Data Based College Students' Comprehensive Quality Evaluation System

Li Li
School of Electronics and Information Engineering
Tongji University
Shanghai 201804, China
lili@tongji.edu.cn

Shaoming Yao
School of Electronics and Information Engineering
Tongji University
Shanghai 201804, China
1333705@tongji.edu.cn

Zhou Ou*
School of Electronics and Information Engineering
Tongji University
Shanghai 201804, China
18817366097@163.com
*Corresponding author

Qijun Chen
School of Electronics and Information Engineering
Tongji University
Shanghai 201804, China
qjchen@tongji.edu.cn

Abstract—As EQO(Essential Qualities Oriented) education has become consensus of society, the evaluation of college students' comprehensive quality stands at the vital position in the work of educational workers. First college students' comprehensive quality evaluation system is built based on some University's information system. Then framework of comprehensive quality evaluation system is designed. Finally, an application based on the framework is accomplished with the database technology and data mining technology. The complete of the application, which works as an auxiliary of related departments' decision-making, shows the process of realizing the framework.

Keywords—comprehensive quality evaluation; data mining; Apriori; association rules

I. INTRODUCTION

With the rapid development of modern science and technology, society has higher criteria for talents. To adapt to the demands of new era, the Ministry of Education has paid more attention to the cultivation of college students' comprehensive quality. At present, the implementation of EQO education has become consensus of the whole society [1].

In the meantime, the information construction of colleges is advancing at an unprecedented speed. The campus local area network with advanced performance covers nearly all colleges. And the construction of data warehouse systems enhances the schools’ management and decision-making ability [2]. However, the amount of data stored by colleges is rising rapidly, which leads to a concern on how to use colleges’ database system efficiently.

Big data era has a vital influence on our work, thinking and life. From Google’s prediction of H1N1 to Taobao’s accurate recommendation, big data era has been deeply into modern life. Data mining technology appears under such a background [3]. With data mining technology and related data analysis algorithms, we can deal with those data and find rules and patterns to help people with their decision-making.

II. THE ESTABLISHMENT OF COLLEGE STUDENTS’ COMPREHENSIVE QUALITY EVALUATION INDEX

A. Design Principles

To ensure the quality of comprehensive evaluation indexes, which is the basis of the comprehensive quality assessment, the system must follow some basic principles [4-5]:

1) Principle of comprehensiveness: Each evaluation index can comprehensively and accurately cover, and describe the characteristics of college students' comprehensive quality.

2) Principle of hierarchy: The principle of hierarchy reflects developmental evaluation thought.

3) Principle of representative: The optimized index can be in accordance with students' personality development and can typically reflect the college students' comprehensive quality.

4) Principle of independence: The index system must be independent of each other in the same level of indicators.
5) **Principle of feasibility**: Each evaluation index should be operable and effective.

B. **The Establishment of a Comprehensive Quality Evaluation Index System**

In this paper, we evaluate students' comprehensive quality based on some University’s information system. As our evaluation objects are numerous and the complexity of data is high, we use hierarchical structure [4-5] to establish college students' comprehensive quality evaluation index system.

The college students' comprehensive quality evaluation index hierarchical structure model (Tab. 1) is divided into three levels according to the design principles of comprehensive quality evaluation index system.

### III. **DESIGN OF COMPREHENSIVE QUALITY EVALUATION SYSTEM FRAMEWORK**

#### A. Design Idea

The design of comprehensive quality evaluation system framework is the core part of this paper. The framework is divided into three parts including the establishment of data layer, model layer and application layer.

The data layer is used to extract the relevant data samples, which depends on a specific application. And the model layer is designed to analysis the data samples and to get the information needed. And the construct of application layer is to meet the demand of users in an accurate and concise way.

Fig. 1 shows some details and a specific relationship between each layer.

#### B. Introduction of Each Layer

1) **Data layer.** The data layer mainly completes the work of data sampling and data storage. Data warehouse of some University information system unifies data information of departments and provides convenience for related research. And Oracle database [6] is adopted and helps us realize the establishment of data layer based on a specific application in the application layer.

2) **Model layer.** This layer mainly introduces data mining technology. Data mining is the process of extracting valuable knowledge hidden in large volumes of raw data. It is a computer-assisted process of digging through numerous sets of data, which are large, incomplete, noisy, fuzzy and random, and extracting the meaning of the data [7-8].

3) **Application layer.** The application layer is mainly designed to provide customers with applications, which includes employment recommendation, students’ behavior survey, graduate students selection, academic performance prediction and so on. And it ultimately shows the required information to the customers in an accurate and concise way. These applications can be used as an auxiliary of related departments’ decision-making.

<table>
<thead>
<tr>
<th>First level index</th>
<th>Second level index</th>
<th>Third level index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Academic quality</strong></td>
<td>Ability of practice</td>
<td>Internship experience</td>
</tr>
<tr>
<td></td>
<td>Professional performance</td>
<td>Grades of practical courses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Social practice experience</td>
</tr>
<tr>
<td>Learning habits</td>
<td>Professional performance</td>
<td>Grades of professional courses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grades of Minors</td>
</tr>
<tr>
<td></td>
<td>Basic skills</td>
<td>The percentage rank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grades of basic courses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Foreign language level test</td>
</tr>
<tr>
<td>Prizes</td>
<td>Innovation ability</td>
<td>Computer level test</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mandarin proficiency test</td>
</tr>
<tr>
<td>Ideological and moral qualities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disciplinary records</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disposition records</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Owing Money records</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades of ideological courses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political status</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Living habits</strong></td>
<td>Diet status</td>
<td>Eating times</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trading volume</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumption type</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumption site</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Card using times</td>
</tr>
<tr>
<td>Family background</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subsidy information</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Physical and psychological health</strong></td>
<td>Physical health</td>
<td>Grades of physical courses</td>
</tr>
<tr>
<td></td>
<td>Psychological health</td>
<td>Sports competitions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mental health status</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Psychological questionnaire</td>
</tr>
</tbody>
</table>
IV. THE DESIGN AND IMPLEMENTATION OF COMPREHENSIVE QUALITY EVALUATION SYSTEM

In this chapter, we will put the college students’ comprehensive quality evaluation system into practice by taking students’ behavior survey as an example and show the process of realizing the framework.

A. The Establishment of Data Layer

Generally, data warehouse adopts three-layer structure [7-8]. Local database is established by connecting the data warehouse of school’s information system. However, the raw data extracted cannot be directly used, which must be preprocessed. Data preprocessing includes data cleaning, data integration, reduction and transformation [7-8]. After preprocessing, the data can be analyzed by data mining tools. We choose the students of grade 11 from SEE(School of Electricity) and SEM(School of Economy Management) as data samples. In horizontal direction, we search for the difference between academies. And in vertical direction, we search for the changing of students in different years. Basic statistics are shown in Tab. 2.

Fig. 2 shows a portion of data samples of SEM_2013 (students of SEM in 2013). The attributes chosen include sex, year of birth, book renal times, subsidy information, campus card use times, consumption and scholarship information.

B. The Establishment of Model Layer

1) The introduction of association rule mining. Mining association rules stands for the process of finding all of the valuable associations existing in the data sets made up of a great number of data that meet certain requirements of minimum support and the minimum confidence [9-10]. Association rules mining is normally a two- step process. The first step is to find all the frequent item sets. And the second step is to generate strong association rules. The rules that meet both the requirements of support and confidence are named strong rules. There are plenty of association rule mining algorithms, such as Apriori, FP-Tree, and FP-Growth.

2) The Introduction and Realization of Apriori Algorithm. Apriori algorithm is one of the most classic algorithms of mining association rules. The iteration of layer by layer is used. The 1_item set is obtained if the support is satisfied. Similarly, the (k+1)_item set is explored by the k_item set until item set no longer satisfies the minimum support. Apriori algorithm can be described in the following pseudocode [9-10].

1. \( L_1 = \text{find frequent 1-itemsets}(D) \)
2. \( \text{for } (k = 2; L_{k-1} \neq \emptyset; k++) \{
\)
3. \( C_k = \text{apriori_gen}(L_{k-1}) \)
4. \( \text{for each } t \in D \{
\)
5. \( C_t = \text{subset}(C_k, t) \)
6. \( \text{for each candidate } c \in C_t \{
\)
7. \( c.\text{count}++; \}
8. \( L_k = \{ c \in C_k | c.\text{count}\geq \text{min_sup} \}
\)
9. \( \text{return } L = \text{all the frequent item sets} \);
In this report, we provide related customers with conclusions which show advice analyzed by horizontal and vertical comparison with association rules.

V. CONCLUSION

In this paper, we first complete establishment of comprehensive quality evaluation index system based on design principles. Then we construct the comprehensive quality evaluation system framework that includes data layer, model layer, and application layer. Next we choose students’ behavior survey to show process of realizing the framework. We use Apriori as our mining algorithm to dig up potential association rules. After analyzing rules from horizontal and vertical direction, we display conclusions to customers with analysis reports. In future work, a more concise and direct way to display the conclusion will be considered and related application software will be designed.

ACKNOWLEDGEMENT

I would like to express my gratitude to all related people and departments that support the research and provide the data.

REFERENCES


[7] Jiawei Han, Micheline Kamber, Jian Pei. Data Mining:Concepts and Techniques, MorganKaufmannPublishers, 2012

