

# Data Mining-Based Design and Implementation of College Physical Education Performance Management and Analysis System

<https://doi.org/10.3991/ijet.v14i06.10159>

Yimeng Fan

Hebei Agricultural University, Baoding, China

Yu Liu

Baoding University, Baoding, China

Haosong Chen, Jianlong Ma<sup>✉</sup>

Hebei Agricultural University, Baoding, China

mayongqing05@126.com

**Abstract**—The purpose of this paper was to effectively apply data mining technology to scientifically analyze the students' physical education (PE) performance so as to serve the physical teaching. The methodology adopted in this paper was to apply ASP.NET 3-layer architecture and design and implement college PE performance management and analysis system under the premise of fully analyzing the system requirements based on Visual Studio2008 software development platform and using SQL Server 2005 database platform. Based on data mining technology, students' PE performances were analyzed, and decision tree algorithm was used to make valuable judgments on student performance. The results indicated that applying computer technology to the management and analysis of college PE performance can effectively reduce the teaching and managing workload of PE teachers so that the teachers concentrate more on the quality of physical education.

**Keywords**—Physical education (PE), data mining; system requirements, decision tree algorithm

## 1 Introduction

Physical quality and PE performance of college students have become the focus of social attention. In addition to the mastery of college culture courses, society and enterprises pay more and more attention to the overall quality of college students such as physical quality; therefore the education department also requires colleges and universities to strengthen the monomial achievement test and physical fitness test for college students. PE performance tests are usually measured by standards such as length and time, while PE performances are converted into specific scores according to education sector standards for recording and analysis, which takes a lot of man-

hours. The workload of PE performance management occupied too much time of PE teachers, which seriously affected PE teachers' teaching quality in PE courses [1].

With the development of computer network technology and the further improvement of computer hardware in colleges and universities, most ordinary colleges and universities have already met the hardware requirements for setting up PE performance management and analysis system based on data mining technology [2]. From the perspective of data query, designing PE performance management and analysis system for colleges and universities can provide quick and convenient query means for PE teachers, students and administrators. For data registration, the efficiency of data statistics, analysis and processing are greatly improved. In addition, the system can effectively analyze PE performances through decision tree algorithm model and make effective judgments on the key links that affect PE teaching [3].

This paper firstly introduced the related computer technology ASP.NET technology and SQL database, and carried out a science popularization of data mining technology and the decision tree algorithm. Then the system requirements, the overall system architecture, system function modules, and system databases were analyzed and designed. Based on the ASP.NET 3-layer program architecture, the design of the PE performance management and analysis system was completed on Visual Studio 2008 software platform, and the operation interface of the PE performance management was presented [4]. At last, the data mining of student PE performance table was carried out by using the decision tree algorithm, and valuable judgments were obtained from the measured PE performances. On one hand, the study of this paper could facilitate the PE teachers to carry out PE teaching management, on the other hand, data mining and data analysis can provide effective guidance to students and teachers and improve scientificity and pertinence of college students' physical exercise [5].

## **2 Introduction to Related Computer Technology and Basic Theories**

### **2.1 ASP.NET**

ASP.NET is called .NET for short, and it is a general term for a series of products offered by Microsoft. The .NET strategy, .NET Framework, .NET Enterprise Server, and .NET development tools make up .NET [6]. The .NET strategy refers to the idea of connecting all the devices on the Internet and using all the software as a web service. The .NET Framework is an environment that provides programmers with specific services and technologies. .NET Enterprise Server refers to server-end products used by .NET Framework applications such as SQL Server 2000.

The presentation layer is responsible for communicating with the owner. The business logical layer is responsible for forwarding the browser request to the data layer and returning the processing result. The data layer is responsible for providing data services to the business logic layer [7].

Figure 1 shows ASP.NET 3-layer program architecture:

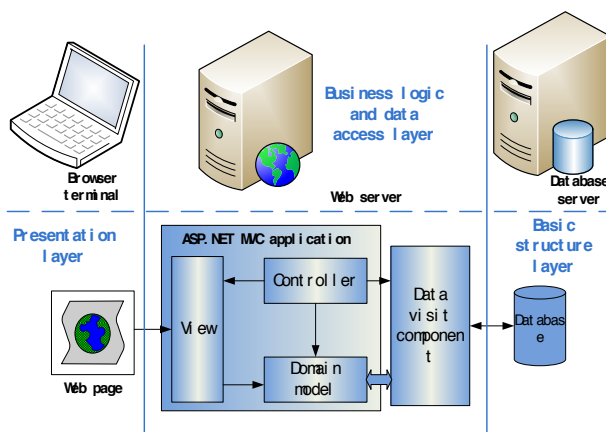


Fig. 1. The system 3-layer structure of .NET

## 2.2 SQL server

SQL Server2005 is a typical relational database with functions of creation, development, design and management. It supports operating systems of all Windows versions, and is widely applied in e-commerce, enterprise data warehouse, management information systems and the like. As a database in the context of Internet, it offers a complete data management and data analysis solution for users' Web applications [8]. SQL Server2005 has the following characteristics:

**Internet integration:** Web applications are fully developable with the XML Database Engine and provide powerful search and query functions.

**Scalability and availability:** The workload of different software systems can be satisfied by configuring multiple SQL databases respectively, while the related SQL database can query and optimize preferentially across multiple servers [9].

**Data warehouse function:** The typical characteristic of data warehouse makes SQL have the advantage of a large capacity, data analyzing and summarizing tools and online processing tools, including data warehouse, OLAP and other tools.

**Enterprise-level database functions:** With stability and security, SQL adapts to the operating environment with the high complexity, and the stability of the operation is also favored by users.

## 2.3 Data mining technology

Data mining technology analyzes and predicts the saved data or database with different algorithms, and excavates valuable information. It has the characteristics of precognitiveness, validity and practicality [10]. The function of excavating potential value from a large amount of data of data mining mainly includes automatic data analysis and trend forecasting, correlation analysis, and clustering.

Compared with traditional data information processing methods, data mining technology has the following characteristics:

- Large-scale databases are objects to be processed [11].
- Decision-making users can query in real time and randomly; data mining technology provides them with interested things.
- Data mining techniques can extract useful rules and make predictions.
- Data mining technology can quickly respond to data changes and provide decision support.
- Data mining discover potential rules and manage and maintain rules as well.
- The found rules do not have to be applied to all the data. When a certain threshold is reached, the rule can be considered as existing.
- Data mining extracts rules based on database or shot materials in data warehouse, and then uses them to guide the present and future actions.

## 2.4 Decision tree ID3 algorithm

This paper selected ID3 algorithm proposed by QuinLan as the decision tree classification algorithm. A set of non-class attributes  $C_1, C_2, \dots, C_n$ , a class attribute  $C$ , and a training set  $S$  were given, a decision tree could be constructed using the ID3 algorithm [12]. The algorithm was described as follows:

```
// return a decision tree
Function ID3 (R: a combination of non-class attributes, C: class attributes, S: a training set)
Begin
  If S is null, a single node with the value Failure is returned.
  If S consists of the same category attribute records, return a single node with this value;
  If R null returns a single node whose value is class attribute value with the highest frequency found in the record; assigns the attribute with the largest gain (D, S) value among the attributes in R to D;
  Assign the value of attribute D to  $\{d_j | j = 1, 2, 3, \dots, m\}$ ;
  Assign a subset of S consisting of records corresponding to D and having a value of  $d_j$  to  $\{s_j | j = 1, 2, 3, \dots, m\}$ ; return a tree, mark it as D, mark branches as  $d_1, d_2, d_3, \dots, d_m$  [13];
  Then construct the tree separately:
  ID3(R- $\{D\}$ , C, S1), ID3(R- $\{D\}$ , C, S2), ..., ID3(R- $\{D\}$ , C, Sm);
End ID3;
```

The ID3 algorithm builds a decision tree based on a set of data sets or data objects with known class attributes and relies on the decision trees to classify unknown type of data. Information gain is the main evaluation standard attribute of the algorithm. This attribute was utilized to determine a decision tree node. By testing all values of the node, each branch of the node is obtained. In this way, the original data set is divided into a number of sub-data sets by branches. If a node contains data rows of the same category, then the node is the leaf node of the decision tree and is marked as

the corresponding category. The construction of this decision tree is repeated until all the nodes do not need to continue branching.

### 3 Analysis and System Design of college PE Performance Management and Analysis System

#### 3.1 System function requirements

Figure 2 shows the use case diagram of the three types of users of the system:

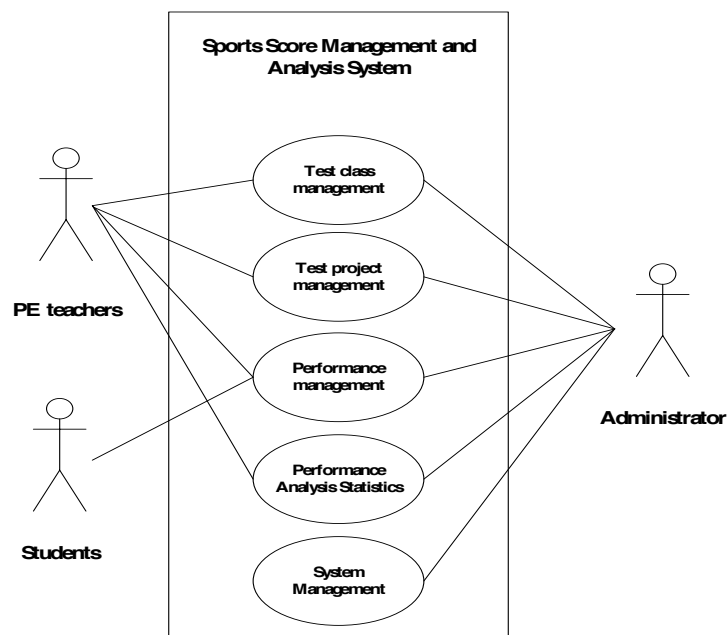


Fig. 2. The system overall use case diagram

The purpose of system design was to realize orderly management, classification and filing of PE performance on one hand, reduce the teaching pressure of PE teachers and shorten the time spent on PE performance management and improve the efficiency and accuracy of registration and recording of students' PE performances on the other hand, in this way, the teachers can have more energy to improve the teaching quality of physical education.

The system had to provide administrators, teachers and students with the rights to register and log in. The PE performance evaluation system should store each type of test standard and score conversion weight under the premise of 100 system. The system should have the function of inputting students' PE performances, and can modify, delete, query, export etc. the performances; the system should have the functions of maintaining, back-uping and restoring basic data information.

As can be seen from the figure, the project where teachers participate in include PE test type, PE test project, PE performance statistics, PE performance management, while students can only participate in PE performance management; in contrast, the system managers participate in all types of project managements [14].

### 3.2 System design

Corresponding to the 3-layer architecture of ASP.NET technology, the overall technical architecture of the system is shown in Figure 3:

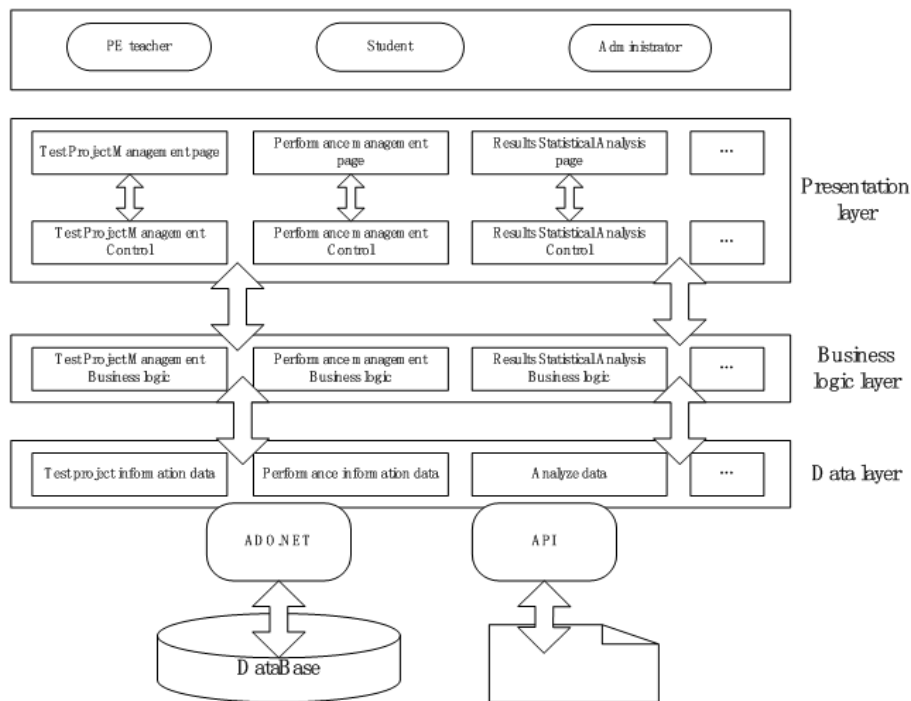


Fig. 3. The System overall technical architecture

The system was divided into three layers: presentation layer, business logical layer and data layer from top to bottom. The presentation layer interacted with the system user to obtain the user's relevant data requirements on PE performance. Logical judgment was performed through the business logical layer, and finally the database was accessed through the data layer, and the relevant data was obtained [15].

**Functional module design of the system:** According to the system requirements and the overall architecture, the functional modules of the system are divided into five parts:

- PE test type management module, its sub-function modules include: add, modify, delete test types, and various types of score permission settings; common test types

include basic knowledge of PE, physical fitness standards, physical fitness projects, special technical assessment and so on.

- Test project management module, its sub-function modules include: add, modify, delete test projects, and permissions settings of various project scores, common test projects include basketball, volleyball, medium length running, ping pong and so on.
- Performance management module, its sub-function modules include: enter, modify, delete, query, export PE performances, and credit system conversion.
- Performance analysis module, the sub-function modules include a performance statistics module and performance analysis module.
- System management module, its sub-function modules include user addition, user modification, user deletion, permission assignment, database backup, and database restoration.

The second-level sub-function modules are independent of each other to form a first-level function module. Each of the first-level function modules has functions of data exchange and data reading. Based on this mode, functional modules form a system to realize the overall function.

**Database design:** The structural design of the database adopted ER entity model. Firstly, all data entities including user information, student information, test types, test projects, test scores and PE performance information were summarized; Secondly, the relationship between the entities and the data information items contained in each entity were analyzed; finally, the data sheet design was obtained.

By transforming the E-R model of the PE performance management system into a relational model and combining with the concept and logical design of the database, various types of data storage forms of the PE performance management and analysis system were obtained, that is, various types of information tables. Table is a basic information structure of relational database SQL.

Table 1 shows the user information table, including user name, user password and user type.

Table 2 shows PE test type information, including PE type name, PE type description, PE type weight etc.

**Table 1.** The information of users

Fieldname	Description	Data type	Data length	NULL	Primarykey
UserId	UserID	Int	10	N	Y
UserName	Username	Varchar	20	N	N
UserPassword	Login password	Varchar	20	Y	N
UserType	Usertype	Bit	10	N	N

**Table 2.** Test type table

Fieldname	Description	Data type	Data length	NULL	Primarykey
TestTypeID	Test Type ID	Int	10	N	Y
TestTypeName	Test Type Name	Varchar	20	N	N
TestTypeDesc	Test Type description	Text	10	Y	N
Test type Weight	Test Type weight	Int	10	N	N

Other table types include test score table, student performance table, score conversion table etc. The forms are similar to the above two tables and will not be described one by one. The form of the above two are no longer describe one by one, and database design of the PE performance management and analysis system was completed based on regulated, accurate and standard principles.

## 4 Overall Implementation of The System

### 4.1 PE Performance management

The overall programming of the system was realized according to the overall architecture design and overall module composition of the system depending on the Visual Studio platform by means of C # language, the interface diagram of PE performance management part was displayed and illustrated. Figure 4 shows PE test type management interface map. Teacher users can add and delete various PE test types in the PE test types. In addition, the weight of each PE test type can be modified.

Figure 5 shows an interface diagram of a student score entry system. PE teachers can enter various sport scores of a student with a specific student ID. As shown in the figure, the student with a student ID of 3A7283945 logged in for sprint, basketball and hurdle levels.

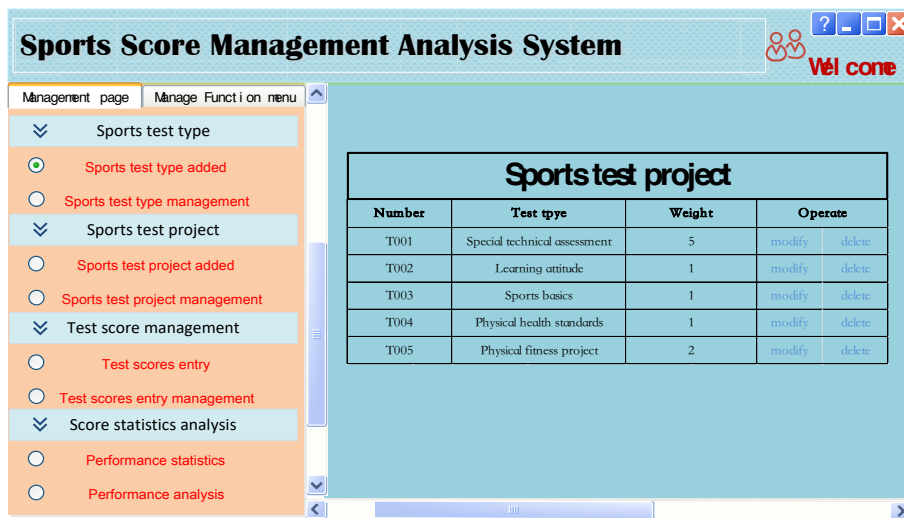


Fig. 4. PE test type management interface diagram



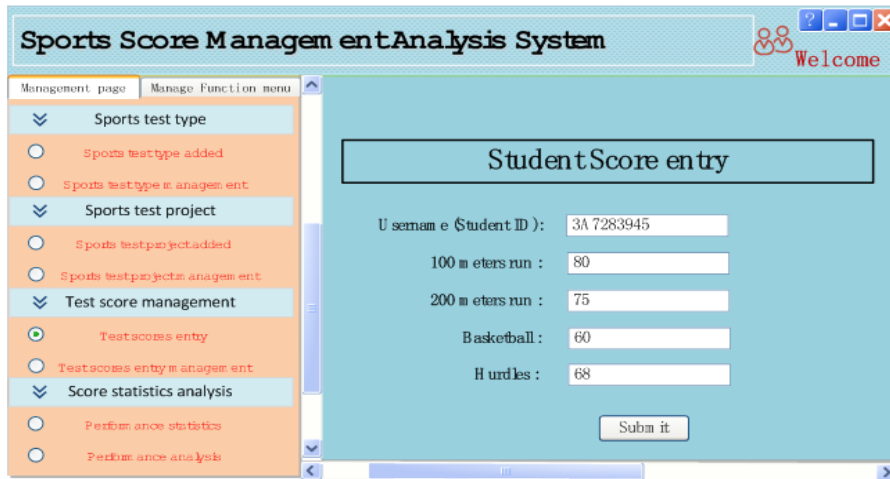


Fig. 5. Student score entry interface diagram

#### 4.2 Data mining implementation and pe performance analysis

Case: The decision tree ID3 algorithm was adopted to analyze that student sprint performance can be determined by which project score.

**The first step:** Regulate the database.

New Excel spreadsheets were created, eight attributes of hurdles, long jump, shot put, basketball, high jump, sprint, long-distance run and volleyball were converted into 0 or 1, where the value 0 represented failed scores < 60, 1 represented qualified sport scores  $\geq 60$ .

**The second step:** Select training example set, and get 159 student score records.

Table 3. Training set data

	Long jump	Long run	Basketball	Short run	Hurdles	Volleyball	Shot	High jump
Pass	80	54	32	30	37	87	37	100
Failing	79	105	127	129	122	72	123	59

**The third step:** Utilize information gain degree to select attributes that can best distinguish the actual training.

**The fourth step:** Substitute into the calculation according to the steps in Section 2.4 to obtain the information gain and tree node. According to the information gain degree of each attribute, the attribute of long-distance run should be chosen as the root node of the decision tree. Since the long-distance run had only two attribute values, 1 represented pass, 0 represented fail. Under the long-distance running node, a branch can be created. A screening function was applied to display the recordings of failed long-distance run and qualified sprint. This indicated that the long-distance run was failed and the sprint was basically failed. Therefore, the branch of failed long-distance running stopped branching.

The calculation results were as follows:

- IF long-distance running scores of students fail and were less than 60
- Then the students have a very high probability of failing in the sprint.
- The calculation process of the accuracy of the conclusion:  $(104-4)/104=96.2\%$
- The calculation process of coverage rate:  $104/159 = 65.4\%$

We can conclude that the degree of learning of a student's long-distance run can have a direct impact on his or her sprint performance. Therefore, teachers can indirectly improve the sprint performance by providing the long-distance score in the training of student sprint.

## 5 Conclusion

In the higher education of college students, physical education is an important link. With the teaching reform of other disciplines in the background of computer technology, how to realize the informationization and modernization of physical education and management in colleges and universities is a hot topic in current physical education. In order to save the time spent in PE teaching management and improve the efficiency of PE teaching management, this paper made full use of ASP.NET technology based on its 3-layer framework to fully study the functional requirements of the system. The system used SQL relational database to build a PE performance management system. And meanwhile the decision tree ID3 algorithm was used based on data mining technology; the analysis of the system's PE performance was realized. The main conclusions and significance of this paper are as follows:

- System construction fully considers the application requirements and functional requirements of physical education; combined with the functional division of system function modules, the system has better practical application values.
- Embedding of data mining function can perform data analysis based on the students' scores entered from the system.
- The establishment of the system is an effective step in the reform of physical education in colleges and universities. At the same time, it can provide relevant advices for teachers' PE teaching.

## 6 Acknowledgement

Application and Development of Intelligent Classroom in College Extracurricular Physical Activities (NKJ201802042).

## 7 References

- [1] Parris, P. D. (2012). Physical education basic instruction in united methodist colleges: a study of applied wesleyan principles. *Advanced Materials*, 24(46): 6234-6239.

- [2] Xu, J. (2012). The probe into the construction of young physical education teachers in common colleges and universities. *Advanced Materials*, 24(16): 2176-2185.
- [3] Le, Y. (2011). Application of computer and multimedia technology in college physical education. *IEEE*. <https://doi.org/10.1109/paccs.2011.5990343>.
- [4] Li, X., Jiang, Y., Lu, N., Liu, J. (2008). Several issues on establishment of computer application management system for college physical education. *Ieice Transactions on Communications*, 76(8): 913-920. <https://doi.org/10.1109/KAMW.2008.4810586>
- [5] Morita, A., Nakamoto, A., Takano, A., Ito, N. (2002). The study of the menstrual cycles of the students of the college for the physical education: report ii: interpretation of one year records of b.b.t. of players of volley and hand ball clubs in the college. *Biochemistry*, 41(42): 12771-12784.
- [6] Foggon, D. (2005). Beginning asp.net 2.0 databases: from novice to professional. *Journal of Drug Targeting*, 13(7): 399-406. <https://doi.org/10.1007/978-1-4302-0742-9>.
- [7] Spaanjaars, I. (2014). Beginning asp.net 4: in c# and vb. *Nucleic Acids Research*.
- [8] Litchfield, D. (2007). Microsoft SQL server passwords (cracking the password hashes). *Plant Physiology*, 140(2): 499-511.
- [9] Li, S. Q., Mi, H. L., Yang, M. D., Song, X. X., Zhao, Y. P. (2002). The application of ms SQL server copying technique in earthquake precursory data integration. *Journal of Seismology*.
- [10] Huang, X., Fan, X., Chen, X., Yang, G., Gan, M. (2016). Bed permeability state prediction model of sintering process based on data mining technology. *Isij International*, 56(12): 2113-2117. <https://doi.org/10.2355/isijinternational.ISIJINT-2016-193>
- [11] Wang, J., Zhang, F. (2015). Research on spatial data mining technology in agricultural spatial decision support system. *Journal of the Institute of Brewing*, 46(2): 73-74.
- [12] Delibašić, B., Vukićević, M., Jovanović, M., & Suknović, M. (2013). White-box or black-box decision tree algorithms: which to use in education?. *IEEE Transactions on Education*, 56(3): 287-291. <https://doi.org/10.1109/TE.2012.2217342>
- [13] Lee, K. M., Herrman, T. J., Bean, S. R., Jackson, D. S., Lingensfelder, J. (2007). Classification of dry-milled maize grit yield groups using quadratic discriminant analysis and decision tree algorithm. *Cereal Chemistry*, 84(2): 152-161. <https://doi.org/10.1094/CCHEM-84-2-0152>
- [14] Xiao, B., Li, L. Y. (2013). Design and realization of management system for university sports scores. *Applied Mechanics & Materials*, 336-338: 2533-2536. <https://doi.org/10.4028/www.scientific.net/AMM.336-338.2533>
- [15] Wang, Q. (2010). Design and implementation of sport score management system in universities & colleges. *Office Informatization*.

## 8 Authors

**Yimeng Fan**, female, 38 years old, who is a teacher of the Hebei Agricultural University and engaged in physical education teaching.

**Yu Liu**, Female, 42 years old, who is a teacher of the Baoding University and engaged in physical training.

**Haosong Chen**, male, 19 years old, who is a student of the college of mechanical and electrical engineering, Hebei Agricultural University.

**Jianlong Ma**, male, 54 years old, who is a teacher of the Hebei Agricultural University and engaged in physical education teaching.

Article submitted 2019-01-15. Resubmitted 2019-02-27. Final acceptance 2019-02-28. Final version published as submitted by the authors.