Design and Implementation of Student Physical Fitness Analysis System Based on Data Mining Technology

Chen Kun Zhang(✉)

Nanchang University College of Science and Technology, Nanchang 330001, Jiangxi, China
zck524065750@163.com

Abstract. The student physique analysis system is based on data mining technology. The main purpose of this project is to develop a system that can predict students' physical health level and provide them with appropriate guidance according to their needs, so that they can improve their health level. Data collected from students can be analyzed to determine their physical health level. In this project, we will use data mining technology to analyze data and develop a model, which can be used as the input of the final output. A model is a set of rules or algorithms used to predict results. This study provides students with opportunities to participate in various activities, such as games, sports and other activities related to their health. The application will help improve students' overall performance by providing sufficient information about their current health status and its impact on other aspects (such as academic, social life, etc.).so it is difficult to give scientific guidance to students through complicated data. Taking the physical test data of colleges and universities as an example, Through data statistics, according to the performance evaluation standard in the national physical exercise standard, this paper obtains five basic indexes: cardiopulmonary function, muscle strength, muscle endurance, softness and obesity. Using decision tree C4 5.

Keywords: Students' physical fitness · Data mining · Physical analysis · System design

1 Introduction

With the comprehensive construction of quality education, physical education in Colleges and universities has also set off an upsurge of educational reform. As early as the national college sports conference held in December 2006, it has been clearly emphasized that "to carry out quality education is not only to pay attention to students' achievement and moral education, but also to cultivate students' strong physique and healthy body and mind". According to the outline of healthy China 2030 plan issued by the CPC Central Committee and the State Council, the development goal of healthy China has been clearly pointed out [1]. In view of the declining trend of the physical health status of China's youth, the National People's Congress revised the sports law of the people's

M. A. Jan and F. Khan (Eds.): BigIoT-EDU 2022, LNICST 465, pp. 85–96, 2023.
https://doi.org/10.1007/978-3-031-23950-2_10

Republic of China twice in 2009 and 2016. Chapter III of the sports law of the people's Republic of China stipulates the status of school sports in talent training, the opening of physical education classes, extracurricular sports activities, sports venues and facilities, the opening of physical education classes Physical education teachers and students have made clear provisions on physical fitness monitoring [2]. Figure 1 below shows the system test and detection process.

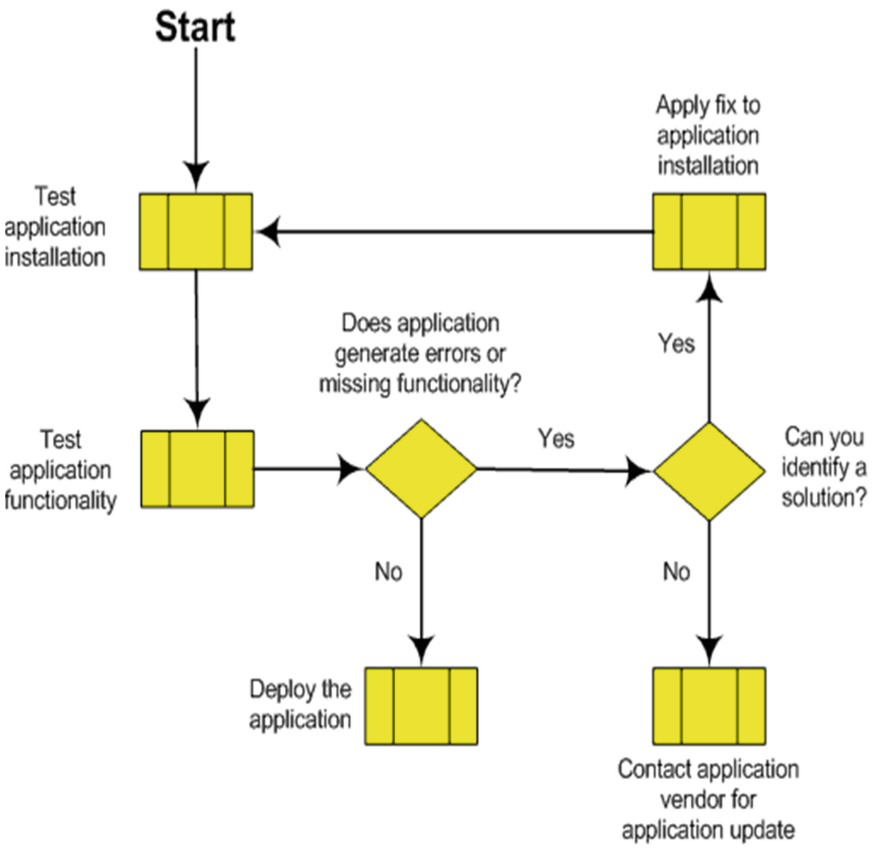

Fig. 1. System test and detection process

The State Council also clearly pointed out that "the physical health of China's youth is an important event related to national rejuvenation and national development. Governments at all levels, Party committees, university organizations and the whole society need to give comprehensive attention and support". However, the survey shows that when the state and society attach great importance to sports construction, there are many reasons for the decline of young people's physical health in China. One of the very important reasons is that young people in China do not cultivate good habits of physical exercise and do not use scientific methods for physical exercise [3].

2 Related work

2.1 Health System Analysis

A healthy constitution is the premise for everyone to carry out other activities. As for the definition of physique, there are different understandings and expressions in different countries. Chinese scholars believe that physique refers to "the quality of human life. A person's physique itself is a relatively stable feature of human form and body function under the comprehensive effect of innate acquisition and acquired learning". Therefore, on the basis of heredity, what level of physical fitness you want to achieve depends on the acquisition of the day after tomorrow [4].

Health means that a person is in good physical, mental and social condition. According to the explanation given by the World Health Organization, health not only refers to whether a person has disease or weakness, but also refers to a person's physical, psychological and social good state [5]. As shown in Fig. 2 below, the framework of healthy constitution analysis is shown.

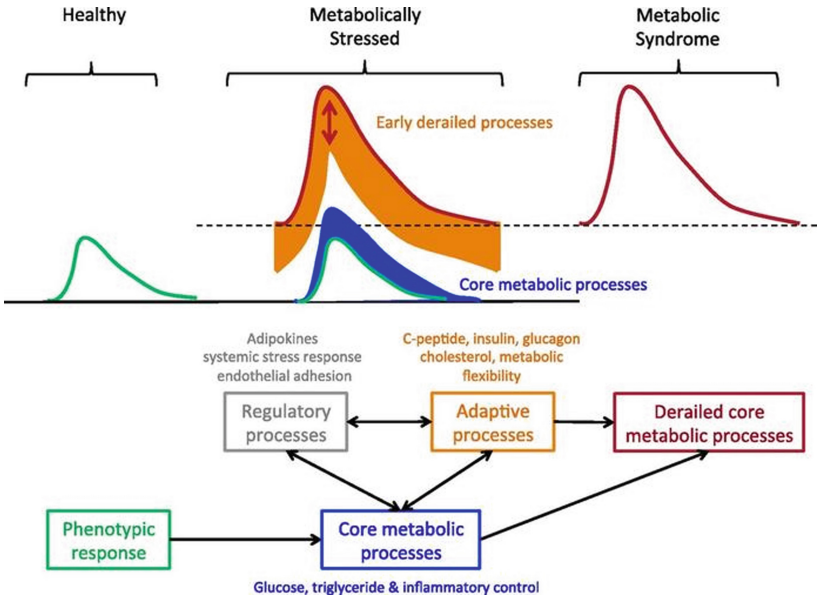

Fig. 2. Analysis framework of healthy constitution

A person's sports ability and physical condition are the main indicators to evaluate his constitution, but also the key to evaluate his health condition. Physical fitness is also called physical adaptability [6]. Exercise ability refers to the ability of the human body to complete special actions coordinated by muscles under the control of the brain. Physical fitness is closely related to sports ability. Figure 3 below shows the framework of the physical fitness test project.

According to the latest standards, there are eight subjects in the physical fitness test of college students. The unique subjects of boys are pull-up and 1000 m running, the

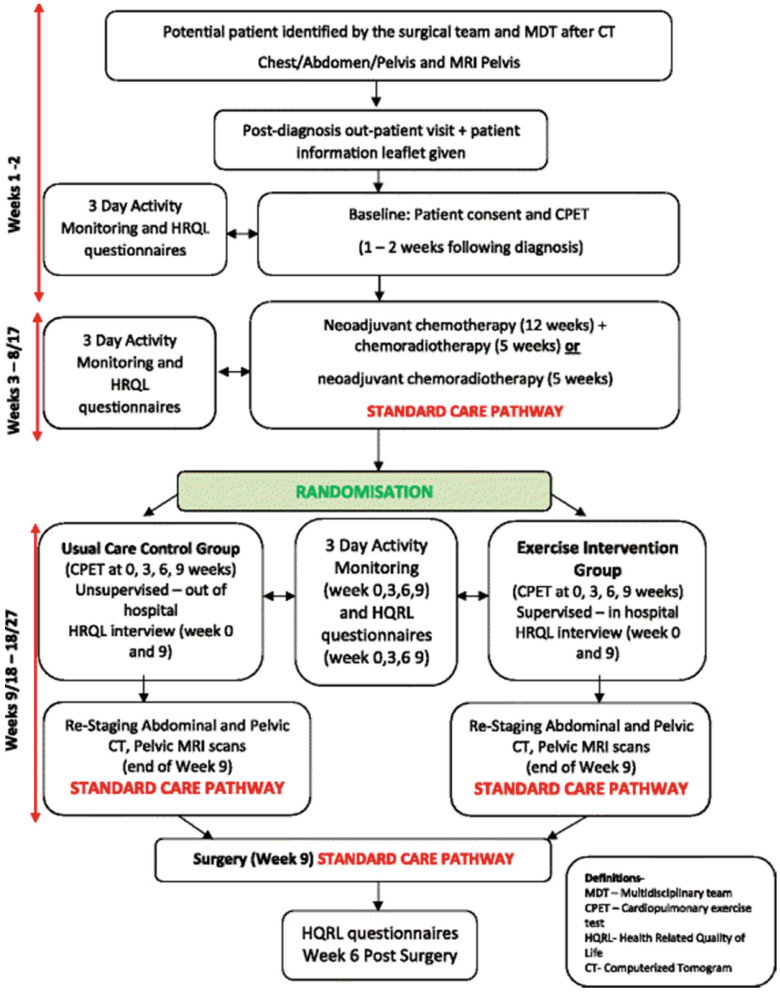

Fig. 3. Physical fitness test project framework

unique subjects of girls are sit ups and 800 m running, and the common subjects of boys and girls are height, weight, vital capacity, standing long jump, 50 m running and sitting posture forward bending. The scoring standard is to convert the scores of various data, remove the test omissions, untested and test free data, and the remaining effective data are statistically divided into four grades: excellent, good, pass and fail, of which the total score of 90 and above is excellent; A total score of 80 to 90 is good; a total score of 60 to 79 is pass; If the total score is less than 60, you will fail.

At present, there are many related researches on fitness testing in China, mainly including the research on the analysis of fitness test performance data and the research on fitness test equipment.

Data analysis and research focus on the following aspects: Research on the current situation of students' physique, research on the causes of students' physique decline, research on the distortion of test data and Research on students' exercise behavior.

The physical health of students is related to the future of the country, and all countries attach great importance to it [7].

Japanese research on physical fitness testing tends to integrate physical factors and living habits. Research by Fuyong et al. Shows that the physical quality and mental state of college students are correlated with exercise habits and sleep habits. Nakajima's research on the correlation between the body mass index and lifestyle factors of freshmen at Fukuoka University shows that the regularity of diet, sleep level, physical exercise, daily fatigue and so on affect the body mass index of students. These studies have provided effective help to promote the physical quality development of college students in Japan. However, the physical fitness test standards implemented in Japan are also very different from those in China, mainly because they refine the age groups and have optional items, such as endurance evaluation. Women can choose endurance running and 20 m round-trip running.

Although the research on physical fitness test abroad is more in-depth, the standards used for physical fitness test are different, and the research focus is also different.

2.2 Data Mining Tasks

At present, the university information management system accumulates all kinds of data every day, and there are more and more scenes of using big data technology for data analysis and mining. Common big data processing technologies include Hadoop ecosystem using MapReduce computing framework, impala providing fast interactive queries, and storm, a real-time stream processing framework with low latency. The data mining information task framework is shown in Fig. 4 below.

Compared with the above technologies, spark technology has advantages in the comprehensive utilization of resources everywhere. With the continuous development of spark, it has become a core library with spark core, including machine learning (mlib), spark streaming, instant query (spark SQL), graph computing (graphx) and other components. Compared with the MapReduce computing framework based on disk computing adopted by Hadoop, the computing speed of spark can reach more than 10 times that of MapReduce [8]. This is because spark is based on memory computing, while Hadoop is based on disk computing. Because spark computing framework uses memory for computing, coupled with its component sparkstreaming, spark has an obvious advantage in real-time stream processing.

The core abstraction of spark is elastic distributed data set, which is called RDD for short. RDD represents a data set that is read-only, partitioned, and provides a series of rich operation modes. Users can operate RDD through rich API commands, which are mainly divided into two types: transformation and action.

In data analysis, this paper uses sparksqli to establish a relational database, uses HDFS to store files, and uses spark computing framework to build a big data analysis platform.

On the other hand, data mining can also scientifically predict the development of things in the future, and then give full play to the maximum value of data.

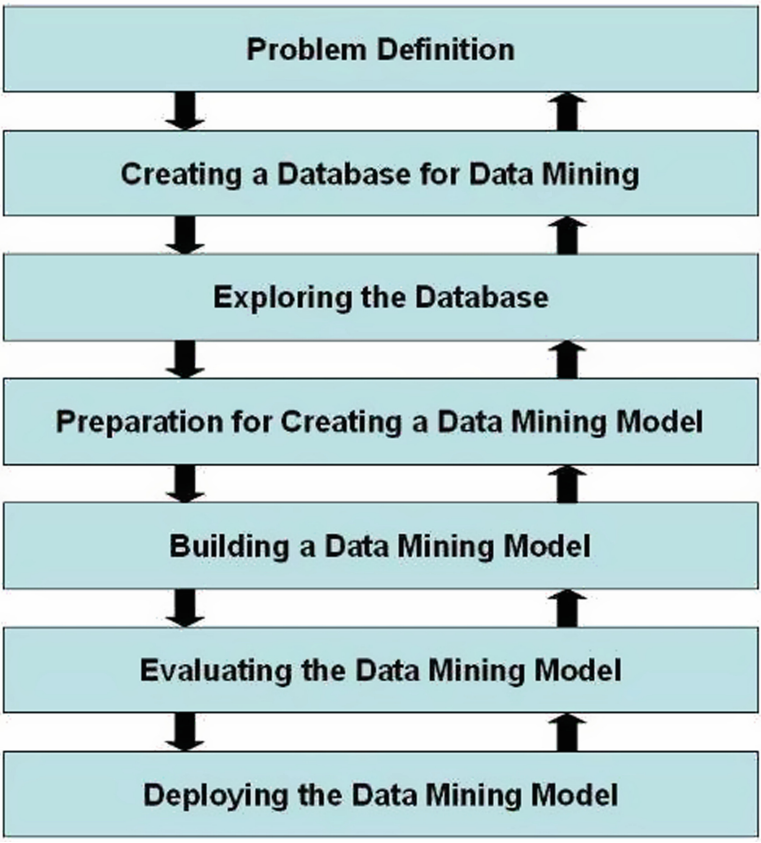

Fig. 4. Data mining information task framework

In the education industry, most studies are aimed at student achievement. Scholars explore the internal relationship between education data and student achievement through data mining technology. Promote the transformation of education mode and gradually change the traditional teaching management mode into a personalized teaching system.The prediction method is to extract the test data set from the original data, and then construct the prediction model to analyze and predict the data, so as to evaluate the knowledge information contained in the unknown data.

Cluster analysis is an unsupervised learning method, which classifies according to the specific characteristics of different individuals. Individuals clustered in the same category have high similarity, but there are differences between different categories. Common clustering analysis algorithms include k-means algorithm, hierarchical clustering algorithm and so on.

2.3 Data Mining Process

The details are as follows:

(1) Select data

After determining the analysis objectives and objects, we need to analyze the data in our area to understand which data can support the business and which data can be used to solve problems. The goal at this stage is to select the appropriate data set and identify the data set most related to the object, so as to lay a good foundation for the follow-up work.

(2) Data cleaning

After confirming the data set, the data needs to be preprocessed. At this stage, it may be found that there are missing data, duplicate values, outliers, etc. in the data set. The missing value generally depends on the degree of missing. If the missing part is less, it can be filled in by means of average value and so on; If there are many missing values, the attribute or this piece of data will be deleted directly.

The general data flow is shown in Fig. 5.

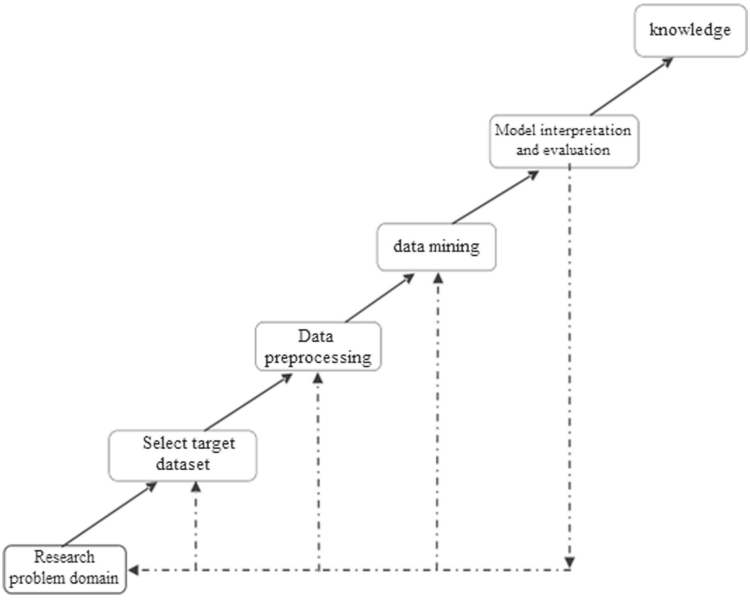

Fig. 5. General process of data mining

3 Application of Data Mining Technology in Students' Physical Fitness Analysis

3.1 Description of Physical Problems of College Students

For the above problems, data mining method can be adopted to analyze the sports test data of college students. The flow chart is shown in Fig. 6.

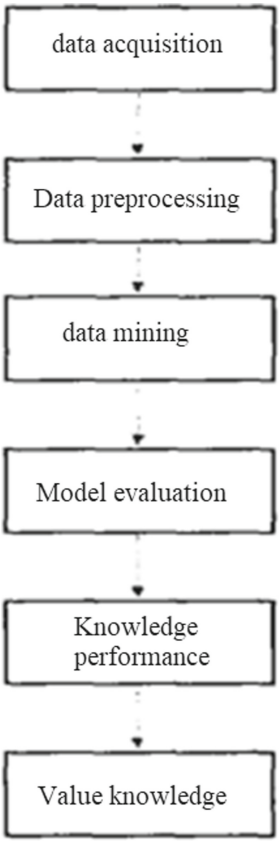

Fig. 6. Data mining flow chart of College Students' physical fitness test

In the whole process, after determining the content and purpose of the project, the relevant data will be collected and preprocessed. The data preprocessing includes four steps: data selection, data cleaning, data integration and data specification. Finally, the relevant data mining model is used to mine the data set, and the data mining results are obtained. Contact the actual content of the project to obtain the corresponding value knowledge[9].

3.2 Data Preprocessing

Data preprocessing is to select and clean up these data. Clean up some data without analytical value and select more appropriate data for experimental analysis, so as to improve the efficiency and accuracy of data mining. In addition, data integration, data standardization and data specification can reduce data redundancy and maintain data integrity, so as to improve the level of data mining.

In the physical fitness test data collection, affected by the measurement accuracy and error of the front-end data collection of the Internet of things and other objective factors,

some data measurement will be inaccurate, or even seriously deviate from the normal situation, and then produce data outliers. The occurrence of outliers will not only affect the performance query work, but also reduce the reliability of the results if the original data is mined, and even the rules cannot be mined or the wrong rules can be mined. Therefore, it is necessary to screen the original data through outliers detection [10].

Physical health intervention measures refer to physical education teachers making sports training plans according to students' grades to help improve students' physical health test results. Due to the different training venues, time and environment of different subjects, this paper analyzes the importance of individual subjects for improving the overall evaluation results, provides suggestions for rapidly improving the physical fitness test results of college students, excavates the links between various subjects, and further verifies them through Pearson correlation coefficient. After getting the relevance between the subjects, when it is limited by the teaching time and teaching venue, and it is impossible to carry out special training for a subject, it can be trained for the subjects with strong relevance and low requirements for teaching venue and time, so as to improve the performance of the above-mentioned poor subjects, which is of great help to formulate physical education teaching plans and improve the efficiency of physical education teaching and performance improvement.

For distance measurement, ultrasonic distance measurement duration of the high level is the time from the transmission to the return of the ultrasonic. According to the time difference between the transmission and reception, the distance between the ultrasonic and the measured object can be calculated.

If the ultrasonic module probe is placed at a distance of 2.5 m from the ground when measuring height, and the distance between the probe measured by the ultrasonic module and the height measurement lifting platform is h, then h = (high level time x sound speed) /2, and then the height h = (250-h) can be obtained. Similarly, in the subjects of standing long jump and sitting forward bending, ultrasonic distance measurement module can also be used to measure students' grades.

3.3 Design of College Students' Physical Fitness Analysis System

In order to meet the use needs of the above administrators, teachers and students, the system design requirements are as follows:

(1) System permission requirements

The system operates as students, teachers and administrators. Therefore, according to the actual needs of the three types of users, it is necessary to distinguish the different permissions of the three users when operating the system. When a user logs in, he / she can identify the user's role through the user name and password. After passing the background verification, he / she can enter the system. This part corresponds to the corresponding permissions of his / her role.

(2) Analysis of students' functional needs

① Students can fill in the physical health self-assessment form in the system and input personal information through the "self-assessment form" system. After

receiving students' personal information, the background will store it in the database and wait for calculation.

② Check the physical test results and evaluation results in the system. The system stores student physical test data and student evaluation data, which need to be converted into a more readable grade standard through corresponding calculation and evaluation standards. Enable students to have a clearer understanding of the results and get corresponding guidance and suggestions.

(3) Analysis of teachers' functional needs

① Check the physical test data and evaluation data of each student, and be able to add, delete, modify and check the data. ② Check the overall physical health analysis results of the whole school. By collecting the physical test information of the whole school students, we can show the physical health distribution map of the whole school students, so that teachers can better carry out physical education activities.

② Using data mining to mine students' data information and interpret the output value of the results of data mining analysis to make it more readable. Teachers can also give corresponding and targeted guidance to different students according to the mining analysis results.

(4) Analysis of administrator function requirements

① Cover all permissions of students and teachers.

② Modify the fields and evaluation system of the physical health self-assessment form, and the administrator can add, delete, modify and check the problems and options in the self-assessment form, and control the score of each option and the overall scoring standard.

③ Control the operation of data mining algorithm and the generation of results.

The technical architecture of the system is divided into three layers: presentation layer, business logic layer and data layer, as shown in Fig. 7.

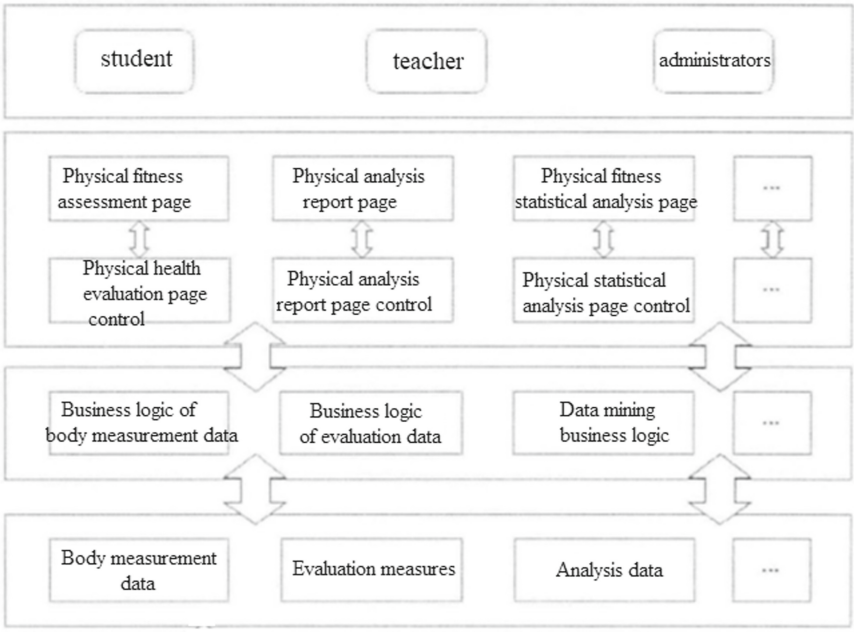

Fig. 7. Overall technical architecture of the system

4 Conclusion

The design and implementation of student physique analysis system based on data mining technology is to make a computerized system that can analyze student data to detect students' performance level, and teachers will use it to evaluate students. The main components of the project are data mining technology, artificial intelligence, database management system and Internet. The system aims to analyze the physical quality of middle school students. The system will be able to collect data from each student and then process it using data mining technology. The main purpose of this project is to create a database that can help us predict individual results based on individual performance in examinations. The database will also provide teachers, parents and students with useful information about their performance so that they can improve their skills accordingly, from the perspective of research methods, this paper only adopts the decision tree and association rule algorithm in data mining, and there are a variety of data mining algorithms that can study and analyze the project, so as to provide a more comprehensive basis and guidance for physical fitness analysis. Third, for the system function, it only involves body test data.

References

1. Zhu, L.: Design and implementation on physical fitness and health standards query system for college students. Syst. Autom. **21**, 223–234 (2018)

2. Mou, R., Yang, J.: Design and implementation of mobile app for athletes' physical fitness monitoring during training. In: 2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA) (2021)
3. Puthoff, M.: Participants' perceptions and the implementation of a physical fitness screen for aging adults. J. Geriatric Phys. Therapy, 44 (2021)
4. Fahrizqi, E.B., Agus, R.M., Yuliandra, R., et al.: The learning motivation and physical fitness of university students during the implementation of the new normal Covid-19 pandemic. JUARA J. Olahraga **6**(1), 88–100 (2021)
5. Wang, T., Park, J.: Design and implementation of intelligent sports training system for college students' mental health education. Front. Psychol. **12**, 634978 (2021)
6. Holman, S.K., Folz, H.N., Ford, B., et al.: Design and implementation of a pilot student wellness program at a school of pharmacy (2021)
7. Allen, W.E., Hosbein, K.N., Kennedy, A.M., et al.: Design and Implementation of an Organic to Analytical CURE Sequence. J. Chem. Educ. **98**, 2199–2208 (2021)
8. Efendi, R., Lesmana, L.S., Putra, F., et al.: Design and implementation of Computer Based Test (CBT) in vocational education. J. Phys. Conf. Ser. **1764**(1), 012068 (2021). (12pp)
9. Ndubuisi, A.A., Chimezie, O.E.: Design and implementation of students' projects allocation system. Int. J. Sci. Res. Publ. (IJSRP) **11**(6), 164–174 (2021)
10. AlAqbi, A., Al-Taie, R., Ibrahim, S.K.: Design and implementation of online examination system based on MSVS and SQL for university students in Iraq. Webology **18**(1), 416–430 (2021)