Physical Fitness Detection System for College Students Based on Correlation Rule Data Mining

Jian Li[1,a], Zhou Zhou[1,b,*]

[1]College of Physical Education and Health, Guangxi Normal University, Guilin, Guangxi, 541000, China
[a]281792143@qq.com, [b]3842919@qq.com
*Corresponding author

Abstract: In the past two years, the social concern about the physical health of college students has become higher and higher, and the status of contemporary college students in terms of physical skills or average physical fitness is not optimistic. Multiple data on college students' physique in the multimedia modeling of Health Cloud Biometrics and Data Management System also confirm that more and more college students' physical fitness is declining. Therefore, more and more colleges and universities began to study how to use the health cloud system of biometric authentication technology and data management system to do students ' personal physical quality data collection and multimedia modeling, and then establish a complete set of college students' physical fitness detection system based on correlation rule data mining and other technologies. This paper first analyzed the biometric authentication technology and data management system in health cloud system, and then the feasibility of the application of the technology and the correlation rule data mining algorithm in the physical fitness detection system of college students was demonstrated. Finally, relevant experiments were conducted to study the improvement of the physical exercise effect of college students before and after using the system, and it was concluded that the physical exercise efficiency of students who used the college students' physical health detection system increased by about 26%. The biometric authentication technology and data management system in the health cloud system complement the information collection and multimedia modeling functions in the physical exercise health detection system for college students.

Keywords: Student Physical Fitness Test; Data Mining; Multimedia Simulation Modelling; Machine Learning

1. Introduction

In the new era, college students should not only focus on the learning of this professional knowledge, but also pay attention to the exercise of their own physical ability. In order to help college students increase the number of physical exercises and help students to exercise more actively, this paper improved the physical health detection system for college students, so as to promote students to actively improve their physical functions and comprehensive health level [1].

College students' physical fitness testing systems can help students gain a deeper understanding of their own health, and can help students choose physical exercise courses or exercise programs [2-3]. Xue, Baohong collected and analyzed the health data of college students from the perspective of cloud computing technology, and designed a scheduler for data mining model. The scheduler gave relevant parameters that can improve data throughput in the process of data mining [4]. Moon, Jeong Eun investigated the factors that affect the physical health of college students and analyzed them through analysis of variance and stepwise multiple regression analysis, which showed that the physical health of college students was positively correlated with health awareness, breakfast habits, and subjective feelings about health conditions [5]. Lu, Xu found that improving physical fitness helps people build up their overall fitness, but the traditional physical fitness test operation is more complex and less efficient. Therefore, he designed a mobile terminal-based health and physical fitness test method for college students, and verified the effectiveness of the method through image acquisition and analysis of the user's physical fitness level [6]. In order to determine the effect of physical activity on cardiometabolic risk factors in college students, Buresh analyzed the height, weight, body composition, lipids, glucose tolerance and blood pressure of several college students. He calculated and analyzed these data for

linear regression, which showed that college students with low levels of health-related physical fitness generally had cardiometabolic risks [7]. In order to compare the gap in the quality of health and life of undergraduates in the relevant sports involved in self-assessment, Snedden used cross-sectional measurements to compare the differences in physical and mental health in the data, which showed that higher levels of exercise enabled undergraduates to have a higher quality of life [8]. Zhu, Qian studied the effect of aerobic exercise on the physical fitness of obese college students, and examined the changes of college students' endurance, body fat ratio, flexibility and body weight on the indicators before and after aerobic exercise, thus revealing the effect of aerobic exercise on the physical health of obese college students [9]. Dallo conducted research on the factors that affect college students' self-assessment of physical health, and randomly sampled multiple college students to examine the relationship between health self-examination, health and academics. Data from the tests suggested that colleges and universities need to take certain intervention measures to improve the physical quality of students [10]. Although the above researches had studied the college students ' physical quality and health testing, they have not proposed a systematic physical health detection method.

The technology related to machine learning algorithms continues to develop and expand, resulting in data mining algorithm. The birth of data mining not only brings considerable convenience to the computer industry, but also brings new ideas for medical modernization in the medical field. Raihan proposed a prediction of the risk of heart disease based on intelligent mobile terminals and data mining technology. Through data analysis of multiple clinical patients, some unique connections between heart disease patients are obtained, and users can view their own risk levels by entering their own data in the smart mobile terminal [11]. Data-based mining techniques can obtain the characteristics of the information needed from a very large amount of data. Al-Dallal used data mining techniques to develop an application to help doctors predict the probability of non-communicable diseases such as diabetes and heart disease, and to apply these data mining models to the collection and retrieval of medical records, which greatly improved hospital productivity [12]. Martinez-Martin studied the feasibility of using smartwatches for the diagnosis of personal health data and used smartwatches to collect users' personal data and health and risk assessment data. The work was largely based on data mining algorithms, and discussed the social feasibility of this approach [13]. Serrano studied users' motivation to use weight loss apps and carried out related activities. Descriptive analysis methods were used to examine the characteristics of users and process these sample data using data mining algorithms, which showed that engaging users in health applications had a positive effect on helping users focus on healthy behavior change [14]. Saeed studied the impact of data mining techniques on processing data related to healthcare products. The findings illustrate the efficient use of data mining techniques in healthcare to help physicians prescribe the most appropriate treatment for patients, while also helping to show patients the correlation between resilience and some activities [15]. Chan used data mining techniques to understand the impact of a patient's health status before traumatic brain injury on the patient's functional gain during inpatient rehabilitation. The data suggested that patients with good health prior to hospitalization also have a good gain during hospitalization [16]. Charles-Nelson used data mining techniques to analyze data from the Health Administration Information System to study the trajectory of patient care after bariatric surgery. The data suggested that about one-third of patients choose to re-enter the hospital in the first year after surgery, and the data from this experiment were useful in helping to reduce the probability of complications after bariatric surgery. There are still few current solutions for using data mining algorithms in the medical industry, so there is still no systematic theory for people to refer to.

Health detection system using association rule data mining algorithm can help students understand their health level timely, so as to choose the exercise mode that suits them. In this way, students can develop good physical exercise habits, thereby improving the average health level of contemporary college students.

2. Application Evaluation of Correlation Rule Data Mining Algorithm in Physical Health Detection System

(1) Application analysis of biometric authentication technology and data management system in health cloud system

Health Cloud Bio-Authentication provides services such as hospital file management and local resident health file management to hospitals and medical institutions near cloud computing bases through software-as-a-service methods. The full use of computing resources by the health cloud can analyze the medical resources required by the medical institutions it covers, which further reduces the

initial capital expenditure of these medical institutions. The multimedia modeling in health cloud biometrics and data management systems is applied to student health detection systems, which also reduces the upfront technology and cost investment of universities. This can greatly facilitate the collection, analysis and storage of students' personal health records, and facilitate the follow-up work. The biometric authentication technology, data management system, and multimedia modeling process in the health cloud system are shown in Figure 1.
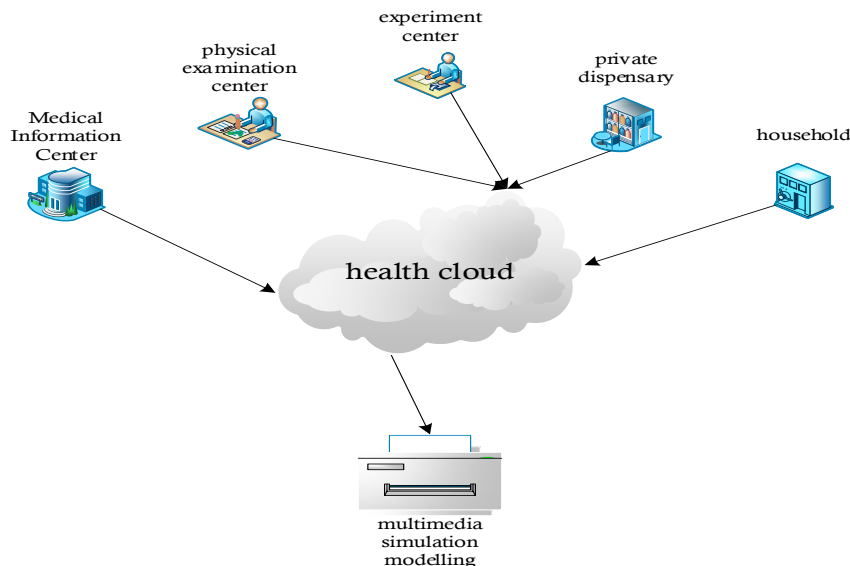

Figure 1: The workflow diagram of biometric authentication technology, data management system and multimedia modeling technology in health cloud system

(2) Data mining analysis

Data mining refers to the process of finding data from a specific database that contains potential value or is needed by users. There are three main steps: data preparation, search for laws between data, and display of laws between data, which can be seen in Figure 2.
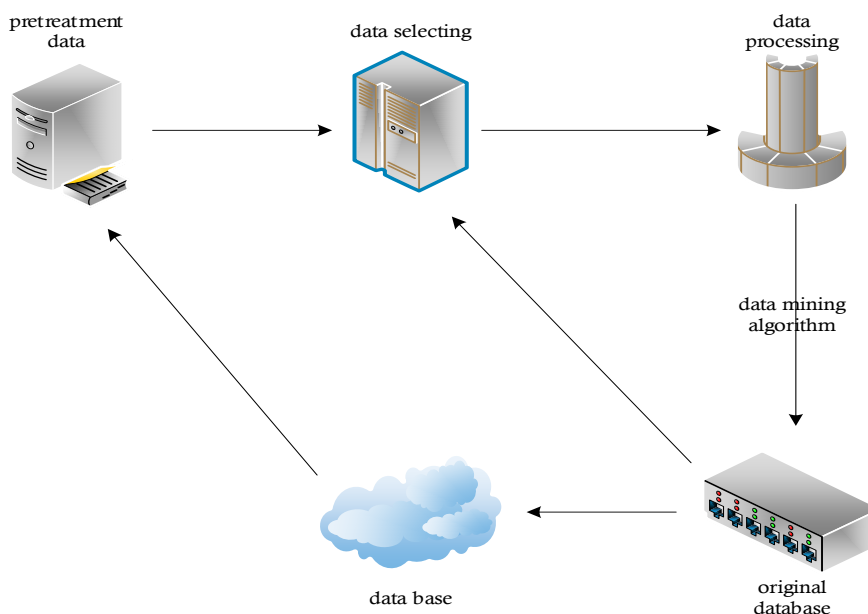

Figure 2: Data mining flow chart

Data preparation is primarily about selecting data from a data source and integrating it into a collection of data for data mining. The search for laws between data is mainly to use the methods in data mining to find the data in the data set integrated in the previous step. The display of laws between data is to show the rules as much as possible in a way that the user can understand.

(3) Association rule analysis

Association rule is the laws between data displayed from data mining to find the correlation between these data, which reflects the correlation and dependency of a transaction with other things. If there is some kind of connection between two or more things, then correlation rule data mining can predict the occurrence of other things related to this matter from the development of one of them. For example, most people in the mall like to buy pillowcases and duvet covers together, so the mall would put the two items together, which can promote sales of both.

3. Application Evaluation of Physical Exercise and Health Detection System

College students' physical health has become a prominent topic of societal concern in recent years. According to the survey, most contemporary college students are not guaranteed to eat on time, and more than half of the students often sleep after twelve o'clock. Therefore, many students have poor sleep quality and dizziness after waking up, which also leads to the inability to ensure the learning efficiency of students in the morning. Contradicting this status quo is that college students are not ignorant of health-related knowledge, and most college students know the positive effects of diet, good work and rest and proper exercise on their own physical health. There are two causes for most college students' poor health. The first is that students think that diseases are far away from them, and health concepts are generally lagging behind. The second is that many students are not firm in willpower, and the various management of colleges and universities is relatively loose and does not require students to keep moving, so many students increasingly dislike physical exercise.

Scientific and reasonable physical exercise can not only enhance and maintain the health of the human body, but also make people more and more like exercise, which is a positive feedback incentive mechanism. Physical exercise is an activity that requires people to adhere to for a long time. It tests people's willpower, and in the process of physical exercise, it constantly forges people's willpower, which can make people have strong willpower. On the other hand, physical exercise can help college students get out of the dormitory and know more students with common hobbies, which is also a good help for college students to expand their interpersonal communication range.

At present, although there are also physical health examination programs for college students, the current physical fitness test items are only a rough test of physical form, physical function and physical fitness. Although these programs can test the physical condition of students to a certain extent, the tests are still not detailed enough. The health test of students adds a number of dynamic function test indicators, and use the biometric authentication technology, data management system and multimedia modeling in the health cloud system to model the students ' various detection indicators in a timely manner. An association rule data mining algorithm is used to find correlations between students' metrics and physical activity. The current physical fitness test program for college students is shown in Figure 3.
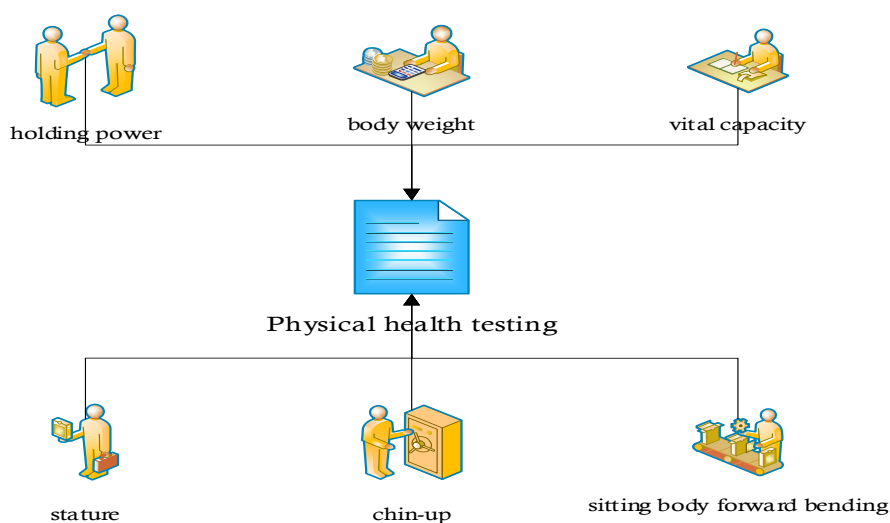

Figure 3: Current college students ' physical health testing project

4. Association Rules Data Mining Algorithm

The health detection system proposed in this paper used data mining algorithm and association rules technology, and used the data mining technology of association rules to find the correlation between students' physical health and various physical exercises in the database. This allows a student's defects to be determined based on the student's physical health test indicators.

This paper first used Apriori's algorithm, which is based on a recursive formula for the two-stage frequent set idea. If the Apriori algorithm is used, it is necessary to calculate the degree of support, confidence and improvement of the research object.

For a simple example, in the shopping mall survey, the degree of support represents the proportion of the number of orders that consumers buy a and b at the same time in the total number of orders, and generally uses $P(a)$ to represent the proportion of orders to buy a, so the Formula (1) of support $Support(a, b)$ can be calculated:

$$Support(a, b) = \frac{P(a,b)}{p(all)} \tag{1}$$

The confidence level $Confidence(a \rightarrow b)$ represents the proportion of consumers who buy a commodity at the same time in the order to buy b product, that is, the proportion of consumers who buy a product and b goods at the same time in only buying a product, which is calculated as shown in Formula (2):

$$Confidence(a \rightarrow b) = P\left(b/a\right) = \frac{P(a,b)}{P(a)} \tag{2}$$

According to Formula (2), the proportion of consumers who buy goods a and b in only buying goods b can be obtained, and its calculation is shown in Formula (3):

$$Confidence(b \rightarrow a) = P\left(a/b\right) = \frac{P(a,b)}{P(b)} \tag{3}$$

The final is the calculation of the degree of improvement. The degree of improvement mainly reflects the correlation of a to b in the data mining association rules. If the degree of improvement is greater than 1, it means that the positive correlation is higher; if it is less than 1, it means that the negative correlation is higher; if it is 1, it indicates no correlation. The formula for calculating the degree of improvement is shown in Formula (4):

$$Lift(a \rightarrow b) = \frac{P(a,b)}{P(a)P(b)} = \frac{P(b/a)}{P(b)} \tag{4}$$

By using the PCY algorithm, the frequency of scanning the database during the calculation of the Apriori algorithm and the memory occupied when selecting candidate sets van be reduced. The PCY algorithm can be said to be an algorithm that further optimizes the Apriori algorithm. The core calculation formula of the PCY algorithm is shown in Formula (5):

$$h(a, b) = \left((order\ of\ a) * 10 + (order\ of\ b)\right) mod 7 \tag{5}$$

Next, the maximum expectation algorithm is used to calculate and classify the data in the database. The sample data is determined first, and the category of each sample is z. It is assumed that there can be m kinds of values for the samples, the task of the maximum expectation algorithm is to calculate the similarity of the parameters in these sample data, which is shown in Formula (6):

$$lnL(\theta) = \ln\left(p(x_1, \theta) \cdot p(x_2, \theta) \cdots p(x_n, \theta)\right) = \sum_{i=1}^{n} ln \sum_{j=1}^{m} p(x_i, z^{(j)}, \theta) \tag{6}$$

The distribution function of x_i pair z is represented by $Q_i(z)$, then $Q_i(z)$ need to satisfie the Formula (7):

$$\sum_{j=1}^{m} Q_i\left(z^{(j)}\right) = 1 \tag{7}$$

$Q_i(z)$ is added to the core calculation formula of the maximum expectation algorithm, and the Formula (6) is simplified based on the Formula (7) to obtain the Formula (8):

$$lnL(\theta) = \sum_{i=1}^{n} ln \sum_{j=1}^{m} Q_i\left(z^{(j)}\right) \frac{p(x_i, z^{(j)}, \theta)}{Q_i(z^{(j)})} \tag{8}$$

Finally, the solution for the lower bound $J(z, Q)$ of the maximum similarity function is shown in Formula (9):

$$J(z, Q) = \sum_{i=1}^{n} \sum_{j=1}^{m} Q_i(z^{(j)}) ln \frac{p(x_i, z^{(j)}, \theta)}{Q_i(z^{(j)})} \qquad (9)$$

After the calculation of the above algorithm formula, the correlation between students' physical health indicators and various physical exercises can be obtained from the database of college students' physical health detection data, so as to facilitate physical education teachers to formulate scientific and effective physical exercise plans for students to improve students' health level.

5. An Experimental Study on the Improvement of Students ' Physical Quality by Student Health Detection System

This paper proposed a scheme to construct a physical health detection system for college students' physical exercise by using biometric authentication technology, data management system and association rule data mining related algorithms in health cloud system. Health Cloud Biometric Authentication technology ensures that the modeling can be tied to each student, and the data management system is responsible for managing the students ' physique data. Finally, association rule data mining technology is used to analyze the correlation between physical health indexes of middle school students and physical exercise items in health database. This section studied the comparison between the health distress and physical health of college students in recent years, and finally simulated the degree of improvement of college students' physical health detection system for college students' health level.

The first is an analysis of the major health problems of the university students in recent years, and the results are shown in Figure 4.
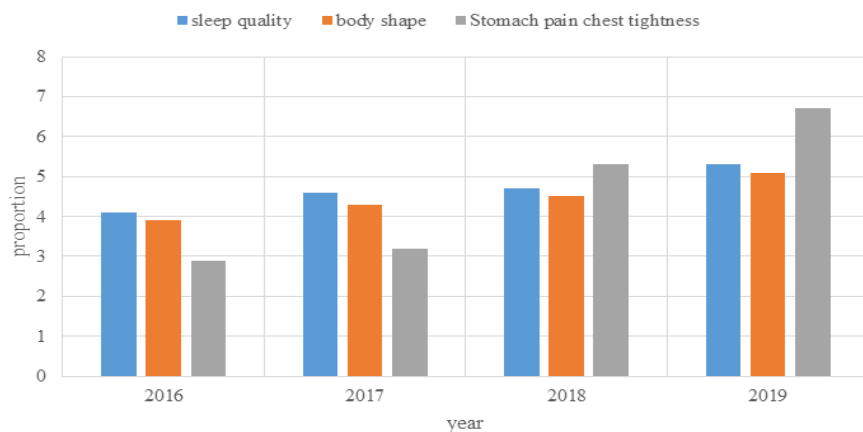

Figure 4: The main health problems of college students

After analyzing the common health problems of college students, it can be seen that the problem of sleep quality in college students has been increasing in recent years. The main reason is that because of the continuous development of smart phones and other entertainment technologies, more and more college students sleep late, which leads to the quality of sleep of college students is also decreasing. In addition, most students in late sleep have the habit of eating snacks, and it often occurs that they sleep until noon due to lack of sleep time at night. Over time, students' physical problems have also increased. Finally, the students' unplanned schedules and eating habits lead to long-term sub-health conditions, so students often suffer from stomach pain and chest tightness caused by lack of sleep and eating disorders. Therefore, college students have a healthy diet and work and rest plan which is very helpful to improve the overall health level of students.

The next is the analysis of students' comprehensive physical quality, and the results are shown in Figure 5.

Figure 5: College students ' physical health level diagram

Figure 5 shows the proportion of students ' physical fitness achievements in three years. From this test report, it was probably clear that only a very small number of students had had excellent physical fitness levels in the past three years; most students were between good and passing, and a small number of students were in a failing health level; the physical fitness level of students had basically not been improved very well. With the passage of time, there were fewer and fewer students with excellent health levels, and more and more college students were in a passing state. Although the number of students who fail was also decreasing, it also meant that the overall physical health level of college students was declining. Therefore, at this time, there is an urgent need for a set of scientific and perfect systems to supervise the physical health level of college students and help relevant staff of colleges and universities to urge students to improve their health level.

Finally, the effect of simulating the physical health detection system of college students on improving the health level of college students is shown in Figure 6.
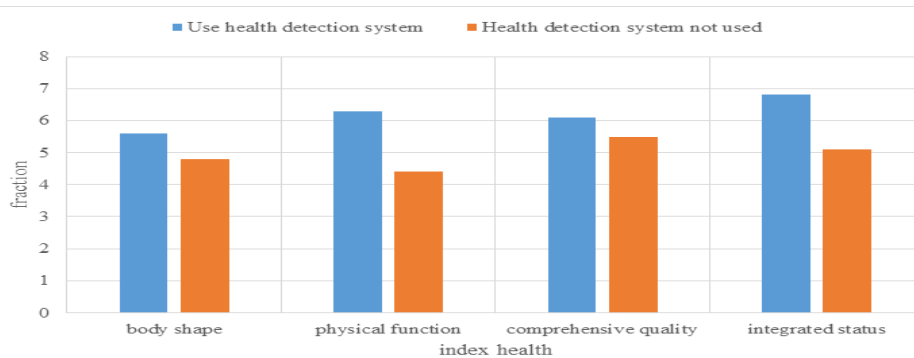

Figure 6: Comparison of the effect of college students ' physical health detection system on improving students ' health level

Figure 6 compares four health indicators after exercise between students who partially use the physical fitness testing system proposed in this paper and students who do not use the physical fitness testing system. It can be seen that the physical fitness indicators of students who used the physical fitness testing system after a period of physical exercise had increased by an average of about 26% compared with those of students who had not used the physical fitness testing system. Students who used the physical fitness test system can obtain more accurate data on body shape, physical function, comprehensive quality and comprehensive status, and formulate a personalized exercise plan based on the shortcomings in these data to systematically improve their own health level. The physical health detection system used health cloud biometric authentication to bind individual students and used a data management system to manage these student health files. After that, the student's health level was multi-modeled based on the student's health measurement data. Finally, the correlation rule data mining algorithm was used to obtain the correlation between various health indicators and physical exercise, which was convenient for comprehensively improving the health level of college students.

6. Conclusions

The current physical health detection system for college students still has many flaws, and it cannot display the physical form and physical function of college students in real time and comprehensively.

Since the reliability of the comprehensive evaluation of college students' physical quality depends on the validity and scientificity of the selected testing system to some extent, it is necessary to screen out the most suitable system indicators for their own development through layers of scientific analysis in the actual physical health test, and to emphasize the integration of theoretical research and actual testing. In this paper, a physical fitness detection system for college students based on association rule data mining and using biometric authentication technology and data management system in the health cloud system was proposed. It was not only convenient for college and university physical education educators to efficiently guide students to carry out scientific physical exercise, but also improve the physical health level of students, so as to do their best to achieve the scientific and efficient physical health test. However, the system is still some distance away from the practical application in various universities, so more researchers are needed to study and explore this aspect, so as to encourage the scientific and efficient improvement of physical health among students.

References

[1] García-Hermoso, A., Alonso-Martínez, A. M., Ramírez-Vélez, R., Pérez-Sousa, M. Á., Ramírez-Campillo, R., & Izquierdo, M. "Association of physical education with improvement of health-related physical fitness outcomes and fundamental motor skills among youths: a systematic review and meta-analysis." JAMA paediatrics 174.6 (2020): e200223.

[2] Fahrizqi, Eko Bagus. "The Learning Motivation and Physical Fitness of University Students during the Implementation of the New Normal Covid-19 Pandemic." JUARA: Jurnal Olahraga 6.1 (2021): 88-100.

[3] Sanjar, Uraimov. "The role of physical quality and abilities in the formation of student teams in futzal." International Journal of Research in Commerce, It, Engineering and Social Sciences 16.06 (2022): 10-16.

[4] Xue, Baohong, and Ting Liu. "Physical health data mining of college students based on DRF algorithm." Wireless Personal Communications 102.4 (2018): 3791-3801.

[5] Moon, Jeong Eun. "A convergence study on factors affecting physical-health status of college students." Journal of the Korea Convergence Society 9.1 (2018): 31-39.

[6] Lu, Xu. "Test method for health-related physical fitness of college students in mobile internet environment." Mathematical Biosciences and Engineering 16.4 (2019): 2189-2201.

[7] Buresh, Robert. "Associations between measures of health-related physical fitness and cardiometabolic risk factors in college students." Journal of American College Health 66.8 (2018): 754-766.

[8] Snedden, Traci R. "Sport and physical activity level impacts health-related quality of life among collegiate students." American Journal of Health Promotion 33.5 (2019): 675-682.

[9] Zhu, Qian, Sheng Feng, and Wen Liang. "Influence of Aerobics Exercises on Physical Health of Obese College Students under the Background of Physical and Medical Integration." Investigación Clínica 61.2 (2020): 837-846.

[10] Dallo, Florence. "Predictors of self-rated health among undergraduate college students in the united states." College Student Journal 52.4 (2018): 431-440.

[11] Raihan, M. "Smartphone based heart attack risk prediction system with statistical analysis and data mining approaches." Advances in Science, Technology and Engineering Systems Journal 2.3 (2017): 1815-1822.

[12] Al-Dallal, Ammar, and Amina Al-Moosa. "Prediction of Non-Communicable Diseases Using Class Comparison Data Mining." Advances in Science, Technology and Engineering Systems Journal 4.5 (2019): 193-206.

[13] Martinez-Martin, Nicole. "Data mining for health: staking out the ethical territory of digital phenotyping." NPJ digital medicine 1.1 (2018): 1-5.

[14] Serrano, Katrina J. "Characterizing user engagement with health app data: a data mining approach." Translational behavioral medicine 7.2 (2017): 277-285.

[15] Saeed, Soobia. "Impact of data mining techniques to analyze health care data." Journal of Medical Imaging and Health Informatics 8.4 (2018): 682-690.

[16] Chan, Vincy. "Data mining to understand how health status preceding traumatic brain injury affects functional outcome: a population-based sex-stratified study." Archives of physical medicine and rehabilitation 101.9 (2020): 1523-1531.