Original Paper

Evaluation Method of Effectiveness of College Students' Employment Guidance Strategy Based on Data Mining

Dong Fei[1] & Jinghan Song[2]

[1] Southwest Petroleum University, Chengdu, China

[2] Langfang, China

Abstract

Aiming at the problem of low confidence in the effectiveness evaluation method of traditional college students' employment guidance strategy, an effectiveness evaluation method of college students' employment guidance strategy based on data mining is designed. The effectiveness evaluation index system of college students' employment guidance strategy is constructed. The system is divided into three levels: target level, criterion level and index level, and the analytic hierarchy process is used to give weight to each index. According to the importance of the index, the effectiveness of college students' employment guidance strategy is evaluated based on data mining. The experimental results show that the confidence of the designed evaluation method is significantly higher than that of the control group, which can solve the problem of low confidence of the effectiveness evaluation method of traditional college students' employment guidance strategies.

Keywords

data mining, college student, employment guidance strategy, effectiveness evaluation method

1. Introduction

With the development and integration of database management and parallel computing technology, data mining comes into being. Data mining is to mine valuable information and knowledge hidden behind data. It has been widely used in many fields, which has also brought great help to people's correct decision-making. It also known as knowledge discovery based on database, was first proposed at the 11th academic conference of American Artificial Intelligence Association held in August 1989. Data mining integrates the theories, technologies and methods of many disciplines, such as machine learning, pattern recognition, database management, AI and mathematical statistics. It has been widely used in many fields, such as finance, market, scientific research and sports.

The transformation of China's higher education from elite to popularization makes the enrollment scale of colleges and universities expand year by year. In order to facilitate the daily management of teaching and employment, colleges and universities have established corresponding information systems using computer and database technology. However, with the increasing enrollment of colleges and universities, the amount of data stored in its information system database is also gradually increasing. Facing the huge data set, the traditional data analysis method has been difficult to meet the real needs of managers. Because the existing information system is mainly based on query, the database can efficiently realize the functions of data storage, query, statistics and sorting. However, these functions can only obtain the surface information of the data, that is, they can not obtain the deeper and more important information hidden behind these data, can not find the valuable rules hidden, and can not predict the development trend. To make full use of these data and improve the utilization of information, we need to use data mining to analyze and process data quickly and automatically.

With the popularization of higher education, most colleges and universities in China have expanded their enrollment to varying degrees according to their own situation. In addition, they have to face the competition of labor-intensive and technology intensive employment force every year, which leads to the increasingly severe employment situation. The difficult employment of graduates is gradually becoming a problem that can not be ignored, but on the whole, the total demand for talents in China is insufficient. According to the statistics of the Ministry of education, there are about 11 million to 12 million new employment opportunities every year (2006-2012), while the number of graduates is about 6 million every year (2006-2012), and more than 6.8 million in 2012. In view of the above data comparison, college graduates should still have relatively large employment space. However, China has obvious deficiencies in dredging channels and employment services, and there is also great room for expansion. For colleges and universities, in terms of college students' employment guidance, although the employment guidance model accumulated year after year has been more comprehensive, they all seem more formal and fixed, which is still difficult to meet the needs for the content of employment guidance. In the outline of the national medium and development plan (2010-2020), it is mentioned that "to meet the needs of national and regional economic and social development... Focus on expanding the training scale of applied, compound and skilled talents...". The outline of Chongqing's medium and long-term urban and rural education reform and development plan (2010-2020) also points out that it is necessary to strengthen employment guidance and career planning education, and guide graduates to establish a correct concept of employment and employment. Broaden employment channels, promote full employment of graduates and improve employment quality... The state and Chongqing have put forward guiding opinions on college students' employment. According to the guidance of these policies, how to improve the effectiveness of guidance strategies is a problem to be solved. Therefore, it takes the information data of guidance strategies as the research object, uses the decision tree classification technology to mine the data, mine the effective information for employment, and then apply the data mining results to employment guidance to provide decision-making basis.

2. Data Mining

With the continuous development of database and the wide application of system, like banks, insurance companies, securities companies and most financial service institutions, many organizations or enterprises have accumulated a large amount of data in their daily business activities (Justin & Ellen, 2019). However, most organizations or enterprises do not convert their data into valuable capital. The reason for this situation is that the information hidden behind these data is not so easy to be found (Pan & Center, 2019). In order to effectively participate in today's market competition, managers must find and reasonably use the valuable information hidden behind the historical data. The process of discovering and using the information hidden behind historical data is data mining.

Data mining refers to a non trivial process of extracting effective, implicit, potentially useful and ultimately understandable information and knowledge from a large amount of data in the database (Zhu, 2019). Data mining is a decision support process closely related to computer science. It can fully and automatically analyze the data of organizations or enterprises and make reasonable judgment and logical reasoning through machine learning, expert system, pattern recognition, database technology, mathematical statistics, parallel computing, visualization technology and other methods. And then dig out the potentially valuable information and knowledge to help management decision-makers adjust market competition strategies, reduce unnecessary risks and make favorable decisions. According to the difference between statistics and non statistics, data mining can be divided into two different types: hypothesis driven and discovery driven data mining (Zhou, Dan, Wen-Li, et al., 2019). The biggest difference between the two is the process of obtaining information.

Data mining is a complete and cyclic process of human-computer interaction. It is usually composed of multiple interrelated steps, such as definition problem, data selection, data preprocessing, data transformation, data mining, pattern interpretation, knowledge evaluation, etc. The schematic diagram of data mining process is shown in Figure 1.
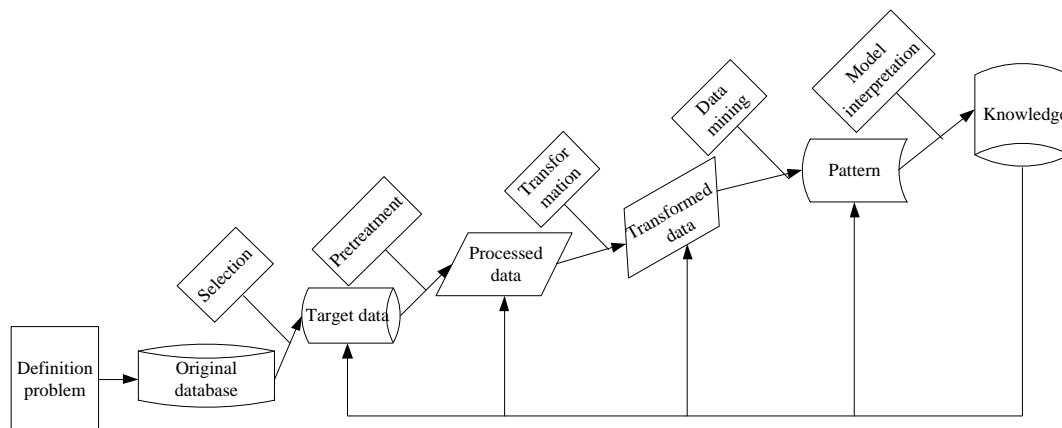

Figure 1. Schematic Diagram of Data Mining Process

As shown in Figure 1, the corresponding steps in the process are described as follows:

(1) Defining problems: accurately finding out the definition of the problem to be mined and clarifying the theme and goal of mining are the cornerstone of the success of data mining. Before preparing useful data, data mining related personnel must clearly understand the target needs of end users, clarify the background knowledge, understand the relevant situation in relevant application fields and the data mining algorithms to be adopted in the process of data mining (Feng & University, 2019). Generally speaking, the final result of data mining is unpredictable, but the problems that need to be explored and mined can be known in advance. The so-called data mining for data mining has a certain blindness, and it is difficult to succeed after all (Xie, QI, Zhang, et al., 2019). Therefore, in the problem definition stage, the close cooperation between data mining personnel and relevant application experts and end users is particularly important.

(2) Data selection: the key of data selection is to determine the target data. According to the needs of end users and under the guidance of experts in relevant application fields, relevant data or samples are extracted from the original database to form a new meaningful target data set. In this process, some database operations are mainly used to process the original data set.

(3) Data preprocessing: further process the data determined in step (2), and check the consistency and integrity of the extracted data set. Remove the useless data deviated from the mining target, clean the noise data, and complete the data type conversion. According to the data characteristics and changes of the selected data set, the vacancy data are deduced and calculated by statistical methods.

(4) Data transformation: reprocess the preprocessed data according to the target task of data mining, so as to eliminate the redundant features in the data. The usual approach is to reduce the number of variables or decision attributes in the data set by using projection or other operations of the database.

(5) Data mining

(a) Algorithm selection: further determine the type of knowledge to be mined through step (1), and then select the corresponding data mining algorithm. Because the data mining algorithms used for different knowledge types are different, such as association rules, classification, clustering, summary, etc., the expression forms of their results are completely different (Li, 2019). After determining the algorithm, selecting the appropriate data mining algorithm also includes selecting the appropriate model and parameters. There are two methods to select the algorithm: one is to select the corresponding algorithm according to the different data characteristics. The other is based on the needs of users. Some users need descriptive results, while others want to get results with as high prediction accuracy as possible.

(b) Establish the analysis model: use the previously selected algorithm to collect data from the preprocessed target data set. Extract the knowledge and information that users are interested in and show it to users in an intuitive and clear form (Lin & Raza, 2019). This is a very important link in the whole process of data mining.

(6) Pattern interpretation: explain the knowledge and information discovered in step (5). Remove redundant or irrelevant knowledge and information found by machine or user evaluation. In addition, if

60

the knowledge and information can not meet the needs of users, it is necessary to find out the reason and return to some previous processing steps for repeated extraction. Such as re selecting data, changing a data mining algorithm or changing some parameter values in the data mining algorithm, so as to mine more effective and accurate knowledge.

(7) Knowledge evaluation: this step is to present the previously mined knowledge and information in a more intuitive form that users can understand. Because the knowledge and information found through the above steps are ultimately user-oriented. Therefore, the most intuitive way that users can understand should be taken as the final form of knowledge and information (Hu, 2019). Therefore, the work of data mining also includes visual processing of knowledge and information and rule transformation.

3. Evaluation Method of Effectiveness of College Students' Employment Guidance Strategy Based on Data Mining

3.1 Constructing the Evaluation Index System of the Effectiveness of College Students' Employment Guidance Strategy

Comprehensive evaluation is the information synthesis of each evaluation index. Whether the evaluation results are objective and accurate depends on whether the information of each evaluation index is accurate, comprehensive and representative. The selection of evaluation index and the design of evaluation index system directly affect the conclusion of comprehensive evaluation. Therefore, what index is selected to describe the evaluated things is the first consideration in multi index comprehensive evaluation (Liao & University, 2019). From the current practice of comprehensive evaluation, many people do not pay attention to this key link, and the randomness in the design of index system is more prominent. The number of evaluation indicators in the evaluation index system is not the more the better, nor is the less the better. Too many, in fact, there will be some repetitive indicators, which will interfere with the evaluation results. Too few, the selected indicators may lack sufficient representativeness, which will lead to one sidedness of the results. Each indicator reflects some information of the evaluation object from a certain aspect it represents. Therefore, how to correctly and scientifically use this information is the first problem to be handled in comprehensive evaluation. Obviously, the selection of evaluation indicators is related to the professional knowledge involved in specific problems, the means we can investigate and obtain, and the experience in comprehensive evaluation. Using the method of qualitative analysis, the pre selection of evaluation indicators can not be simply selected at will, but must follow certain principles. There are mainly the following 10 principles: guiding principle, hierarchical principle, system integrity principle, relative independence principle, comprehensive index priority principle, comparability principle, measurability principle, stability principle, intuitiveness principle and operability principle. Following the above basic principles, specific evaluation indexes can be preliminarily selected through qualitative analysis, and then a preselected evaluation index system can be established. The index system of college students' employment guidance strategy constructed is shown in Table 1.

61

Table 1. Effectiveness Evaluation Index System of College Students' Employment Guidance Strategy

Target layer	Criterion layer		Index layer	
Employment level A	B1	Specialty setting and demand	C1	Degree of fit between specialty setting and social needs
			C2	Signing rate at the end of each school year
			C3	Enrollment rate at the end of each school year
			C4	Average starting salary per employment
	B2	Infrastructure and team building	C5	Per student expenditure per academic year
			C6	Number of employees per 100 students
			C7	Number of computers per 100 students
			C8	Copies of employment information per 100 students
			C9	Number of papers published per employee per academic year
	B3	Market organization and effect	C10	Number of employment information provided by colleges and departments in each academic year
			C11	Number of employment practice bases signed by colleges and departments in each academic year
			C12	The college and department hold special job fairs every academic year
			C13	Number of students' intentions to participate in special job fairs in each academic year
	B4	Employment guidance and services	C14	Number of employment policy lectures held by students of colleges and departments in each academic year
			C15	Proportion of students taking employment guidance courses
			C16	Hours of employment consulting services per academic year

Combined with Table 1, it is the evaluation index system of the effectiveness of college students' employment guidance strategy.

3.2 Calculate the Evaluation Index Weight of the Effectiveness of College Students' Employment Guidance Strategy

On the basis of constructing the evaluation index system of the effectiveness of college students' employment guidance strategies, calculate the weight of the evaluation index of the effectiveness of college students' employment guidance strategies. This paper adopts the method of expert scoring, in which the degree of compliance between professional setting and social needs, the signing rate at the end of each school year, the average student funds per academic year, the number of employment

personnel per 100 students and the number of employment policy lectures per department student per academic year are selected as the key indicators of this evaluation, and their weights are calculated.

(1) Degree of fit between specialty setting and social needs. This index can be expressed by the ratio of the number of graduates of a certain specialty to the number of graduates required by the society, which reflects the adaptability of the specialty setting of colleges and universities to the social needs. Calculation method:

$$I = A / B \times 100\% \tag{1}$$

In formula (1), I refers to the degree of compliance between professional setting and social needs; A is the number of graduates of a major; B is the number of graduates required by the society.

(2) The signing rate at the end of each school year refers to the ratio of formal signing and flexible employment of colleges and departments by the end of each school year. The larger the index, the higher the employment level of colleges and departments. Calculation method:

$$r = (s + c) / Y \times 100\% \tag{2}$$

In formula (2), r refers to the signing rate at the end of each school year; s is the number of formal signatories; c is the number of flexible employment; Y is the number of college graduates.

(3) The per student expenditure of each academic year refers to the ratio of the expenditure for employment of colleges and departments to the number of graduates in each academic year. It reflects the importance attached by colleges and departments to the employment of graduates. The higher the proportion, the greater the investment of colleges and departments in employment and the higher the degree of attention. Calculation method:

$$q = \frac{u}{e} \tag{3}$$

In formula (3), q refers to the average student expenditure per academic year (yuan / person); u refers to the employment funds of colleges and departments in each academic year; e is the number of graduates.

(4) The number of employed persons per 100 students refers to the proportion of college and department personnel engaged in employment work specially or part-time to the number of college and department graduates. Calculation method:

$$g = b / 100h \tag{4}$$

In formula (4), g refers to the number of employed persons per 100 students (person); b is the number of personnel engaged in employment in Colleges and departments; h is the number of college graduates.

(5) The number of employment policy lectures held by students of colleges and departments in each academic year reflects the employment guidance and services provided by colleges and departments to students. Calculation method:

$$t = \frac{p}{m} \tag{5}$$

63

In formula (5), t refers to the number of employment policy lectures held by students of colleges and departments in each academic year; p indicates the number of employment policy lectures held by colleges and departments in each academic year; m is the number of students in the Department. Thus, the weight of the evaluation index of the effectiveness of college students' employment guidance strategy is obtained.

3.3 Effectiveness of Employment Guidance Strategies for College Students Based on Data Mining

Based on the above calculated evaluation index weight of the effectiveness of college students' employment guidance strategy, judge the effectiveness of college students' employment guidance strategy. However, due to various reasons, the data in the original data set is often incomplete, which may contain some false information or wrong information. In addition, some data may be meaningless for data mining analysis, and so on. Therefore, before data mining, it is often necessary to do appropriate data preprocessing for the original data, otherwise unexpected error results may occur. In view of the characteristics that the data set attributes after mining, analysis and preprocessing of employment information data are all classification attributes, this experiment plans to use predictive classification algorithm to construct classification model for analysis. Classification is an important data mining method. Its purpose is to build a classification model (often called classifier) according to the characteristics of sample data. The model can map some new samples to a given category. The classification methods mainly include decision tree, k-nearest neighbor method, support vector machine, Bayesian method and neural network method. Because the decision tree is easy to understand and implement, its classification model is easy to test and evaluate, and has high reliability. In addition, the decision tree can mine feasible and effective patterns from large data sets in a short time. Therefore, the decision tree method is used to construct the classification model.

After the above analysis, this paper intends to use ID3 algorithm and improved ID3 algorithm to construct the corresponding decision tree classification model for the same data set. The results of each model are compared, and then a classification model with good classification effect is found for the data mining analysis of employment information data set, and the corresponding results are obtained, so as to provide decision support for college students' employment guidance.

3.4 Evaluation on the Effectiveness of College Students' Employment Guidance Strategy

This paper tests the consistency according to the effectiveness data of college students' employment guidance strategies. Set explicit variables that can be used for direct measurement to generate dynamic evaluation formula. Get the strong and weak relationship between each evaluation index and the effectiveness of the final college students' employment guidance strategy. The effectiveness evaluation formula of college students' employment guidance strategy based on data mining is shown in formula (6).

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix} = \begin{bmatrix} 0 \\ c_1 \\ c_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ c_5 \\ c_6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ c_{14} \\ 0 \\ 0 \end{bmatrix} = \alpha \tag{6}$$

Formula (6) is the evaluation result of the effectiveness of college students' employment guidance strategy based on data mining. Combined with the effectiveness evaluation system of college students' employment guidance strategies and corresponding index factors, each factor adopts the Crete 5-level positive scoring system in the survey. So far, the effectiveness evaluation of college students' employment guidance strategy has been realized.

4. Experiment

4.1 Experimental Preparation

The research data of this paper comes from the employment information database of the employment guidance center of a university in Chongqing, mainly based on the employment information data of a certain session of graduates. The original dataset is an excel format file containing 4927 records. The attribute fields mainly include college, student number, graduation destination, company name, company affiliation, company address, company nature, employment status, remarks, company, dispatch certificate number, file processing method, gender, educational background, major, teacher logo, training method, school system, family, political outlook, enrollment time, graduation time, original company name, place of origin, candidate number, file receiving unit, QQ number and contact number. It takes the effectiveness of guidance strategy as the evaluation content by designing an example analysis. Firstly, it is used to evaluate the effectiveness of guidance strategies. The confidence is tested and recorded by Kinect software, and it is set as the experimental group. Then use the traditional methods to evaluate the effectiveness of college students' employment guidance strategies. Similarly, test and evaluate the confidence through Kinect software, record it, and set it as the control group. It can be seen that the main content of this experiment is to test the evaluation confidence of the two evaluation methods. The higher the evaluation confidence value, the higher the evaluation accuracy of the evaluation method. Through 6 experiments, the experimental data were recorded according to the evaluation confidence measured in the experiment.

4.2 Analysis and Conclusion of Experimental Results

Compare the evaluation confidence under the two evaluation methods, and the comparison results of evaluation confidence are shown in Table 2 below.

Table 2. Comparison Results of Evaluation Confidence

Number of experiments	Evaluation confidence of experimental group	Evaluation confidence of control group
1	99.85%	80.25%
2	99.82%	80.36%
3	99.80%	80.74%
4	99.81%	80.27%
5	99.82%	80.91%
6	99.79%	80.05%

As shown in Table 2, the confidence of the evaluation method designed in this paper is significantly higher than that of the control group, which has obvious advantages.

5. Conclusion

There are still many meaningful topics about the application of data mining technology in college students' employment guidance and the research of decision tree classification algorithm. Next, the further work directions are as follows: Although the improved ID3 algorithm overcomes the disadvantage that the original ID3 algorithm prefers attributes with a large number of values to a certain extent. However, it is highly targeted and can not be widely applied to many fields. How to solve this problem and establish a more perfect decision tree classification model needs further research. This paper only uses a single employment information data of college students, but does not involve some data similar to personal basic information and achievement information of relevant college students. In fact, these data have an impact on employment. Therefore, in the next step, building a more complete data warehouse, data mining and analyzing the impact of employment factors is also a topic worthy of research. For the same data set, different mining topics can be found, and the preprocessing methods and modeling algorithms will be different. Next, using other data mining algorithms such as association rules to mine the employment information data set is also worth studying.

References

Feng, J., & University, H. F. (2019). Innovative Ways of College Students' Employment Guidance under the New Media Environment. *The Guide of Science & Education*, *09*, 174-175.

Hu, R. A. (2019). Employment Situation of College Students and Guidance Strategies of Counselors. *Journal of Hengshui University*, *21*(01), 110-114.

Justin, M. P., Ellen, J. Y. (2019). Goodwill without Guidance: College Student Outreach Practitioner Training. *Journal of chemical education*, *96*(3), 414-422. https://doi.org/10.1021/acs.jchemed. 8b00882

Li, J. (2019). On the Improvement of Employment Guidance Service System for College Students. *Journal of Jiangsu Institute of Commerce*, *03*, 80-82.

Liao, S. R., & University, L. N. (2019). Suggestions on Dehydration of Employment Guidance Course in Colleges and Universities. *Science & Technology Vision*, *01*, 143-146.

Lin, B. Q., & Raza, Y. M. (2019). Improvements in the decision making for Cleaner Production by data mining: Case study of vanadium extraction industry using weak acid leaching process. *Journal of Cleaner Production*, *16*(5), 417-428.

Pan, M., & Center, E. G. (2019). Analysis of the Role of Career Planning in College Students' Employment Guidance. *Hetao College Forum*, *33*(1), 89-96.

Xie, W. Y., QI, G. C., Zhang P., et al. (2019). Exploration on the Reform of College Students' Career Development and Employment Guidance Course with the Goal of Curriculum Ideological and Political Education. *Science & Technology Vision*, *33*, 38-39.

Zhou, S. Z., Dan, L. I., Wen-Li, Y. U., et al. (2019). Research on the Employment Guidance Path for College Students. *Management & Technology of SME*, *07*, 118-119.

Zhu, M. (2019). Research on Ideological and Political Education in Employment Guidance for College Students. *The Theory and Practice of Innovation and Entrepreneurship*, *02*(02), 193-194.