

Teaching Quality Monitoring of Higher Education Based on Data Mining

Xiaobing Jiang

Dongfang College, Shangdong University of Finance and Economics, Taian, China.

122666860@qq.com

Keywords: Higher Education, teaching quality, data mining, quality monitoring.

Abstract. In order to improve the teaching quality of higher education and improve the quality assurance system of higher education, based on the Internet + and under the background of big data, data mining technology is applied and a method for monitoring the teaching quality of higher vocational education is established. A variety of methods are used to clean up the monitoring data, get high-quality monitoring data, and construct the content system of school teaching quality monitoring. Then, according to the mode of "teaching input - teaching process - learning result", the data items which have high correlation with the quality of teaching are used as the core index of teaching quality, and the quality monitoring model of higher education is constructed from the core index of teaching quality. Moreover, the quality monitoring index of higher education is constructed in accordance with the weight of various data. Finally, a normal monitoring system for teaching quality of higher education is realized. The research shows that the teaching quality monitoring system proposed can effectively monitor the quality of teaching. The results can be used to amend the existing teaching quality evaluation indicators, and improve and enrich the five-in-one national higher education quality assurance system.

1. Introduction

With the rapid development of Internet technology, the great social value, economic value and scientific research value of big data have aroused great attention from all walks of life in the whole society. If the big data can be used effectively, it will greatly promote the development of the social economy and the level of scientific research. At the same time, the big data technology is also pregnant with unprecedented opportunities. The normal monitoring of the basic state of teaching is the main content and basic form of teaching evaluation. According to the internal rules of the teaching work and the big data technology of the Internet, the basic state of the higher education is reflected by the data, and the data analysis helps the school (Fang, Wu, Luo, Meng and Liu, 2013). The basic status data of teaching is an important measure to promote the development of university connotation and ensure the continuous improvement of the quality of teaching. It is also an important part of the state monitoring system of the quality of higher education in the national five-in-one evaluation system.

According to the analysis and research of the basic state data of higher education, the indexes of perfecting the quality evaluation of higher education are improved based on the results of data analysis, and the new methods and new ways of the teaching quality monitoring are formed. In addition, the main factors influencing the quality of teaching are analyzed and the teaching quality monitoring data model is constructed. It is of great significance to improve the scientific level of the quality monitoring and evaluation of higher education. It also plays an important role in providing comprehensive, accurate and specialized data support for the quality self-evaluation and improvement of higher education, as well as the formulation of the policies and measures for improving the quality of higher education.

2. State of the Art

In 1975, the United States established a specialized teaching quality management organization, the American Higher Education Accreditation Committee (COPA), and in 1993, the American Higher

Education Accreditation Committee renamed the National Higher Education Bureau for Identification Policy Committee (NBP). NBP has three main aspects in evaluating the quality of higher education teaching: Teachers' teaching, scientific research and public service. The quality monitoring of higher education in the United States is mainly based on certification, and the quality of higher education is guaranteed through various types of accreditation of non-governmental organizations and the ranking of various institutions. Schools or majors voluntarily participate in and consciously conduct self-quality assessment (Sun, 2014). The quality monitoring of higher education in Britain is universally recognized in the world. The quality monitoring of higher education in the Britain is monitored by the internal management and self-assessment of the quality of the higher education, the higher education evaluation carried out by the government's Higher Education Fund Committee and the higher education committee representing the Committee of the University Presidents. The quality audits and social evaluations carried out by the body consultative organization are composed of four levels of assessment, a typical social assessment of the British Higher School Rankings issued by the Times (Qu and Li, 2014). The quality evaluation of higher education in China has begun to try to make use of information technology since 2007, especially big data and Internet + technology to build the database of the basic state of teaching that can reflect the teaching quality of colleges and universities. It has been accumulated for 10 years in a continuous period, and the status data of many schools have been put into storage, but the analysis and application are not very well open. Therefore, in view of this situation, data mining technology is applied to analyze and utilize data, and a monitoring system for teaching quality of higher education is proposed.

3. Method

3.1. Data Mining.

Data mining is a new technology, which differs greatly from traditional data analysis methods. The traditional data analysis is often only the query, processing and analysis of the database, and the data mining is often not clear when it is analyzed. As a result, the final result of the data mining is much unknown, but in the meanwhile, it has the validity and practicality (Gou, 2013).

Clustering algorithm is a kind of unsupervised machine learning algorithm. The essence of the clustering algorithm is to group the data set which is not known beforehand, so that the data in the same group are as similar as possible and the data in different groups are as different as possible. The purpose is to reveal the true situation of the data distribution. The current clustering algorithms are roughly divided into five categories: partition method, hierarchical method, density-based method, grid-based method and model-based method.

K-means algorithm is a typical representative of clustering algorithm. In essence, the algorithm is based on the average value of objects in clusters. In order to achieve the global optimum, the partition-based clustering requires exhaustive division of all possible partitions (Li, Hu and Zhao, 2014). The algorithm process is to first select k objects from the entire n sample as the center m_i ($i=1, 2, \dots, k$) of the initial cluster. Using the formula (1), the distance $d(p, m)$ between each object P in the data set to k cluster center can be calculated. Then, the minimum of each object P is found and classified into the cluster the same as m_i . After traversing all the objects, the formula (2) is used to recalculate the value of m_i as a new cluster center. Reassign the objects in the entire dataset to the most similar clusters. Finally, repeat the above steps until the square error criterion is the minimum, as shown in (3).

$$d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{in} - x_{jn})^2} \quad (1)$$

$$m_k = \sum_{i=1}^N \frac{x_i}{N} \quad (2)$$

$$E = \sum_{i=1}^k \sum_{P \in C_i} |P - m_i|^2 \quad (3)$$

Analytic hierarchy process (AHP) is a hierarchical decision-making method combining qualitative analysis with quantitative calculation. The step is to construct the hierarchical structure model firstly. The hierarchical structure model is divided into the highest level, the middle level and the lowest level, which are used to refine the research topics from top to bottom. The highest level is the analysis of the target preset by the problem, which is called the target layer; the middle layer is also called the standard layer, including the middle link or element to realize the design of the target, which can contain multiple levels; the lowest level indicates the measures and schemes for the realization of the target, and is also called the scheme layer. Then, the structure judgment matrix is constructed. The judgment matrix is a matrix which is used to show relative importance by comparing the importance between the opinion layers. The relative weight vector W can be obtained by using the judgment matrix. The method of calculating weight mainly includes the sum method, the root method, the characteristic root method, the logarithm least square method, and so on, and the most widely used is the characteristic root method. At last, it is the hierarchy and consistency check. Consistency checking must be carried out when calculating weight vectors. The consistency of each layer of elements is compared by the consistency ratio (CR), and the consistency ratio is $CR=CI/RI$, in which CI is the consistency index and RI is the random index (Zhou, 2014).

3.2. Monitoring Data Processing.

In order to improve the quality of data, the data collected from higher education quality is checked according to the logical relationship between data items. After data verification, the data obtained are greatly improved in both quality and normalization, and some misinformation can be corrected. However, data verification cannot make it ruled out because of the data anomaly caused by users' understanding of data connotation and statistical time errors. Therefore, the core data related to the teaching quality are compared with the school's corresponding types, levels, properties and national regional norm data, and the data items which are larger than the norm are marked out and confirmed again.

According to needs, different data items are aggregated according to different classifications, and data cube is constructed. Statistical units with the data items not belonging to schools are summed up, and the data of school level are counted through COUNT, SUM and AVERAGE functions. The time data, such as the year of birth and the time for admission to the teachers, are calculated according to the statistical time points of the year into the data of the age or unit of the year, and the data are coded according to the needs. For example, the age can be divided according to the actual age: $\text{age} \leq 18$, $18 < \text{age} \leq 35$, $35 < \text{age} \leq 45$, $45 < \text{age} \leq 55$, $55 < \text{age} \leq 65$, $65 < \text{age}$, and $\text{age} \leq 18$. In addition, according to abnormal data, they are removed, and others are classified and coded in accordance with the interval range.

Teaching process is directly related to teaching input and data in learning outcomes. After careful analysis of the characteristics of teaching and the existing data, on the basis of the principle of "input - output", referring to various teaching evaluation indexes and according to the model of "teaching input - teaching process - learning result", the data items which have high correlation with the quality of teaching are selected as the core index of teaching quality. The hierarchical structure model of higher education teaching quality is constructed, as shown in Figure 1.

The standard judgment matrix of the model is the result of the importance of three factors: the teaching input factor, the teaching process factor and the learning result factor. The judgment matrix of the teaching input, the teaching process and the learning result is A , in which $C.R. = 0.0707$ and the judgment matrix satisfies the consistency test.

$$A = \begin{bmatrix} 1 & 1/3 & 3 \\ 3 & 1 & 4 \\ 1/3 & 1/4 & 1 \end{bmatrix} \quad (4)$$

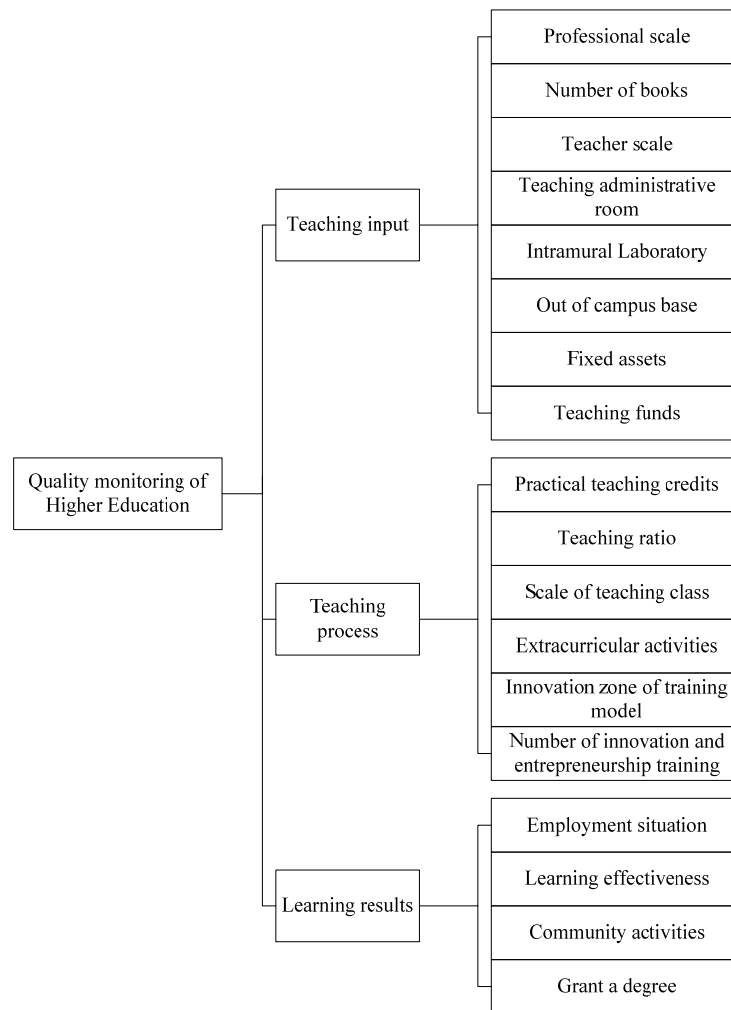


Figure 1. The hierarchical structure model of higher education teaching quality

Sub criterion level is divided into teaching input sub criteria level, teaching process sub criteria level and learning result sub criteria level. The judgment matrix of the teaching input sub criterion layer is the professional scale, the number of books, the size of the teachers, the teaching administrative room, the laboratory, the outside base, the fixed assets and the teaching funds. The matrix A_1 is judged according to the importance result, in which $C.R. = 0.967$ and the judgment matrix satisfies the consistency test.

$$A_1 = \begin{bmatrix} 1 & 9 & 1 & \frac{1}{4} & 1 & 1 & 2 & \frac{1}{2} \\ \frac{1}{9} & 1 & \frac{1}{5} & \frac{1}{9} & \frac{1}{6} & \frac{1}{8} & \frac{1}{5} & \frac{1}{4} \\ 1 & 5 & 1 & 1 & 1 & 1 & 5 & 3 \\ 4 & 9 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 6 & 1 & 1 & 1 & 1 & 1 & \frac{1}{3} \\ 1 & 8 & 1 & 1 & 1 & 1 & 1 & \frac{1}{3} \\ \frac{1}{2} & 5 & \frac{1}{5} & 1 & 1 & 1 & 1 & 1 \\ 2 & 4 & \frac{1}{3} & 1 & 3 & 3 & 1 & 1 \end{bmatrix} \quad (5)$$

The judgment matrix of the sub criterion level of the teaching process is the practice teaching credits, the teaching class scale, the teaching ratio, the extracurricular activities, the training mode innovation area, and the number of innovative entrepreneurship training. Moreover, the matrix A_2 is

judged based on the importance results, in which C.R. =0.0758 and the judgment matrix meets the consistency test.

$$A_2 = \begin{bmatrix} 1 & 9 & 3 & 6 & 9 & 9 \\ \frac{1}{9} & 1 & \frac{1}{9} & 1 & 2 & \frac{1}{2} \\ \frac{1}{3} & 9 & 1 & 3 & 9 & 7 \\ \frac{1}{6} & 1 & \frac{1}{3} & 1 & \frac{1}{2} & 1 \\ \frac{1}{9} & \frac{1}{2} & \frac{1}{9} & 2 & 1 & 1 \\ \frac{1}{9} & 2 & \frac{1}{7} & 1 & 1 & 1 \end{bmatrix} \quad (6)$$

The judgment matrix of the learning result sub criterion layer is the employment situation, the school effect, the community activity and the degree situation. The matrix A3 is judged according to the importance result, in which C.R. =0.0579 and the judgment matrix satisfies the consistency test.

$$A_3 = \begin{bmatrix} 1 & 9 & 9 & 3 \\ \frac{1}{9} & 1 & 1 & \frac{1}{9} \\ \frac{1}{9} & 1 & 1 & \frac{1}{9} \\ \frac{1}{3} & 9 & 9 & 1 \end{bmatrix} \quad (7)$$

According to the judgment matrix, the relative weight vector W is determined by the power calculation method of the judgment matrix, and the corresponding weights are determined. The weight of each data item is shown in Figure 2.

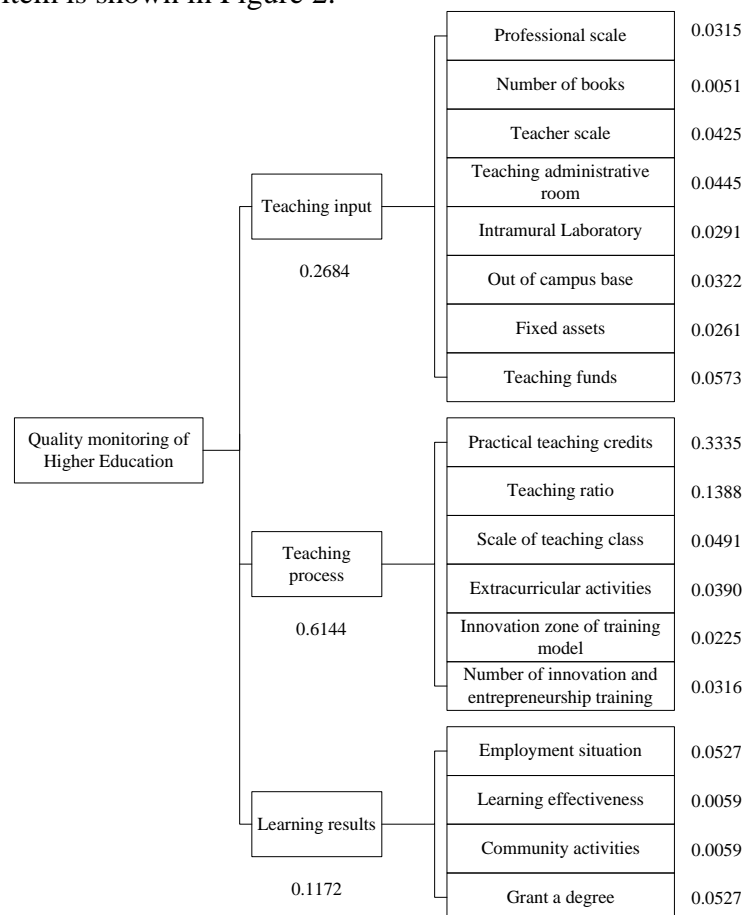


Figure 2. The weight of the hierarchical structure model of higher education teaching quality

According to the weight of each data, the monitoring index of teaching quality is set up and recorded as HI, and the formula is:

$$HI = \sum_{i=1}^{18} x_i \omega_i \quad (8)$$

In the above formula, x_i suggests the data item and ω_i refers to the weight of each data item.

In 2016, the National University data is selected as the research object. In accordance with the quality monitoring index of higher education, the changes of the provinces can be seen through the changes of the monitoring index. The changes of the teaching quality monitoring index in some colleges and universities in 2017 are shown in Figure 3.

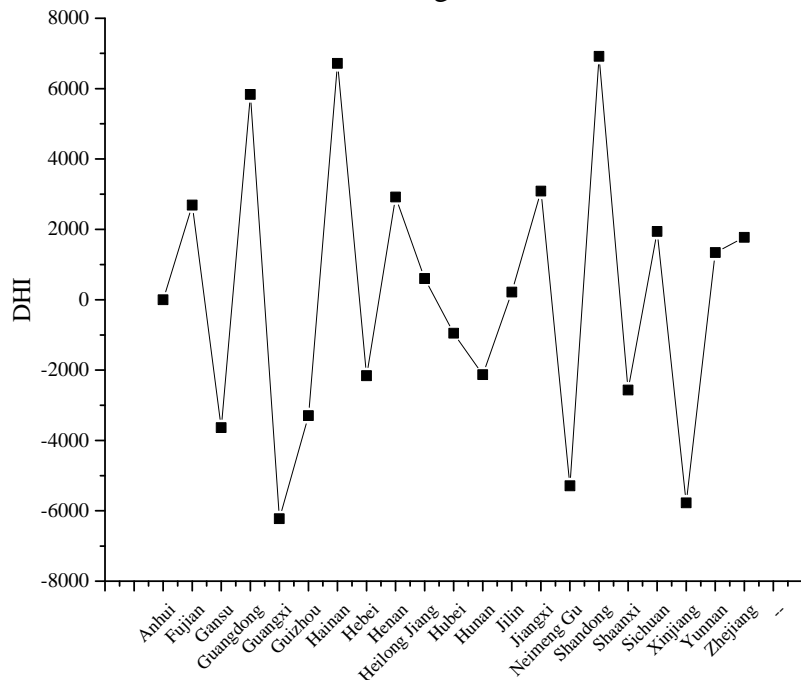


Figure 3. Changes in monitoring index of education and teaching in some universities in 2017

3.3. Design of Teaching Quality Monitoring System.

The monitoring system structure includes data integration layer, data analysis layer and data presentation layer, so as to realize the processing and analysis of teaching quality data. The bottom is the data integration layer, which mainly integrates the data filled by universities in the collection system into the major database (data warehouse), including the refined professional data and the expanded evaluation data, the survey data in the satisfaction survey system, and the public data on other public platforms, through the ETL data cleaning and extraction. In this process, data cleaning, data transformation and data integration are mainly completed. The middle layer is the data analysis layer, mainly using the data in the major database, through the multi-dimensional data analysis and mining, to generate various uses of the analysis library. It includes the query of various data, statistical analysis results of all kinds of data, various indicators and quality monitoring warning data, various quality reports generated by the use of data, and professional evaluation authentication data. The top layer is the data display layer, which mainly provides data services to the colleges and universities, the provincial education authorities, the state education authorities, various evaluation institutions, evaluation experts and the public. There are both the multi terminal data information publishing platform and the other work platforms, such as evaluation management system, engineering professional certification management system and other work platforms to achieve data resource sharing.

The functions of the monitoring system include the general situation of the school, the whole situation of the school, the situation of each teaching unit, the professional situation and the data inquiry, and the overall functional structure of the system is shown in Figure 4.

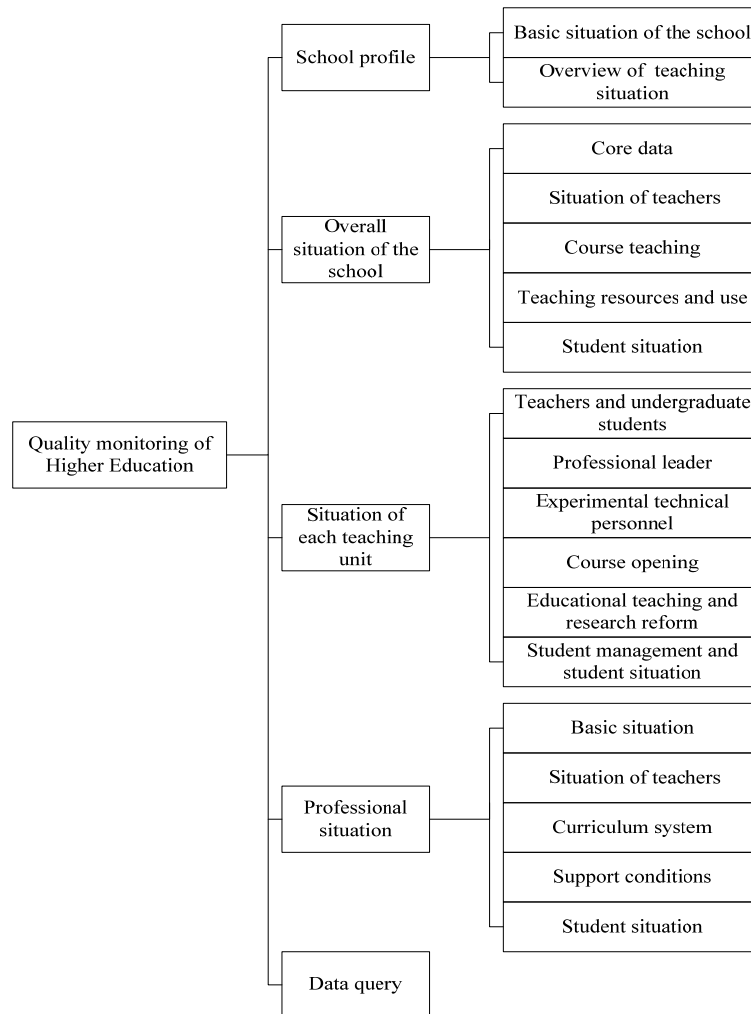


Figure 4. Overall functional structure diagram of the system

4. Results

The annual data of a university are selected to test the function of the main modules. After introducing the data into the database, the test function of each module is realized, and the test information is recorded in detail, including the test modules, the problem of the function realization, the cause of the problem and the severity, as well as the subsequent improvement method.

After testing the function of major modules, the system basically meets the goal of system requirement analysis, and all functional modules can run normally. In terms of function, the basic function points of CRUD have been realized in the general situation management of the school, the whole situation of the whole school, the situation of each teaching unit, the professional situation and the data query. In view of security, users will automatically jump to the login page if they access the system directly through URL without login. All system resources need to be authenticated by users before they can be accessed. For ease of use, the system interface is simple and easy to operate, realizing the ease of users to access and manage the system.

5. Conclusion

The data items which have high correlation with the quality of teaching are analyzed by data mining technology, and the quality monitoring model of higher education is constructed based on the core of the teaching quality. Additionally, the quality monitoring index of higher education is built according to the weight of various data. Finally, a normal monitoring system for teaching quality of higher education is realized. Through the analysis and implementation of the quality monitoring data of higher education, this paper provides a basis for perfecting the quality assurance system and

improving the quality of teaching in colleges and universities. What's more, it provides a reference for the scientific decision-making of the educational authorities. The research results can be used to amend the existing teaching quality evaluation indicators, and improve and enrich the five-in-one national higher education quality assurance system.

References

- [1]. Fang, M., Wu, W., Luo, J., Meng, K., &Liu, Y. (2013). Research on Teaching Quality Monitoring System of Management by Objectives. *Proceedings of the International Academic Workshop on Socialness*, 50 (1), 848-851. Doi: 10.2991/iaw-sc.2013.192.
- [2]. Gou, G. (2013). Evaluation of Mathematical Teaching Quality for the Universities based on Data Mining. *Journal of Convergence Information Technology*, 8 (5), 661-668. Doi: 10.4156/jcit.vol8.issue5.76.
- [3]. Li, AF. Hu, Y., & Zhao, WJ. (2014). Research on Teaching Quality Evaluation Using Data Mining Technique. *Advanced Materials Research*, 926-930, .4582-4585. doi.org /10.4028/www.scientific.net/AMR.926-930.4582.
- [4]. Qu, H., & Li, X. (2014). Application of Data Mining in the Assessment of Teaching Quality. *Lecture Notes in Electrical Engineering*, 269, 1813-1819. doi 10.1007/978-94-007-7618-0_212
- [5]. Sun, P. (2014) The Application of Cloud Model in Teaching Quality Monitoring and Evaluation. *Applied Mechanics & Materials*, 519-520, 1688-1691. Doi: 10.4028 /www.scientific.net /AMM.519-520.1688.
- [6]. Zhou, LM. (2014). Application of Data Mining Technology in Sports Teaching Quality Evaluation. *Applied Mechanics & Materials*, 543-547, 4694-4697. doi.org /10.4028/www.scientific.net /AMM. 543-547. 4694.