# Research on the Application of Information Technology in Psychological Education of College Students

Qiudi Xing[1(✉)] and Xuefang Chen[2]

[1] Fuyang Preschool Teachers College, Fuyang 236015, Anhui, China
[2] Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310028, Zhejiang, China

**Abstract.** The mental health level of college students not only directly affects their own growth, but also affects the stability of the campus, and then affects the social harmony and the improvement of the quality of the whole people. Therefore, the psychological problems of college students have aroused widespread concern in the society. The psychological intervention of college students has become a hot spot in the research of college students' mental health. With the development and maturity of data mining technology and the successful application in all walks of life, the advantages of this technology in discovering hidden rules or patterns in data are incomparable with other technologies. This paper collects and reviews a large number of relevant literature, and discusses the application of data mining technology and data analysis of college students' psychological problems.

**Keywords:** Data mining · Data preprocessing · Decision tree · Psychological problems of college students

## 1 Introduction

Data mining, also known as knowledge discovery, is to discover hidden mineral resources knowledge from assive data. It is a comprehensive application of statistics, artificial intelligence, database and other technologies. Using the tools and methods of data mining, valuable knowledge can be extracted from the rich data, otherwise the vast "data ocean" will become the "data grave" of lack of information [1].

Some research shows that most of the students' weariness, dropout, suicide and hurting others are caused by mental health problems, and the number of students with poor mental health has been on the rise. According to a survey of 126000 college students in China, 20.3% of them have psychological problems, mainly manifested as terror, anxiety, obsessive-compulsive disorder, depression and neurasthenia.

According to the survey, the current college students' psychological problems mainly include three aspects: psychological confusion, psychological obstacles and psychological diseases. Among them, the students with psychological confusion are more common.


M. A. Jan and F. Khan (Eds.): BigIoT-EDU 2021, LNICST 392, pp. 487–492, 2021.
https://doi.org/10.1007/978-3-030-87903-7_60

Although they are mild psychological problems, they do not affect their health. However, if the minor problems can not be adjusted and dredged in time, they will develop into mental disorders. If psychological barriers are not timely adjusted and treated, they will develop into mental diseases. Mental illness will seriously affect their physical and mental health and all-round development, and even lead to malignant events.

## 2 Data Mining Technology

### 2.1 Cluster Analysis

Clustering is to classify data objects into several classes or clusters according to the principle of "maximizing the similarity within a class and minimizing the similarity between classes". The similarity of objects in the same class is very high, but the differences of objects in different classes are very high. Clustering analysis is the process of classifying according to some similarity of data and analyzing the formed multiple classes. Clustering methods mainly include hierarchical method, partition method, grid based method, density based method, model-based method, etc.

### 2.2 Classification of Data Mining System

Data mining technology comes from many disciplines, which will have an impact on data mining, as shown in Fig. 1.

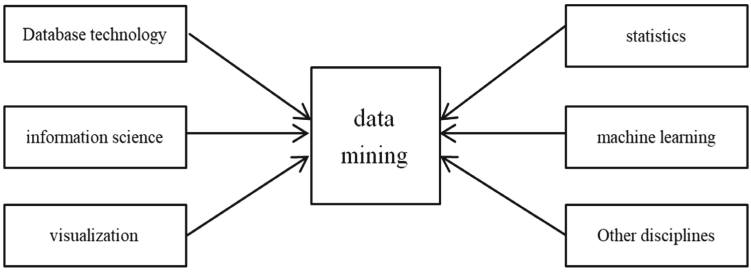


**Fig. 1.** Multiple disciplines influence data mining

Because data mining is an interdisciplinary subject, data mining will produce many different types of data mining systems [2]. Accurate classification of data mining system can provide scientific basis for users to choose the most suitable data mining system.

### 2.3 Decision Tree

In data mining, decision tree is mainly used for classification. Each node represents the distribution method of the top-level tree, and each node represents the distribution method of a class, and each node represents the distribution of a class. According to different characteristics, the decision tree uses tree structure to represent the classification, which

is used as the basis for generating rules. The main advantages of decision tree are simple description, fast classification speed, easy to understand the generated model, and high precision. It is widely used in all kinds of data mining systems. Its main drawback is that it is difficult to construct a decision tree based on multiple variables (Fig. 2).

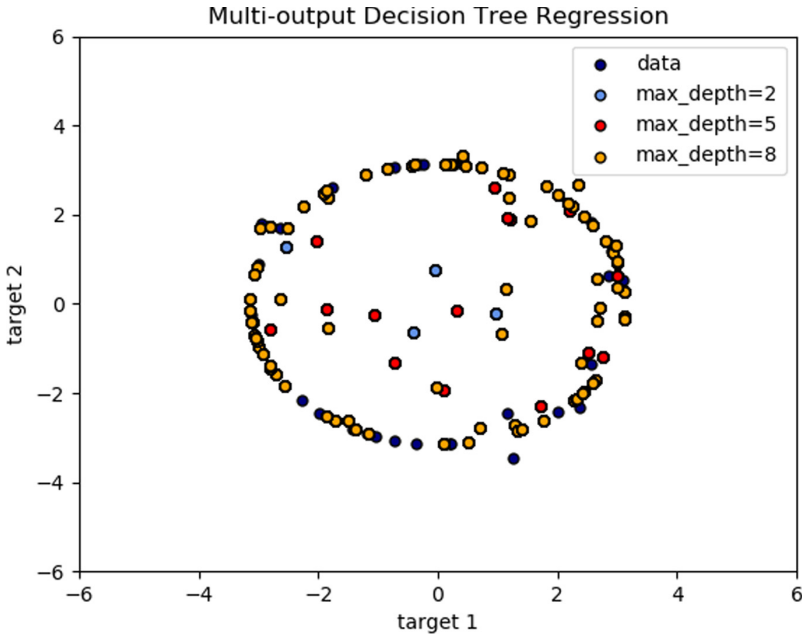


**Fig. 2.** Decision tree

## 3 Application of Data Mining in the Analysis of College Students' Psychological Problems

Attribute selection measures, also known as splitting rules, determine how to split samples on a given node. Here are two popular attribute selection metrics: information gain and gain rate [3].

1. Information gain.

Let node n store all samples of data partition D. The expected information required for the classification of samples in D is given by the following formula:

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i) \tag{1}$$

Where $p_i$ is the probability that any sample in D belongs to $C_i$.

The expected information required for sample classification of D based on attribute A can be obtained as follows:

$$Info_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} \times Info(D_j) \tag{2}$$

Where $\frac{|D_j|}{|D|}$ is the weight of a subset of value $a_j$ on attribute A.

Classification is actually to extract information from the system to reduce the confusion of the system, so as to make the system more regular, more orderly and more organized. The more chaotic the system, the greater the entropy. Obviously, the optimal splitting scheme is the splitting scheme with the largest entropy reduction.

In this chapter, according to the requirements of decision-making analysis of college students' mental health education, the whole process of classification and mining of college students' psychological problems is fully realized. The first is the determination of mining objects and data mining objectives: decision tree model of whether students have interpersonal sensitivity symptoms or not. Then preprocess the data to get the training sample set. According to the characteristics of the training sample set, C4.5 algorithm of decision tree is selected to construct the decision tree model of whether students have interpersonal sensitivity symptoms and prune it. Then the classification rules are extracted from the decision tree model and analyzed. Finally, the accuracy of the model is evaluated. This paper also compares the original tree with the pruned tree in terms of scale, extracted classification rules and classification accuracy. The conclusion is that the pruned decision tree model is simpler, easier to understand and has higher classification efficiency than the directly generated decision tree.

## 4 Design and Implementation of College Students' Psychological Data Management System

The collection and analysis of students' psychological evaluation data is a necessary basic work for colleges and universities to carry out mental health education. With the rapid increase of enrollment and the improvement of the connotation of psychological data analysis, more and more psychological data need to be analyzed and processed more deeply. Although some well-known and powerful psychological assessment software has appeared in China, these software are expensive and have not applied data mining technology. Therefore, it is necessary to develop a college students' psychological data management system based on data mining technology and BS mode, so as to improve the work efficiency of psychological evaluation data collection and increase the depth of psychological data analysis [4].

The student function module is oriented to students, which mainly realizes the collection of students' basic information and psychological evaluation information, and establishes psychological files for students. After students enter the system, the system generates a dynamic psychological file for them. Students can modify their personal password by modifying the password sub module; modify the personal basic information by modifying the basic information sub module; at the same time, the system can

collect the students' basic information; through the psychological evaluation sub module, online psychological self-assessment can be realized, and the evaluation results can be viewed, and the system can collect the psychological evaluation data (Fig. 3).

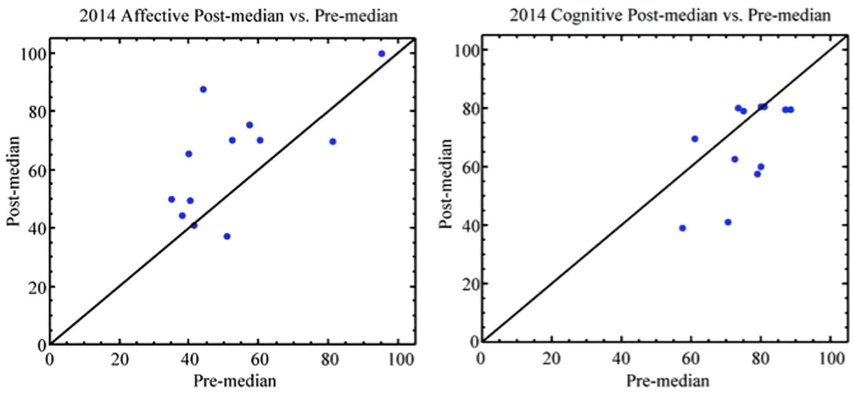


**Fig. 3.** Design and implementation of college students' psychological data management system

The administrator function module is used to retrieve the administrator, mainly to modify personal password, customize psychological questionnaire, information management, data mining, psychological prediction and other functions. The administrator realizes the management and import of the questionnaire through the self-defined psychological questionnaire sub module; realizes the self-determination of the students' basic attributes, the management of the students' basic information and the management of the psychological evaluation information through the information management sub module; realizes the data preparation, the generation of the decision tree and the generation of the classification rules through the data mining sub module; through the psychological prediction sub module, students' psychological problems can be predicted.

## 5 Conclusion

This paper analyzes the key technologies of data mining, deeply studies the classification in data mining, and analyzes and compares several commonly used classification algorithms, which provides the basis for the application of decision tree algorithm in the analysis of college students' psychological problems. The classification rules are extracted from the optimal decision tree model, which provides an important reference for the school psychological consultation work, and realizes the classification and prediction of new data by using the model, which provides a scientific basis for the early warning and intervention research of college students' psychological problems.

**Acknowledgements.** This is the achievement of Anhui Province top talent cultivation funding project excellent young backbone talents domestic visit project (NO. gxgnfx2019111).

## References

1. Guojun, M., Lijuan, D., Shi, W.: Principle and Algorithm of Data Mining. Publishing House of Tsinghua University, Beijing (2005)
2. Peng, W.: Research on the Application of Data Mining in the Identification of College Students with Difficulties. Northeast Normal University, Changchun (2011)
3. Weiping, L.: Application of Data Mining Technology in Secondary Vocational Enrollment Information System. Northwest Normal University, Lanzhou (2008)
4. He, G.D.: Application of data mining in college students' psychological problems. Wirel. Internet Technol. **2**, 196–197 (2013)