Data Analysis and Application Research in Sports Field under Big Data Environment

Peng Ding
School of Physical Education, Wuhan Business University, Wuhan, China
Email: 949808799@qq.com

Abstract. This paper expounds data mining theory and sports data analysis theory, such as data mining concept, mining method and mining flow. Taking the data analysis in basketball and track and field as an example, this paper discusses the application of data mining in the field of sports industry, mass physique monitoring and physical education data analysis at home and abroad, and analyzes the influence of big data on sports, mainly changing the methods of competitive sports training, promoting the innovation of sports research methods and the transformation of sports communication methods, and promoting the development of the world sports industry.

1. Introduction
With the rapid development of information technology and the wide application of database management system, people accumulate more and more data. There is a lot of important information behind this explosion of data that people want to analyze at a higher level to make better use of. Therefore, database management system can be used to store data, machine learning and statistical methods can be used to analyze data, and the knowledge behind a large number of data can be mined. The combination of them promotes the emergence of data mining. Data mining is to help decision-makers to make correct decisions through highly automated analysis of original data and inductive reasoning. With the development of data mining technology, data mining algorithms with good accuracy, time complexity and space complexity have been proposed, making data mining technology widely applied in many fields. Data mining technology has become a research and application hotspot in the society.

In the face of big data processing, special data mining techniques have emerged to meet the increasing amount of data in sports business data processing. This new data mining processing technology will be of great help and vital importance to sports statistics. Sports data mining processing technology for these huge data can be analyzed out of the useful data for us to identify. And it can make a clear and clear processing and display of the interrelation of various factors among various data. And it can make a clear and clear processing and display of the interrelation of various factors among various data. Data mining processing technology can deeply understand and analyze various physical health conditions, sports industry data analysis, sports competition and competition data processing, and can provide some valuable data and data analysis results in school physical education teaching data. So as to find some sports laws with great value. Then in the actual application of data mining processing technology, how to carry out the in-depth application of data mining processing technology so as to facilitate the management of the overall sports data processing and data analysis. This is one of the problems to be solved at present, and this problem is of great significance to help better use of data mining processing technology in the active application of sports big data processing.
2. Data Mining Technology

Data mining processing technology is a process of searching out the special relationship between the hidden data from the large amount of accumulated data collected by oneself instinctively. With the changes of The Times, the amount of data is increasing day by day. There are still many undiscovered data hidden in the accumulated data, but they are very useful data. As long as the information is utilized, it can be maximized and maximized benefits can be generated. Data mining processing technology is not a single discipline application, but a product of the integration of multiple disciplines, which involves many aspects and disciplines.

Data mining processing technology is different from traditional statistical computing. The biggest difference is that the final inference of statistics is only a final inference of hypothetical significance, that is to say, after a certain assumption and on the basis of a certain amount of data to carry out the reverse inference of the result; then the data mining processing technology is the inference of final data based on a large amount of data, which is an extraction mode and hypothesis from a large amount of accumulated data. Then in the actual application of data mining processing technology, how to carry out the in-depth application of data mining processing technology so as to facilitate the management of the overall sports data processing and data analysis. This is one of the problems to be solved at present, and this problem is of great significance to help better use of data mining processing technology in the active application of sports big data processing. This is also one of the reasons why data mining technology can be applied in sports industry or physical education industry. Only by making good use of data mining techniques can we better obtain the results of data analysis.

3. Application of Data Warehouse and Data Mining Technology in Data Analysis of Sports Field

3.1. Application of Data Mining in Data Analysis of Sports Events

There will be huge data in the technical and tactical analysis of various competitions in competitive sports. The traditional competition analysis generally adopts artificial way to observe athletes' competitions, and then carries out statistics on relevant data. The results will be obtained after the practical experience research and analysis of sports researchers. Therefore, there is a serious lack of science, but also some limitations. With the continuous development of computer technology, manual input is shrinking, and the technical and tactical analysis requirements are more refined, which further promote the development of technical and tactical data acquisition and analysis technology. Sports researchers use technical and tactical data collection and analysis system to collect a large number of technical and tactical data, processing these data, using a certain data model and method to further mine and analyze technical and tactical data, and finally evaluate and interpret the model, and apply the data analysis conclusions to specific training. Through the analysis of the athlete's technical and tactical characteristics and related sports data, effective analysis results can be obtained on the respective strengths and weaknesses of the two sides in the competitive sports competition, and the technical and tactical strategies for restraining others. The effective analysis results of technical and tactical strategies can provide reference for coaches to make training plans and implement competition tactics. Therefore, technical and tactical analysis plays an important role in the balance of victory.

The typical applications of data mining technology in sports analysis are the Advanced Scout developed by IBM used by NBA teams, inside-Edge widely used by various teams in the American Professional Baseball League and PASW data analysis tool widely used in European Football Leagues. This study takes basketball and track and field as examples to illustrate the application of data mining in sports event analysis.

3.1.1. Basketball

The most widely used data analysis tool in NBA League is the Advanced Scout tool. For example, in a game between magic and the heat in 2010, the coach used the advanced scout system to analyze that when the magic played Hardaway and Shaw in the starting lineup at the same time, the team scored -17 points every 48 minutes, while the magic team scored +14 points when Hardaway and substitute Armstrong were present at the same time. The Magic decided to let Armstrong enter the starting lineup
and finally won the game smoothly.

3.1.2. Athletics
The success of data mining in track and field is the story of Wilkins, an American athlete, who won the Gold medal in Montreal Olympic Games. After analyzing Wilkins' throwing techniques and throwing movements with data mining tools, Ariel, an American sports scientist, corrected the defect that Wilkins didn't use the braking leg strength in throwing. He helped Wilkins improve his throwing skills, which raised his score by more than 3 m and helped Wilkins win the Montreal Olympic champion and set a new world record.

3.2. Research on the Development of Sports Industry
For the sports industry, big data promotes the generation of more industry values, promotes the coordinated development of the sports industry, and speeds up the scientific and technological development of sports industry. The focus of big data is to excavate the potential value of data, integrate and analyze data information, so as to make it a management resource that enterprises can use. According to data information, changes in market demand and development trends can be grasped in time, and consumers in the sports market can be finely analyzed to provide consumers with attractive sports products. Private customized sports products and services will also become the development trend of the industry. The development prospects of the sports fitness and medical industry are very broad. With the help of Internet of Things sensor social software or mobile phone applications, users' information and data can be obtained. Currently mobile phone users installed Weibo, WeChat, Alipay, QQ, Facebook and other customer applications can diagnose the body data and exercise health of users, and provide sports advice and programs, which can attract more consumers to participate in the fitness and medical industry.

At present, China has formed a number of sports industry bases. Under the background of big data, mutual benefit and sharing of resources can achieve a multi-party win-win situation. Universities and scientific research institutions provide talents for sports enterprises and realize the combination of supply and demand. Enterprises can save investment in scientific research and transform the theoretical achievements of universities into the productivity of sports enterprises. Data processing institutions can obtain more data resources in the accumulation area and expand the order volume. Intermediary institutions become the bridge linking the development of the sports industry. The coordinated development of sports industry can achieve mutual benefit and win-win situation for all parties. There are many successful cases of big data promoting the development of sports industry. The British Reuters Group, which is engaged in information and news, financial services, management and investment and other businesses, provides service support to well-known media and news agencies around the world through special networks, and provides support for the operation of media, newspapers and enterprises such as CNN, New York Daily and even Xinhua News Agency. Open customized information, information and other services for major customers to meet the individual needs of customers. Establish a database of event personal information and use statistical data to launch an online ticketing service for sports fans. In addition, Tencent Sports, in conjunction with IBM, deeply explores the hot spots of World Cup fans from various dimensions of events, fans and culture, and creates wonderful programs to win fans' resonance and bring them different experiences.

3.3. The Field of Public Health Monitoring
Since 2000, China has carried out the national physical fitness monitoring once every five years. This monitoring project is a very large and systematic project, among which the adolescent physical fitness data monitoring was carried out earlier than in 1985. With the comprehensive development of physical fitness monitoring, this field has accumulated a large amount of data while collecting a large amount of monitoring data. Although the traditional statistical method can obtain the corresponding statistical analysis results of data, it cannot reveal the correlation between the conclusions. The use of data mining technology and human medical knowledge can have important guiding significance for the health status, disease situation of the whole people, as well as the change of national lifestyle and the
progress of sports and fitness awareness. In the use of data mining analysis to test the physical health of students, the data mining mode can be set. Firstly, the problems can be clearly defined, the feasibility of the objectives can be analyzed, and the criteria for measuring the results can be set. The second step is to establish a data warehouse, which specifically refers to the collection and maintenance of data, which can be divided into three databases, namely, student information, physical fitness test and physical education performance. Then analyze the data, find the data field set that has a great impact on solving the problem, and decide whether to export the definition field. The next step is to preliminarily model the data, record and filter the field variables in the database according to the definition problem, and implement the transformation according to the current variables to get new variables and fields. The main purpose is to prepare the data for building the model. After the model is established, the corresponding mining algorithm is selected to process the data, and various models are also investigated, and the investigation results are used as a high-precision mining method to judge and select solutions. The meaning of evaluation and interpretation model is that after the establishment of the model, the results need to be evaluated. If there are defects, the model needs to be revised. Meanwhile, the value of the model needs to be effectively explained and finally implemented.

3.4. Application of Data Mining in Data Analysis of Physical Education Teaching

Data mining technology is widely used in the field of physical education teaching, such as teaching management, teaching evaluation, curriculum setting, teaching methods, teaching material selection and so on. For example, using the management rules in data mining, we can establish the prediction model of test results according to the data of different students' gender, age, body function and so on, and find the potential knowledge to improve teaching methods and improve the effect of physical education and training. For example, this study selects the relevant data in 2014~2019 “China Sports Statistics Yearbook” to model the research, taking “the number of world champions "as the research object, taking” the outstanding sports team funds "," the sports capital construction expenditure "," the national level outstanding sports team member "," the coach total number "and" the rank referee "as the related factors. By establishing the data mining association model, it is concluded that human input is the basis of excellent performance output, and human input plays a decisive role in improving performance, while financial input is the key factor of excellent performance output. Therefore, in the two categories of indicators, the indicators related to human input are more important than those related to financial input.

4. Conclusion

Compared with the application in other fields, data warehouse and data mining technology have been used in the sports field a little later, but the results have been considerable. With entering the era of big data, sports researchers in competitive training, public physical health and sports industry, in order to ensure that a large number of sports monitoring data, technical and tactical data and national physical fitness data can be fully mined and analyzed in the field of sports, this study needs to be carried out in more depth. Such as the establishment of sports data platform, relying on the data platform research and mining more data resources in the field of sports, provide the corresponding data sharing platform for sports researchers. To effectively promote digital tools, and finally realize the popularization of digital tools, and develop more data mining software on the basis of data acquisition, so as to ensure that any adult can operate the interface without mastering professional knowledge and skills. The data analysis technology and the data of related industries are combined to further develop the scientific application of data warehouse and data mining technology in the field of sports, and to provide reference value for the development of information technology in the field of sports.

5. Acknowledgement

2020 Hubei Leisure Sports Development Research Center Open-ended Fund Project Established Item Project number2020Y022.

The research center of sports and health, Wuhan Business University program number: 2019TD002.
6. References

[1] Xin Lijuan. Application of big data in the development of sports industry [J]. China New Communications, 2018(3): 108.

[2] Yang Fanjing, Chen Zhihui. Data analysis and application research in sports field in the era of big data [J]. Journal of Xi’an University of Arts and Sciences (Natural Science Edition), 2017 (5): 113-117.

[3] Xie Xiangyang. Research and application of data mining in sports data analysis [J]. Contemporary sports science and technology, 2013, 3 (23): 9-10.

[4] Huang Qian, Shi Yong. Application of data mining in sports training guidance [J]. Journal of Guangzhou Institute of physical education, 2009, 6 (29): 106-110.

[5] Yu Daifeng, Zhong Yaping, Yu Yaguang, etc. Application of data mining technology in data analysis of human muscle strength - A case study of human muscle strength test data [J]. Sports science, 2010, 30 (2): 70-74.

[6] Liu Ke. Opportunities and challenges for China’s sports development in the era of big data [J]. Contemporary sports science and technology, 2017 (16): 250-252.

[7] Zhang Chao. Research on the implementation status and Countermeasures of national physical fitness monitoring under the background of big data [D]. Beijing: Beijing Sport University, 2017.

[8] Wang Qi, Yan Xiaoyan. Opportunities and challenges of China’s sports development in the era of big data [J]. Sports and Science, 2016 (1): 75-80.

[9] Xue Hongqiang. A brief analysis on the application of data mining in the physical health test of students [J]. Journal of Xi’an University of Aeronautics, 2014(6) : 87-90.

[10] Qiu Naimin. New path of sports training scientific research in the era of big data [J]. Journal of Shandong Institute of physical education, 2015, 31 (4): 83-89.