can find out the scores of students in various physical assessments and the relevance of various score data by data mining on various aspects of students' physical health, and reveal the problems of students in physical exercise (Li, 2019). At the same time, the results of data mining can be used to predict the physical health status of students, give early warning to students, make students pay attention to physical health status, and adjust physical exercise status.

## 2. RELATED WORK

### 2.1 C4.5 Algorithm

Decision tree C4.5 algorithm is a classification algorithm in data mining. Its algorithm idea is simple, classification rules extraction is convenient and easy to understand, and it has been widely used. However, the traditional C4.5 algorithm has a better classification effect when the samples of each category in the data set are relatively balanced. In the unbalanced data set, because the proportion of a few samples is small,

### Quick Response Code



### Access this article online

Website  
[www.aiem.com.my](http://www.aiem.com.my)DOI:  
[10.7508/aiem.04.2024.282.285](https://doi.org/10.7508/aiem.04.2024.282.285)

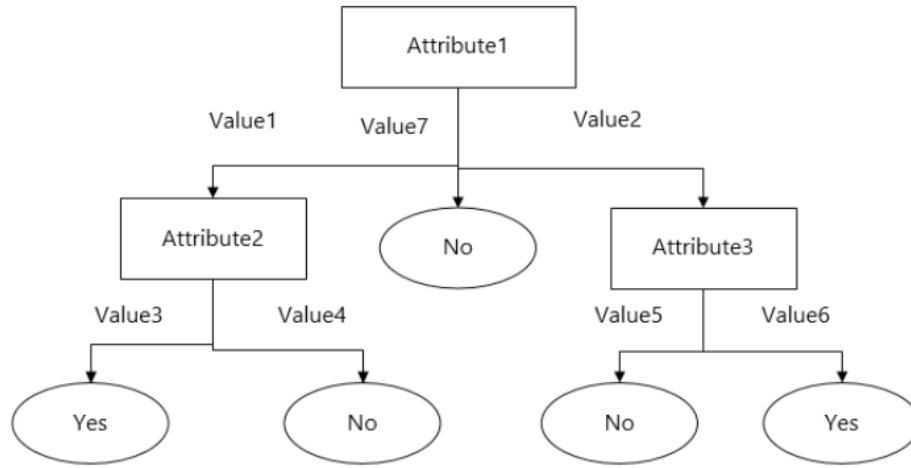


Figure 1: Schematic diagram of decision tree

it can not provide enough classification information for the classifier. In order to ensure the overall classification accuracy of the algorithm, the classifier will pay more attention to the classification accuracy of the majority of classes and ignore the classification of the minority of classes (Yang, 2021). This leads to the phenomenon that although the overall classification accuracy of C4.5 algorithm in unbalanced dataset classification is high, the classification accuracy of a few classes is very low. Figure 1 is a schematic diagram of the decision tree.

ID3 directly uses the gain of information entropy, which will lead to larger information entropy due to more category values of an attribute, so it is easier to be selected as a dividing point.

$$e_k(t) = y_d(t) - y_k(t) \quad (1)$$

$$\diamond x_{k+1}(t) = P^{-1}(t)(f(t, x_d(t)) - f(t, x_{k+1}(t))) \quad (2)$$

C4.5 Deformation of information entropy using information gain rate. The information gain rate takes into account the „cost“ of partitioning information, and can partially offset the effect of the number of attribute values. Therefore, this is one of the important improvements of C4.5 over the ID3 algorithm. On the basis of ID3 algorithm, C4.5 algorithm also adds the processing of continuous attributes and missing attribute values, and has more mature methods for tree pruning.

Nowadays, many data studies have found that these data warehouses and data mining systems established by foreign enterprises have achieved good economic benefits to a great extent. From the perspective of the development of data mining, the research focus of data mining at this stage has not only stayed on the concept of data mining and its application in the system, but also paid more attention to the combination of multiple disciplines (Liu, et al., 2020). At present, data mining urgently needs a systematic scientific theory as its system support, but as far as the current development is concerned, this application field still focuses on theoretical research, and the products that really apply data mining are still very rare.

## 2.2 Current Situation of Sunshine Sports

The problem of adolescent physical decline has become a global problem. It is not only a unique situation in China, but also a serious global social reality. In order to deal with the problem of teenagers' physical decline, many countries have begun to take relevant measures to promote students to strengthen physical exercise and improve their physical condition and health level.

In our country, the study of school physical education has always attracted much attention, which is also an important field of physical education research. Since the promulgation of the decision on the development of sunshine sports for hundreds of millions of students across the country, sunshine sports have had a great impact throughout the country.

Since the implementation of „Sunshine Sports“ in China, it has received extensive attention from the academic community. The academic community has not only studied a wide range of „Sunshine Sports“, but also achieved rich results. At the same time, the majority of sports workers are taking this as an opportunity to constantly find ways and means to strengthen students' physical exercise and improve students' physical health on the platform of school physical education and by means of physical exercise (Gu, et al., 2021).

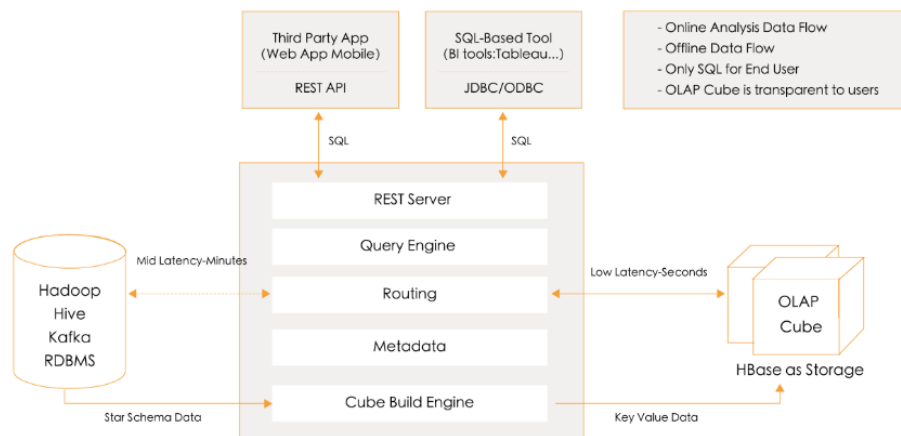
At present, in addition to the annual physical health test of students, there are not many systems that can urge students to carry out „Sunshine Sports“ in China's „Sunshine Sports“, and there are also few systems that use high-tech means such as sunshine sports card swiping terminals to realize the autonomy and intelligence of sunshine sports. The management of „Sunshine Sports“ in Colleges and universities mostly stays in the manual input or paper-based management mode, and the informatization of „Sunshine Sports“ has not been widely promoted throughout the country.

The application of data mining technology in the field of sunshine sports system has not yet taken shape. Colleges and universities mainly use databases in the information construction of sunshine sports system, and do not analyze and mine the data collected by sunshine sports and students' physical health tests to obtain more valuable information (Wang, et al., 2019). However, along with the implementation of „Sunshine Sports“ in higher education institutions, the test data on „Sunshine Sports“ and students' physical fitness in the information management system shows an increasing trend year by year. It is particularly important to use data mining techniques to analyse and mine the test data of „sunshine sports“ and students' physical fitness. Design and implementation of University sunshine sports system based on C4.5 algorithm

Because the system deals with a large number of student assessment score data, how to store these data, and how to input them conveniently and quickly, and how to execute them in the system quickly and output them stably is the key to the operation of the whole system. In order to ensure the smooth operation of the system and the consistency and integrity of the data, it is first necessary to select a database that can run stably. Secondly, we should design a perfect data structure, and finally, we should ensure the smoothness of the network (Hai-Rong, et al., 2018).

At present, the most used database products for developers are SQL Server 2000, Oracle, Sybase ASE and DB2. The following is a simple analysis of the advantages and disadvantages of these four mainstream databases.

First, from the perspective of openness. SQL Server 2000 can only run on windows, and its openness is tightly limited. However, for the operating system, the stability of the system plays a very important role in the database. Oracle can run on all platforms. Support all domestic and foreign industry standards. This product adopts a completely open



**Figure 2:** General structure of sunshine sports system

strategy, which makes it easier for customers to choose the most suitable solution. Sybase ASE can also run on all platforms. However, the early Sybase and OS integration was very low, so it needed to download the necessary patches to operate normally. In addition, in the environment of multiple platforms, there will be some problems when it runs. DB2, like them, can run on all platforms. Its advantage lies in the storage of massive data.

From the perspective of easy operation, SQL server operation is very simple, with only a graphical interface and no DOS interface. Oracle is very complex. It provides both GUI and command line, which is applicable to multiple platforms. Sybase ASE is very complex. It provides both GUI and command line. If you want to use it, you'd better use the command line (Wang, et al., 2019). The operation of DB2 is very simple. It provides both GUI and command line, and is applicable to multiple platforms.

### 3. SIMULATION ANALYSIS

The system adopts ASP technology to support the technical realization of the system. ASP is not a language used to write scripts, but it is a technology that can realize dynamic websites. Its working principle is to provide a running platform on which script language programs can run, so as to achieve the effect of dynamic display of the website.

The files of programs written with ASP technology have the same extension asp. This extension is an important symbol that distinguishes other technologies. If you are browsing a web page and see that the last file name of the URL address of the web page is ASP, so we can be sure that this website is developed with ASP technology. There are many tools for writing ASP programs. It can be said that any tool that can write can write asp programs, such as Notepad and dreamwave (Chen, et al., 2020). The general architecture of sunshine sports system is shown in Figure 2.

The network layer serves as the entry point to the Sunshine Sports System and provides diverse service options for a variety of different user groups. This part is implemented by ZK, and the main service targets include physical education teachers, students, system administrators, school management, and various other users. The service layer plays a crucial role in the overall system. Various business requirements, such as business rules and business processes, are implemented at this layer. This layer of structure can also be referred to as the business logic layer and is realised by springs. The service layer is responsible for receiving page requests and interacting with the underlying data after logical processing (Wang, et al., 2018).

The Dao layer, also known as the data persistence layer, can directly operate the database. The Dao accepts requests from the service layer, obtains data from the data layer, or writes data.

The entity layer is the sunshine sports system database, the database mainly consists of user data, attendance data and physical measurement data. Between the road layer and the entity layer are hibernate as a support.

### 4. CONCLUSION

In terms of data analysis and processing technology, data mining technology is one of the most rapidly developing technologies, and the decision tree algorithm is one of the most important branches of data mining algorithms and has made some achievements in the field of scientific research. Applying decision tree to the sunshine sports system is the focus of this thesis. In the university sunshine sports system constructed on the basis of C4.5 algorithm, "sunshine sports" is regarded as a strategy to encourage students to do physical exercise and enhance their physical fitness. In order to better operate the sunshine sports and build an efficient sunshine sports system, it is also a need for the development of the information age.

### FUNDING

Shanghai Education Science Research General Project (Project No.: C2-2020042) and Shanghai Philosophy and Social Science Planning Project (Project No.: 2021ZTY001).

### REFERENCES

- Chen, N. Y., Chen, L., Xiong, W., 2020. Situation plotting algorithm and system implementation based on Web. *Journal of Physics: Conference Series*, 1584(1): Pp. 10.
- Gu, H., Jin, C., Yuan, H., et al. 2021. Design and implementation of attitude and heading reference system with extended kalman filter based on MEMS multi-sensor fusion. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 29(Supp01): Pp. 157-180.
- Hai-Rong, L. I., Fang, Z. C., 2018. Design and implementation of recommendation system based on Mahout and collaborative filtering algorithm. *Journal of Inner Mongolia University of Science and Technology*.
- Li, S., 2019. Sunlight sports system design and implementation based on decision tree. *Techniques of Automation and Applications*.
- Liu, X., Liao, L., 2020. Research and implementation of performance evaluation system for physical education teachers based on load weight algorithm. *CIPAE 2020: 2020 International Conference on Computers, Information Processing and Advanced Education*.
- Wang, B., University, S., 2018. Research and implementation of spam filtering system based on Naive Bayes algorithm. *Electronic Design Engineering*.
- Wang, Q., Feng, W., Zhuang, C., et al., 2019. Research and implementation of acquisition algorithm based on multi-correlator parallel frequency search for long code.
- Wang, Q., Zhou, G., Rong, H., et al., 2019. Research and implementation of

sign-in system based on mutual difference jump algorithm. Journal of Hefei University of Technology (Natural Science).

of Physics: Conference Series, 1883(1): Pp.6.

Yang, Z., 2021. Research on the application of university teaching management evaluation system based on Apriori algorithm. Journal

Zhang, X., Cheng, Z., Zhang, C., 2020. Design and implementation of a web crawler system based on an adaptive page-rank algorithm. IOP Publishing Ltd. IOP Publishing Ltd, 012021: Pp. 6.

