Design of University Financial Decision-Making Platform Based on Data Mining

Yueting Tang*, Yulin Lan
Weifang Engineering Vocational College, Shandong, China

*Corresponding author: tangyt@wfec.com

Abstract. Under the environment that the government and the society attach great importance to, the management of colleges and universities should also keep pace with The Times. Colleges and universities should carry out scientific and standardized reforms from ideas, objectives to methods and other aspects. The database data such as the university information system provides the data foundation for data mining, greatly improves the financial data processing capacity of colleges and universities, and can make financial decisions on this basis. By analyzing the financial data of colleges and universities, this paper constructs the financial data warehouse, establishes the analysis model and combines the data mining technology to study the financial decision support system of colleges and universities.

Keywords: data mining, university finance, decision-making platform

1. Introduction
Along with the reform of higher education and public fiscal system reform gradually deepening, financial management of colleges and universities is facing more complicated situation, the government and the school financial management in colleges and universities are also put forward higher request, environment, profound changes have taken place in financial management in colleges and universities financial work is gradually from the business accounting type to analysis type, management transformation. Data mining, also known as knowledge discovery in databases, generally refers to the process of searching hidden information from a large number of data through algorithms. In western countries, the research on financial analysis of colleges and universities mainly focuses on the performance evaluation of colleges and universities. Because of the diversity and uncertainty of the sources of funds in colleges and universities, it is necessary to pay attention to the economic benefits and financial analysis of funds. In teaching practice, data mining tools suitable for actual needs should be developed according to practical application. Financial analysis should also keep up with the needs of financial development and strengthen the depth and breadth of analysis.[1-2]

School students financial decision-making system, the author of this paper the problems existing in the database, through the data warehouse design, combining with the clustering algorithm based on density clustering of students tuition amount, by decision tree algorithm for year of college, professional, and tuition fees behavior prediction, in colleges, the number of various professional fees and tuition proportion of deep mining, until the excavation to the most detailed information.[3-5]
2. University financial decision-making platform for data mining

Data mining has become a new information processing technology. Its main feature is to extract, transform, analyze and process the massive data in the database, and extract the key data for decision making. And the key data or information we get from data mining must be beneficial to the development of the whole data mining field, but also beneficial to social life, to meet the actual needs, so that data mining technology is more meaningful. By studying the process of data mining, we can understand the process of data mining and find out ideas and methods to solve problems. The process of data mining can be understood through the flow chart of data mining. The process of data mining is shown in Figure 1.[6-9]

![Data mining process](image)

Fig. 1. Data mining process

By studying the demand for student payment analysis in the university financial decision-making system, DBSCAN algorithm, a density-based clustering algorithm, is selected to cluster the tuition amount of students, and ID3 algorithm, a decision tree algorithm, is used to predict the tuition level and tuition proportion of students. DBSCAN algorithm is used to cluster the data and obtain the maximum set of points connected with density through density collection. The area with sufficient density to be queried is divided into clusters, and arbitrary shape clustering points can be found in the spatial database with noise and null values. DBSCAN is a density-based clustering algorithm that looks up samples in this radius. Minp is a constraint on the number of samples found in radius n. The process of programming is shown in Table 1.[10]

### Table 1. Programming process

|   |   |
|---|---|
| (1) | As long as \( n \geq \text{minp} \), the sample point you find is the core point. |
|    | private int x; |
|    | private int y; |
|    | private boolean isKey; |
|    | private boolean isClassed; |
| (2) | All points in the text are extracted and stored in the Points List. |
| (3) | Check whether the given point is a core point and determine the following variables: |
|    | A list of LST storage points |
|    | P points to be tested |
|    | E radius |
|    | Minp density threshold |
|    | TmpList temporarily stores points that have been accessed |
| If an item specified in object set B does not exist in object set A, it is added to object set A. |
|    | boolean merge=false; |
|    | for(int index=0; index< b.size(); ++index) |
|    |   |
|    | if(a.contains(b.get(index))) |
|    |   |
|    | merge = true; |
|    | break; |
if(merge) {
    for(int index=0; index<b.size(); ++index) {
        if(!a.contains(b.get(index))) {
            a.add(b.get(index));
        }
    }
}
return merge;

(5) Find all the clusters that can be reached directly.

(6) Merge all the direct clustering, that is, find out the indirectly accessible points and merge.
int length = resultList.size();
for(int i=0; i<length; ++i) {
    for(int j=i+1; j<length; ++j) {
        if(Utility.mergeList(resultList.get(i), resultList.get(j))) {
            resultList.get(j).clear();
        }
    }
}
return resultList;

(7) Find out all the clustering.

(8) Input the result list of clustering that has been completed.

3. Experiment

For the clustering and experimental results based on DBSCAN algorithm, a set of two-dimensional test samples are given as follows for easy understanding: The initial data are shown in Table 2:

Table 2. Training sample set

| Training sample field 1 | Training sample field 2 |
|-------------------------|-------------------------|
| 2                       | 3                       |
| 2                       | 3                       |
| 1                       | 3                       |
| 1                       | 1                       |
| 0                       | 2                       |
| 5                       | 6                       |
| 6                       | 6                       |
| 7                       | 7                       |
| 6                       | 5                       |
| 10                      | 2                       |
| 9                       | 4                       |
| 10                      | 6                       |
| 10                      | 5                       |
| 11                      | 3                       |

For the above data, it can be visually displayed through the chart, as shown in Figure 2:
Cluster radius and cluster density were set by DBSCAN algorithm: 3 and 3 respectively. The clustering results are shown in Table 3 as follows:

**Table 3. Clustering results**

| The 1ST cluster  | (2, 3.0) |
|------------------|----------|
|                  | (2, 3.0) |
|                  | (1, 3.0) |
|                  | (1, 1.0) |
|                  | (0, 2.0) |
| The 2ND cluster  | (5, 6.0) |
|                  | (6, 6.0) |
|                  | (7, 7.0) |
|                  | (6, 5.0) |
| The 3RD cluster  | (10, 2.0) |
|                  | (9, 4.0) |
|                  | (10, 5.0) |
|                  | (11, 3.0) |
|                  | (10, 6.0) |
|                  | (12, 6.0) |
|                  | (12, 7.0) |

By comparing the position of sample points shown in Figure 2 with the clustering result of the algorithm, it is not difficult to find that the clustering result can achieve the expected classification effect well. College financial decision software design mainly for the leaders to make decisions, first of all, through the system log in to view to all the students' tuition information, then according to the number of query results in mining to each college tuition, and the proportion of tuition, further excavations can be a corresponding professional college tuition information, and so on, can dig, step by step to get leaders want information. The design framework of the system is shown in Figure 3, which is divided into three levels.
The system is designed using Java language, the development tool is Myeclipse, the whole process is the Web development process, need to understand Java framework knowledge, JSP syntax, script elements, hidden objects and tag library. It creates the actual Web application, implements application security measurements, and integrates the database into its OWN JSP structure. Myeclipse is the Java development tool used in the design process of financial system in colleges and universities, and its interface is shown in Figure 4:

Fig. 4. Myeclipse development tool interface

By changing the density threshold and radius threshold, the clustering results can be analyzed. In this paper, the amount of data tested on the financial system of colleges and universities is relatively large, so the density threshold is 1, radius threshold is 100 clustering. This algorithm is sensitive to threshold change. Through testing, when the density threshold is 1, the clustering result data with arbitrary change of radius threshold will not be lost and can be processed well.

4. Conclusion
Data mining is a multidisciplinary field, which combines the research results of database technology, artificial intelligence, machine learning, statistics, knowledge engineering, information retrieval and other latest technologies, and its application is very extensive. The "student payment system" of Qiqihar University has a large amount of data, so it is suitable to use data mining technology to analyze students' payment behavior, and provide a visual reference basis for some related management and decision-making work. Data mining technology can find useful knowledge from a large amount of data, which will be of great significance to the decision-making of college teaching management.

References
[1] Feng Li, Shilun Ge, etc. Research Index of Talents in Universities and Colleges based on Data Mining [J], Science and Technology Management Research.2009, 9:19-198.
[2] Lili Wu, Bo Zhang, etc. Analysis and implementation of intelligent analysis module of university scientific research management system [J], computer engineering and design.2010,31(11):2641-2642.
[3] Wuqi Gao, Fengju Kang et al., Process Improvement and model Application of Data Mining [J], Microelectronics & Computer.2011, 7(28) :9-10.
[4] Lanfang Yu, mei shi. Research on data warehouse design of university management information system [J]. Science and technology bulletin. 2012,2(28):190.
[5] Yue He, Dayong Chen, etc. Research on user browsing interest path based on web data mining [J], computer engineering and application.2012,48(7):106-108.
[6] Ester Metal. A densith-based algorithm for discovering clusters in large spatial databases with noise[C]. In:Pro of 2nd Int'1 Conf on Knowledge Discovering in Databases and data mining(KDD-
96).Portland:AAA I Press,1996.

[7] Liang Cao and Hong Shu are equal. DBSCAN improved algorithm based on correction of observed data bias [J]. Computer engineering and design. 2009,30(13):3160-3162.

[8] Boyuan Liu, Wenhui FAN, Tian-yuan XIAO. Analysis of research status of decision support systems [J]. Journal of system simulation. 2011, 7(23).

[9] Wenwei Chen. Decision support system and its development [M]. 2nd edition. Beijing: Tsinghua University Press, 2000.

[10] GAGorry, MS Scott Morton. A framework for management information systems [J]. Sloan Management Review (S0019-848X), 1971, 13(1):50-70.