



A Follow-Up Investigation on the Measurement Data of College Students' Mental Health Effectiveness Based on Big Data Analysis

Deming Qiu^(✉)

Harbin Finance University, Harbin 150030, 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. College Students' psychological intervention has become a hot topic in college students' mental health research. With the development and maturity of big data analysis technology and its successful application in all walks of life, this technology has incomparable advantages in discovering hidden rules or patterns in data. In this paper, the effectiveness of mental health measurement data tracking research.

Keywords: Big data analysis · Mental health · Effectiveness · Tracking analysis

1 Introduction

Data mining, also known as knowledge discovery, is to discover hidden mineral resources knowledge from massive 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.

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 [1]. 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. 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 [2–4]. 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 [5]. 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.

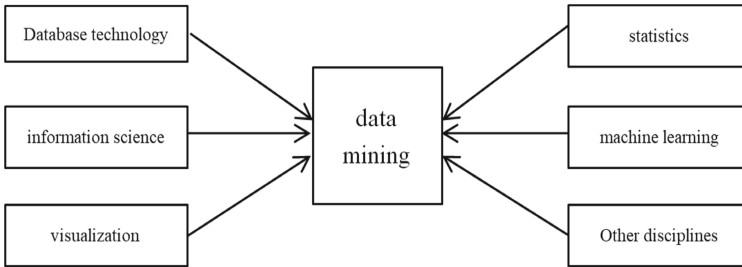


Fig. 1. Multiple disciplines influence data mining

Because data mining is an interdisciplinary subject, data mining will produce many different types of data mining systems [6]. 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 [7–9]. 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.

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.

1. Information gain.

Let node n store all samples of data partition D [10]. The expected information required for the classification of samples in D is given by the following formula:

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

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:

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

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 [6, 7]. 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 [8]. 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 Simulation for 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 [9, 10]. 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.

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 (see Fig. 2), 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.

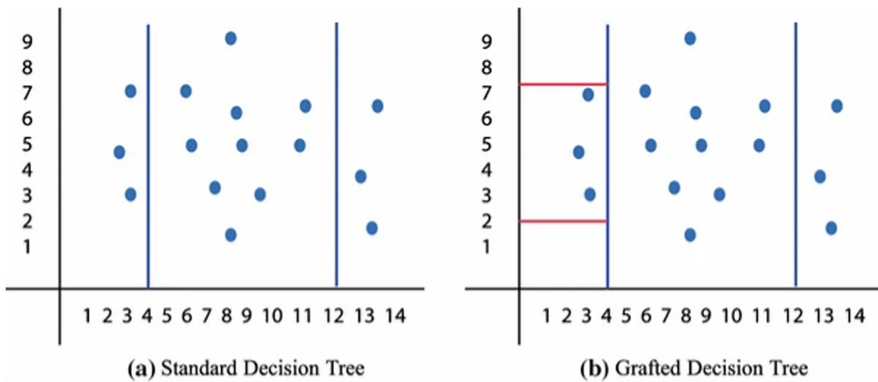


Fig. 2. Decision tree with grafted and standard

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 [11]. 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 (see Fig. 3).

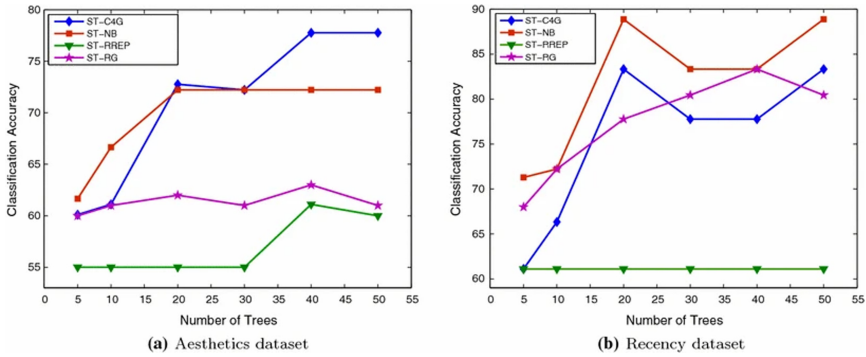


Fig. 3. Compare with the different dataset for the decision tree

5 Phase Difference Detection Technology of CMF System

5.1 Introduce

According to the basic principle of Coriolis force measurement, the output signals of the left and right vibration detectors on the U-tube of CMF are two sinusoidal signals [12, 13]. When there is flow in the measuring tube, because of the Coriolis force, the phase difference between the two signals is very small and increases with the increase of the flow in the measuring tube. The mass flow rate of fluid has a linear relationship with the phase difference. In order to get the accurate flow rate, the accurate phase difference must be measured. Phase difference detection technology is one of the key technologies of CMF system. Which phase difference detection technology is used directly determines the composition of the system and the accuracy of the instrument.

5.2 Common Phase Difference Detection Technology

(1) After adding, subtracting or multiplying two sinusoidal signals of the same frequency by vector method and phase multiplication, the sine or cosine value of the phase difference can be obtained through a low-pass filter, and the value of the phase difference can be obtained after anti sine or anti cosine operation [14]. According to the different operation methods, they are called vector method and phase multiplication. These two similar methods are suitable for the measurement of high frequency sinusoidal signal with single spectrum and large phase difference. If the frequency of the signal is low, then the low-pass filtering is difficult to achieve very accurate, and if the phase difference is small, the cosine value is close to 1, As shown in Fig. 4 and the sensitivity is very low. In addition, the signal must have a single spectrum, noise and interference have a great impact on the measurement results.

(2) In practice, the most commonly used detection methods are zero crossing phase detection method and improved two-way zero crossing phase detection method. The basic method is to shape the two signals after zero crossing comparison, then get the phase pulse signal by phase discrimination, and finally get the phase difference by high frequency counting. There is also a process of offset pulse square wave correction in

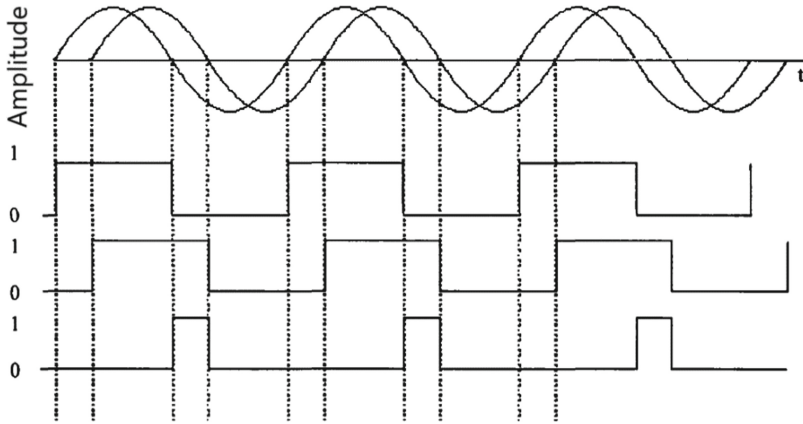


Fig. 4. Waveform diagram of zero crossing phase discrimination

the two-way zero crossing phase detection. The two-way zero crossing phase detection method can well solve the influence of temperature drift and comparator offset on the measurement error [15]. However, the noise and interference also have a great influence on the measurement error, and when the phase difference is small, the phase difference accuracy will be reduced due to the influence of counting clock accuracy.

(3) There is another phase detection method based on least square method (LMS). Through some matrix transformation and calculation, the amplitude and phase angle of two signals can be obtained simultaneously. The calculation of this method is more complex, and it can get high accuracy ($<1\%$) when the noise and interference are small and the signal spectrum is single. However, the measurement accuracy will deteriorate sharply with the distortion of signal waveform and the increase of noise and interference, and the actual implementation is also more complex, which requires a higher processor [16–19]. These methods have their own advantages and disadvantages, but they can not overcome the influence of signal zero drift and waveform distortion on the measurement error.

6 Research on CMF Signal Processing Method based on DPLL

6.1 Introduction of Application Technology

In practical application, the signal is not an ideal sine signal because of various harmonic interference and noise in industrial field. The frequency of the output signal of Coriolis mass flowmeter sensor falls into the frequency range of many industrial noises, and the amplitude of the output signal of the sensor is relatively small, which is not significantly greater than the amplitude of the noise signal in many cases, which limits the sensitivity of the flowmeter, makes it difficult to obtain the useful signal, and leads to the reduction of the measurement accuracy [20]. In addition, the vibration frequency of the U-tube varies with the fluid density, so that it is not equal to the driving frequency, resulting in the frequency change of the output signal of the sensor. In this way, the spectrum of the

output signal of the sensor can not be a single spectral line corresponding to the sinusoidal signal, and its composition will become very complex. At present, the commonly used zero crossing equal phase measurement methods are based on amplification, filtering, shaping and counting. The use of analog filter will inevitably change the amplitude and phase of the sensor output signal, which is undesirable, because the phase delay between the two signals is used to calculate the basic information of fluid characteristics [21]. This leads to the final measurement of the phase difference of the sine useful signal output by the synthetic wave rather than the sensor, and the measurement results are easily affected by the waves and random noise, resulting in measurement error.

6.2 Signal Processing based on DPLL

6.2.1 General Overview

Phase locked loop (PLL) is a closed-loop phase control system. Its theoretical basis is automatic control theory. The theory of synchronous detection proposed by dbellesize in 1932 first published the description of PLL. Since the 1960s, PLL began to be digitized, and some digital loops and integrated loops appeared [22]. The common characteristics of these digital PLLs are high reliability, stable performance and easy integration. Now, PLL has become an indispensable part of communication, radar, navigation, deep space exploration, electronic instruments and other equipment. The reason why it can be widely used is that it has a unique narrow-band tracking performance, and can complete the tasks of frequency synthesis, modulation and demodulation, synchronous extraction, velocity and moment measurement, micro frequency conversion and so on. Narrowband tracking performance includes two meanings: one is the tracking function of input signal; the other is the narrowband filtering performance of input noise.

The DPLL consists of digital phase detector, digital loop filter and digital controlled oscillator. According to the type of digital phase detector used in the loop, digital phase locked loops can be divided into four types: lead lag digital phase locked loop (ll-dpll), trigger digital phase locked loop (ff-dpll), zero crossing detection digital phase locked loop (zc-dpll) and Nyquist digital phase locked loop (nr-dpll).

6.2.2 Down Sampling Filter

In this system, although the required useful signal frequency is very low, only about 100 Hz, but the signal is complex, high-frequency noise and interference have great influence. In order to eliminate the influence of noise and interference after AD sampling, the sampling rate of ad is much higher than twice of the highest frequency of the useful signal, so we need to pay attention to the spectrum aliasing of the sampled signal. In the traditional down sampling system, in order to avoid spectrum aliasing, the input sampling data must be de aliased first, so that the highest spectrum component contained in the signal is lower than the Nyquist frequency after sampling, $f/(2 m)$. In this system, the fundamental frequency component of the signal after digital mixing has moved to the zero position, and the harmonic interference has been eliminated by integral comb filtering, so the sampling frequency can be reduced very low [23]. The anti aliasing filtering is not needed before sampling.

6.3 Simulation Analysis Data

Before the simulation analysis, it is necessary to obtain the parameters of all filters in the whole DPLL loop, in addition to the frequency offset $\Delta - y$ data table. It is very easy to get the filter parameters by using the computer-aided design method under the MATLAB platform. We will not repeat here, but only discuss the acquisition of 4-y data table in detail. After that, we will simulate and analyze the data of several different cases to verify and analyze the performance of the whole DPLL system. In the simulation process, the most important thing is the evaluation of frequency tracking and the calculation accuracy of the final phase difference results, which will be verified and analyzed later. All the work of this part is completed in the MATLAB simulation software platform [13, 14]. The hardware system based on CPLD and MCU overcomes the contradiction between the low processing speed of MCU and the high speed of data acquisition, and greatly eliminates the redundant data collected by AD converter, which reduces the volume of the system and the difficulty of MCU software system development. The MCU system completes the final data processing, calculates the phase difference and mass flow [24–26]. Because the data processing is relatively simple, the software system of single chip microcomputer is designed with assembly language, which realizes many functions including data processing, display and communication modules, and enhances the data processing and auxiliary functions of CMF system, especially the communication between CMF and PC.

7 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. A New Characteristics of the Ideological Development of Poverty University Students in Universities and Colleges by the Provincial Basic Business Fee Project of Harbin Institute of Finance (2018-KYWF-E017).

References

1. Mao, G., Duan, L., Wang, S.: Principle and Algorithm of Data Mining. Publishing House of Tsinghua University, Beijing (2005)
2. Wang, P.: Research on the Application Of Data Mining in the Identification Of College Students with Difficulties. Northeast Normal University, Changchun (2011)
3. Li, W.: Application of Data Mining Technology in Secondary Vocational Enrollment Information System. Northwest Normal University, Lanzhou (2008)

4. Le, G.D.: Application of data mining in college students' psychological problems. *Wireless Internet Technol.* **2**, 196–197 (2013)
5. Zhou, H., Jiang, Z., Wang, D.: *Maneuvering Target Tracking*. National Defense Industry Press, Beijing (1991)
6. He, M., Wang, Z., Zhu, J.: Real time field elimination method for multi-sensor target tracking. *Acta Astronaut. Sin.* **23**(6), 34–37 (2002)
7. Zhu, J.B.: Data fusion technology of incomplete trajectory measurement. *Sci. Bull.* **45**(20) (2000)
8. Sun, H., Ding, X., Chen, X.: Real time field culling method based on data fusion for range optical measurement data. *J. Command Technol. Coll.* **10**(4) (1999)
9. Zhu, X., Han, R., Yang, R.: Outlier elimination method based on fuzzy prediction system. *Syst. Eng. Electron. Technol.* **28**(3) (2006)
10. Xu, Z.: *Principle and Application of Total Station*. PLA Press, Beijing (2003)
11. Luo, Y.: *Application of Measuring Robot in Fast Feed Dynamic Tracking Measurement*. University of Information Engineering, Zhengzhou (2003)
12. Yang, Y.: *Adaptive Dynamic Navigation and Positioning*. Surveying and Mapping Press, Beijing (2006)
13. Ni, W.: *Phase Difference Measurement Method and Device for Coriolis Mass Flowmeter*. Institute of Automation, Hefei University of Technology, Hefei (1999)
14. Xu, K., Lv, X., Chen, R.: Research on key technology of DFT method in signal processing of Coriolis flowmeter. *Ind. Instrument. Autom. Device* (5), 7–10 (1998)
15. Xu, K., et al.: Development of excitation circuit for Coriolis mass flowmeter. *J. Hefei Univ. Technol. (Nat. Sci. Ed.)* **23**(1), 37–40 (2000)
16. He, H., Liu, Q., Xu, Y.: Hardware circuit analysis of remote transmitter rft9712 for Coriolis mass flowmeter. *Autom. Instrum.* **18**(10), 32–34 (1997)
17. Sara, D., Fateme, D., Marzieh, A.: The relationship of maternal anxiety, positive and negative affect schedule, and fatigue with neonatal psychological health upon childbirth. *Contraception Reprod. Med.* **6**(1) (2021)
18. Hu, Z.: *Research on local government emergency management of public health emergencies*. Anhui University of Finance and Economics (2021)
19. Cheng, K., Xu, J.: Improving the mental health service system for Tibetan adolescents. *China Soc. Sci. J.* 2021-04-01 (005)
20. Sun, Q.: Tolerance is a magic weapon for mental health. *Modern Health Preserv.* **21**(7), 52 (2021)
21. Yang, H.: Problems and measures of moral education of primary school head teachers in the Internet age. *Knowledge Base* **7**, 58–59 (2021)
22. Chen, L.: On the influence of family factors on College Students' mental health and countermeasures. *Knowledge Base* **7**, 177–178 (2021)
23. Zeng, J.: Analyzing the realistic demands and paths of adult mental health education. *Knowledge Base* **7**, 181–182 (2021)
24. Lu, B.: Old things and new play are beneficial to the mental health of the elderly. *Shanxi Elderly* **4**, 64 (2021)
25. Educating people with heart, cultivating wisdom, cultivating heart, moistening life. *Mental Health Educ. Primary Second. Schools* **10**, 2 (2021)
26. He, Y., Yu, G.: The influence of cultural trend and social transformation on the mental health of Chinese teenagers. *Primary Second. School Psychol.*