Eng Int Syst (2019) 4: 167–175

Engineering
Intelligent Systems

Emotional Data Mining and Machine Learning in College Students' Psychological Cognitive Education

Xiaofan Shi*

School of Economics and Management, Beihua University, Jilin, Jilin 132013, China

Emotion analysis method in AI technology can quickly and effectively analyze the emotions in text, and it is an important way to identify the emotional information of students' feedback text. In this paper, the author analyzes the application of emotional data mining and machine learning in college students' psychological cognitive education. Result shows that there is a relationship between Academic Emotion and students' learning achievement. Using artificial intelligence technology to discover Academic Emotion in educational texts can more easily understand students' learning behavior and solve students' psychological problems. Through the use of computer access and mobile terminal access to the two access methods, psychological test system provides a variety of choices for teachers and students. Therefore, applying the mobile terminal to the psychological testing work has exerted the portability advantage of the mobile terminal, so that the psychological testing work can be performed anytime and anywhere, which not only facilitates the teachers and students, but also improves the work efficiency.

Keywords: Emotional data mining; Machine learning; Psychological cognition; Neural network

1. INTRODUCTION

Emotion is closely related to cognition, decision-making, communication and so on. It is of great significance to mine and analyze emotions. In the field of education, researchers pay more attention to academic emotions(Zhang et al., 2017). A large number of studies have shown that academic emotions are closely related to academic achievements. Understanding students' academic emotions is conducive to better teaching intervention and management. Emotional analysis based on artificial intelligence technology has been applied in word-of-mouth analysis, public opinion analysis and other fields(Bi et al., 2015). However, due to the unique application of emotional analysis in the field, it is difficult to achieve better results by directly applying emotional analysis in other fields to academic emotional analysis(Ulukuset al., 2015; Bi & Wang, 2016). At the University stage, there is a lack of strict norms and order, and close management of teachers. They need to make their own decisions, learn by their own constraints, and live by their own abilities. Therefore, college students are experiencing an unacceptable sense of powerlessness and confusion without goals(Jiang et al., 2015). They may have various problems in thought, life, psychology and learning, and their physical and mental health will be affected to varying degrees. One of the basic tasks of individual life is development and adaptation. The key to individual survival and development lies in adapting to society(Cattivell & Sayed, 2011). For teenagers, school adaptation is an important part of their social adaptation.

At present, there are many predictive factors for school adaptation, which factors will play a major role and which time point will play a role, which are the urgent scientific problems

*Corresponding author: Xiaofan Shi, Email:shixiaofan666@126.com

to be solved in practice, which need further investigation(Ling et al., 2016). In this study, machine learning method is used to investigate the impact of many individual and environmental factors on school adaptation at different time points, trying to make a preliminary exploration to solve the above problems(Li et al., 2016). Machine learning is essentially a cross-disciplinary subject, which is popular in today's multi-disciplinary exploration and research. It extensively draws on the achievements of artificial intelligence, probability statistics, computational complexity theory, cybernetics, information theory, philosophy, physiology, neurobiology and other disciplines.

Psychological motivation is the motivation to guide a person to a variety of behavioral activities. Therefore, understanding the individual's motivation and response during psychological testing is valuable for psychological testing(Chitra et al., 2013; Castelli et al., 2017). Zvi et al. performed experiments to observe the effects of different people under different motivations. Their experiments show that when people have anti-motivation in their hearts, the psychological reaction will change significantly, and when people have a cooperative attitude, they will not misjudge. On the basis of this, Elaad added external stimulus conditions to study, and also obtained the same conclusion as Zvi. In 1820, Brown et al. proposed the law of psychology, and the test has a theoretical basis(Andrea et al., 2015; Leiva et al., 2016). At present, the testing technology is more and more diversified, and the emergence of the test scale has brought the psychological testing work to a new level(Mohan & Govardhan,2013). The PHES scale is an internationally recognized scale. PHES is used to diagnose patients with MHE with a sensitivity of 96% and is internationally known as the gold standard. The test contents of the PHES scale are: NCT-A, NCT-B, DST, LTT, SDT (Djenouri et al., 2015). The test scale 15FQ+ is often used as a personality assessment tool in research, for example, Jonna Moutafi of the University of London et al. used it to assess personality. It was used by Sertse Moyo of the University of Stellenbosch to verify the personality structure. Since the 15FQ+ scale contains intellectual attitudes and emotions, the use of 15FQ+ is even broader. For example, Nicolene Barkhuizen and others at Northwestern University in South Africa used it to measure personality in research. Verona Rosemarie Solomon of the Western Cape University studied the predictability of 15FQ+ in job performance. Chloe M. Hammans of Massey University in New Zealand used it to measure personality traits. 15FQ+ is also used in the research work of educational psychology (Bayindir et al., 2016). For example, Saffiya Nagdee of the University of Pretoria uses it to evaluate students' personality in research work. As test techniques become more mature, test conclusions are increasingly recognized as evidence. Some regions of the United States have unconditionally recognized that psychological test conclusions can be used as evidence in cases(Fang et al., 2013; Ebesu et al., 2017). These areas believe that the court can decide whether to use the test results as evidence, whether or not the testee agrees to use the results as evidence before the test. Article 321 of the Japanese Criminal Procedure Law of Japan also stipulates that psychological test results can be used as evidence on the premise of voluntary consent of the parties.

Based on the network platform and based on the B/S structure of the psychological test system for college students, after the teacher is authorized by the identity security certification, the test results of the students' psychological tests can be viewed, so that the students' mental health status can be found in time, and the students in need can be intervened in advance to achieve the prevention-oriented goal. After the student is authorized by the identity security certification, he can choose the test task to conduct psychological test through the psychological test system anytime and anywhere and view his test results. In this way, students can test their mental health status, actively participate in health consultation and receive guidance from psychological teachers, which is of great significance to improve students' self-prevention awareness.

2. RESEARCH METHOD

2.1 ERP Analysis System Design

The ERP analysis system of this subject is actually based on brain electrical amplifier, which consists of two parts: data acquisition and data analysis.

(1) The data acquisition section includes an electrode cap, a lead wire, an amplifier, and the like.

According to the demand analysis of EEG signal acquisition described in the previous section, this topic selects the UE-B type 16-lead brain electric amplifier produced by Beijing Zhongke Xintuo Instrument Co., Ltd. as the acquisition amplifier of the system. The amplifier is connected to the computer via a USB interface and powered via the USB interface. It is fully compliant with the testing requirements of the psychological test system in terms of amplifier parameters. The main parameters of the amplifier are as follows: 16-bit A/D, common mode rejection ratio $\geq 98dB$, time constant is 0.3s, high frequency filtering is 30Hz, input range is $\pm 15mv$, noise is $\leq 2.5\mu Vp-p$, sampling rate is 1000 times / second, isolation is 2500V. The electrode of this amplifier is the $Ag/AgCl$ electrode, and it uses the traditional square wave calibration method.

An additional reason for choosing this EEG amplifier is that the amplifier provides a dynamic link library, and its programming is relatively simple, and it has a small form factor and reasonable price.

The main function of this part is to collect the weak bioelectrical signal, and amplify the signal through the amplifier, and finally transmit it to the computer for further processing through the USB interface.

(2) Data analysis section

This part is a very core part of this topic, and its hardware carrier is a conventional computer system. The main functions are to collect EEG signals from brain electrical amplifiers, store and display EEG signals, implement signal processing algorithms such as ERP feature waveform extraction, and achieve signal matching between signal analysis systems and stimulation systems.

Similar to the hardware and software of the stimulus system, the analysis system development platform still uses

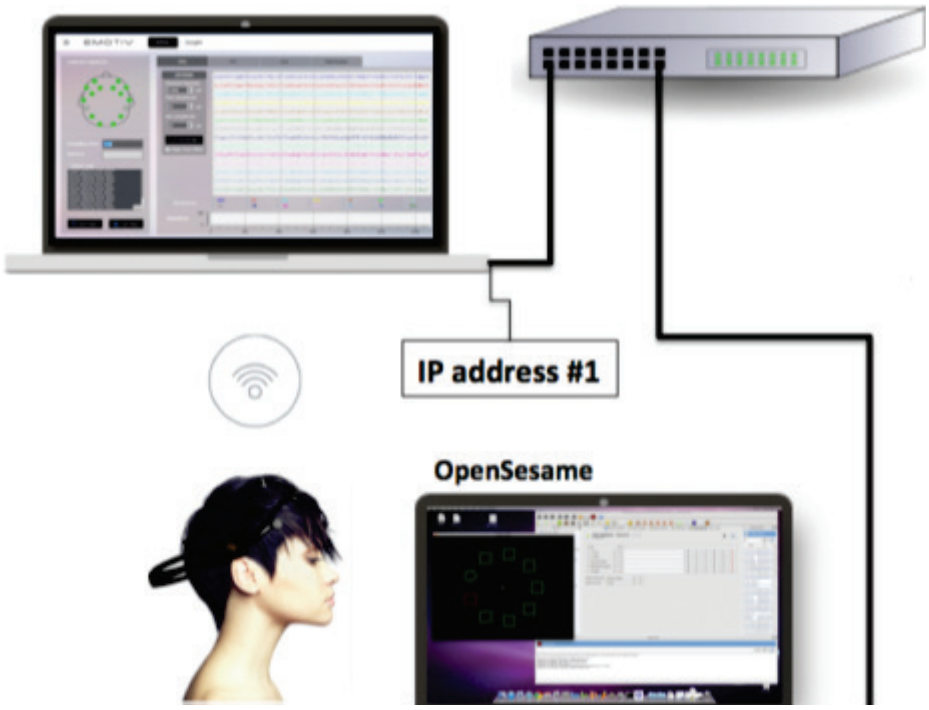

Figure 1 Software system structure.

the Windows XP operating system, but the development environment uses the Visual C++ development environment, and Visual C++ has better ability in the underlying hardware than Visual Basic. At the same time, Visual C++ is an integrated development environment closely connected with Win32. The Visual C++ development system can be used to develop a variety of applications. From the underlying software to the upper-level user-oriented software, Visual C++ can be used for development.

The development environment for selecting Visual C++ as the analysis system has the following points:

(1) Visual C++ has a strong ability to access API communication functions, facilitating data communication between upper-level applications and lower-level hardware.

(2) Visual C++ and Matlab have a good application interface engine, which can facilitate Visual C++ to transfer the signal processing pressure to the obvious Matlab and make up for its lack of signal processing.

(3) Visual C++ has many successful routines for system development, which can speed up the development of this part.

Therefore, the development will choose Visual C++ as a development tool for signal processing systems.

2.2 Software Design

According to the top-down project development principle of software engineering, a prototype of the signal analysis system software system is proposed first, as shown in Figure 1. The signal analysis software system is divided into the following five subsystems by nature.

(1) Data acquisition subsystem: The main task is to receive EEG signals transmitted by peripheral devices from the data interface.

(2) File management subsystem: The main task is to realize the effective management of the file information of the test, including the file output function of data such as file query, editing and EEG waveform.

(3) Signal processing subsystem: The processing subsystem consists of two parts, one is the EEG data preprocessing subsystem, which completes the filtering of the signal; the other is the EEG data analysis subsystem, which realizes the extraction and analysis of the ERP signal.

(4) Data display subsystem: The main task is to display various data information, including EEG waveforms, ERP waveforms, etc.

(5) Data Storage Subsystem: The main task is to implement the storage function of all data in the system.

The following is a detailed description of the design of the subsystem from two aspects.

(1) Subsystem function design

The above is the design of the top layer of the system—the description of the system structure. Next, the system design will be further refined according to the given software structure and requirements analysis. The system design will be introduced through the extraction of functional modules.

(2) Data acquisition subsystem design

Data collection is the first process of the analysis system. Since the focus of this topic is on the software system,

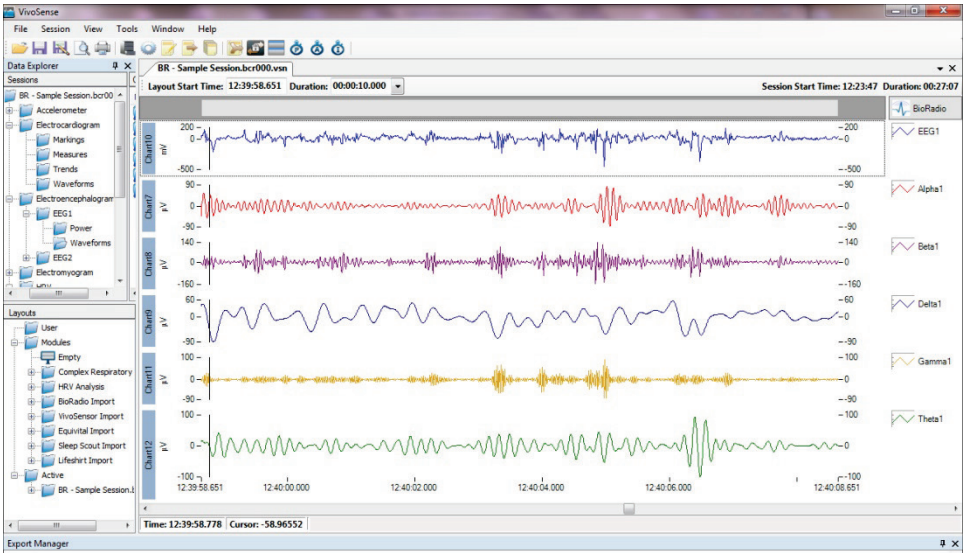

Figure 2 Analysis system interface.

only the acquisition design of the application layer is introduced. First, the system analyzes the various actions to be completed in the data acquisition subsystem, and then uses the dynamic link library provided by the amplifier manufacturer to design the relevant variables and use multi-threading to complete the data collection. Figure 2 shows the waveform interface displayed by the analysis system.

(1) Data collection sub-thread

The function BOOL ReadData(HANDLE hDevice, short*pBuffer, ULONG*nCounts) is called in the thread. This function is used to read the amplifier data, hDevice is the device handle of the amplifier, pBuffer is the buffer address for storing data, and nCounts is the number of data read. After opening the child thread, the buffer size is set to 320 bytes to ensure that the latest data is read within 10ms.

(2) Drawing child thread:

The drawing process is as follows. First, the device environment class CDC in the MFC class library is used to set the brush properties (color, line type, etc.) of the drawing. Second. The buffer data is temporarily stored in short gm_pBufferTemp, and the data size is 16 data, which accounts for 32 bytes. Finally, the MoveTo and LineTo functions are used to draw a line, and only one point of one of the data is drawn at a time, and 16 pieces are drawn and then recycled, thereby completing the graphic output of the 16-conductor data.

3. SYSTEM DESIGN

The module of this system can be roughly divided into two parts: server management module and client test module. Each of these sections can be subdivided into multiple modules. The system function module is shown in Figure 3.

(1) The administrator logs in. In order to increase the confidentiality of the test data, only the administrator can view the data record and perform various operations on the data. Moreover, when the administrator logs in, a password is required, and the password is used for control. (2) The message displays. When testing online, administrators need to understand the various information of online quiz users. The information display system is to help administrators understand all the conditions of the online test. (3) data maintenance. Each test user needs to query or print the test results, which is performed in the data maintenance module. In addition to printing queries, various operations of data such as modification, deletion, etc. can be performed in this module. (4) The test data results sometimes need to be exported for statistical processing. The function of the data import and export system is to export the data into a spreadsheet form or import data of other formats into the system for various processing of the data. (5) User login. Each user must log in with their own username and password when conducting a test, which is beneficial to the confidentiality of user information. (6) Registered. If the user is using the system for the first time, the user must go to register. Registration refers to the registered username (unique) and password. (7) Online test. This is the main part of the computer test to replace the paper and pencil test, the computer presents the test questions one by one, and records the tester's reaction according to the tester's choice. (8) Result query. After the registered user logs in, the test can be performed, and each test result can be viewed at the end of each test. (9) Help. In order to help users to understand the main functions of the system and how to use the system, the system has set up a help module.

Only the administrator can enter the server, and only the administrator has the authority to view the user's test record. For the client, only the tester can view his own test records, and this is for the confidentiality of the data and to ensure the privacy of the tester. In this system, it is realized by administrator login and user login module. Both the administrator and the user login module must use the password

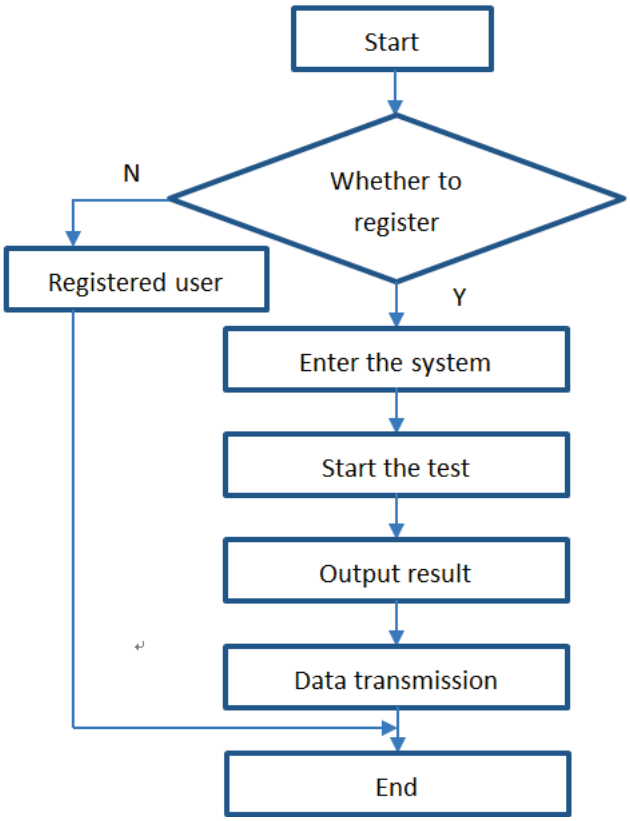

Figure 3 Test flow chart.

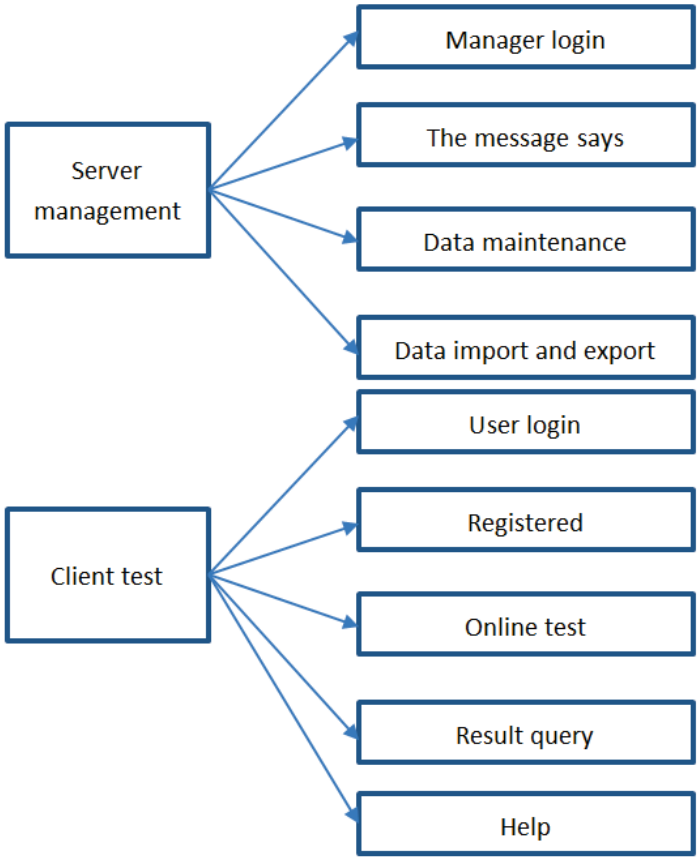

Figure 4 System function module diagram.

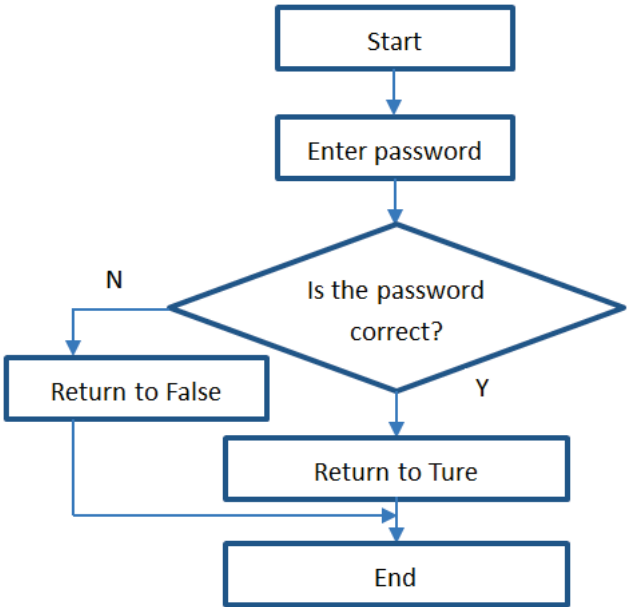

Figure 5 Password detection function flow.

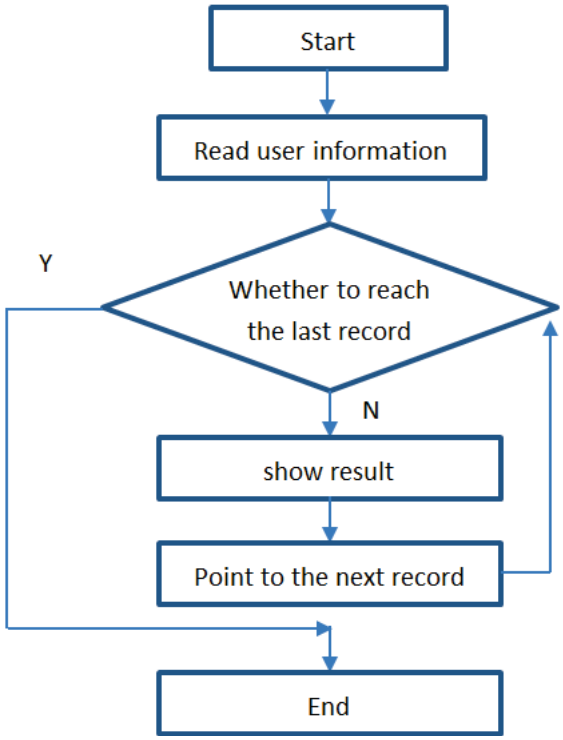

Figure 6 User information display flow chart.

to complete, which requires a function that can complete the detection of the password and return the result. When password verification is required, the password detection function is called, and the return value of this function is used to determine whether the password entered by the user is correct, as shown in Figure 5.

In the data maintenance module, all tester information is maintained, mainly including modification, deletion, and addition of tester information. First, all the tester information should be displayed in the interface, and each tester information is read out from the database at one time and displayed on the interface. The specific operation process is shown in Figure 6.

The system information table is used to store some data necessary to maintain the normal operation of the system, such as the administrator password, the number of current online testers, etc., as shown in Table 1. The system information table is relatively simple, only two fields need to be stored, and the items of the information and the corresponding values are respectively stored. However, considering that different data has different types, the "data value" field is divided into two, which are used to store different data types (text

Table 1 System information table structure.

Field name	Type	Length	Description
Item	text	50	System information item
Value1	text	50	Value (text type)
Value2	Long integer text	–	Value (value type)

Table 2 Descriptive statistical analysis.

Influencing factors of course quality	Test group 1	Test group 2	Test group 3
Course offered	0.5344	0.478	0.7443
faculty	0.3454	0.3765	0.454
Teaching system	0.5565	0.9676	0.5654
Student	0.3656	0.3765	0.8434

Table 3 Descriptive Statistics of the Dimensions of College Students' School Adaptation Scale.

factors	Learning adjustment		Self adjustment		Career adaptation		Interpersonal relationship		Emotional adaptation	
	M	SD	M	SD	M	SD	M	SD	M	SD
Test group 1	38.31	4.11	28.92	5.51	30.52	6.31	16.57	4.11	32.21	6.33
Test group 2	37.80	3.83	28.10	5.34	31.16	6.87	16.47	3.83	31.38	6.14
Test group 3	36.25	3.83	28.56	5.76	30.09	7.15	16.77	3.67	31.87	6.56

and numbers). Moreover, the Item field is set to the primary key.

Relationship describes the contiguous relationship and one-to-one correspondence between tables in the database. In the table designed above, each table records a specific range of content, but the content is also related.

4. ANALYSIS AND DISCUSSION

The system structure design is a design process based on the perspective of the developer. The criteria are structural rationality, easy implementation and maintainability, high efficiency and stability. However, functional design is a design process based on the perspective of the system user. The design criteria mainly include two aspects: one is to meet the user's required functions, and the other is to provide a humanized interactive interface.

Psychological testing has a long history, and traditional testing methods mainly include questionnaire testing and conversation testing. However, these test methods are only suitable for occasions with a small number of people. This system is a psychological testing system developed according to the characteristics of the number of college students. With the popularity and application of computers and smart mobile terminals and the popular use of smart phones, people's learning life is increasingly dependent on computer networks and mobile phone networks. The use of smart mobile terminal devices has penetrated into all areas of our daily life and learning work. Since the system of the present invention uses devices such as computers and smart phones that have been popularized, and popular network platforms and adopts a modular design method based on B/S structure, the system development cost is low, and the maintenance is convenient at a later stage. The system has a long-life cycle and a relatively low cost, and reduces the use of paper materials, which meets the requirements of environmental protection and sustainable development. The use of this psychological testing system will greatly reduce the investment of college teachers in labor and save manpower and material resources. Moreover, students are free to use any of the system to select test tasks for psychological testing anytime, anywhere, to improve student initiative and test authenticity. Therefore, this method is especially suitable for college teachers to carry out psychological testing work for college students.

In order to test the application performance of this method in quantitative evaluation of influencing factors of college students' psychological education course quality, the simulation experiment was carried out. The scale of the statistical sample is 2400, the training set is 100. The number of learning iterations and the iterative steps of the quantitative evaluation of the factors affecting the quality of psychological education courses are 1000 and 20 respectively. The influencing factors of curriculum quality in science education are: curriculum setting factors, faculty factors, teaching system and students themselves factor, respectively, and the descriptive statistical results of the influencing factors are shown in Table 2.

According to the results of descriptive statistical analysis, the factors influencing the quality of college students' psychological education course are modeled, and the contribution weight of each factor is analyzed, and the weight of influencing factors of college students' psychological education curriculum quality is obtained. The results of the analysis are shown in figure 7.

Figure 7 shows that the accuracy of modeling the influencing factors of college students' psychological education course quality by using this method is good, and the

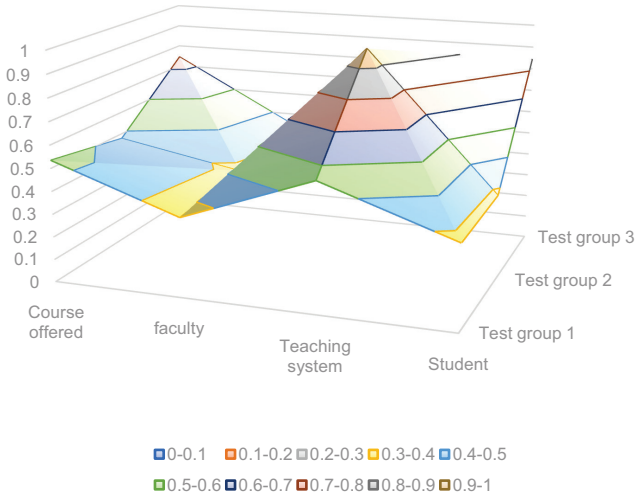

Figure 7 Weight distribution of influencing factors on the quality of psychological education courses for college students.

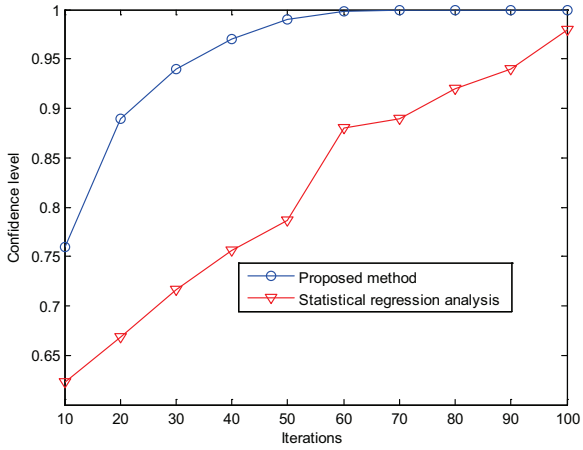

Figure 8 Comparison of confidence levels.

confidence degree of quantitative analysis of influencing factors of psychological education curriculum quality is tested by different methods. The comparative results are as shown in figure 8, and the analysis figure 2 shows that the accuracy and confidence level of using this method to model the influencing factors of college students' psychological education course quality are better.

Although mobile websites and computer websites are different on the surface, they all run on the same computer server and share and process the same database data. Therefore, teachers and students on campus only need to choose one of the access methods to operate. The mobile website based on the smart mobile terminal access is an important extension of the computer website accessed based on the computer mode, and the interface design is more concise than the computer website. The application of mobile websites to the psychological testing of college students can fully compensate for the lack of mobility and real-time of a single computer website. The combination of a mobile website based on smart mobile terminal access and a computer website accessed based on computer access has made up for the shortcomings of the two and has fully utilized the advantages of both. The mobile terminal connects to the Internet network through a data network or a WLAN network and accesses the computer server through the TCP/IP protocol for operation. Introducing intelligent mobile terminals into the psychological testing of college students can fully exert the portability advantages of the intelligent mobile terminal and the high efficiency advantage of the computer server in data processing, bring a new sense of freshness and experience to the user, improve the work efficiency, and promote the development of psychological testing work in colleges and universities.

5. CONCLUSION

The traditional college students' psychological cognition assessment methods mainly include three types: assessment of scales, survey visits, and observation and analysis. With the development of educational informatization, computer-aided technology is also gradually applied to the field of psychological cognitive assessment of college students. Moreover, the traditional method of assessing college students' psychological cognition relies heavily on the descriptions of third parties other than the evaluator, and the resulting evaluation results are not necessarily accurate. To this end, this paper designed a computer-aided evaluation system for

college students' psychological cognition. Through the use of computer access and mobile terminal access to the two access methods, CAD-based college students psychological test system provides a variety of choices for teachers and students, but also bring freshness and experience, and fully exploits the advantages of this system. With the popular use of smart phones, mobile Internet access has become more and more popular. Therefore, applying the mobile terminal to the psychological testing work has exerted the portability advantage of the mobile terminal, so that the psychological testing work can be performed anytime and anywhere, which not only facilitates the teachers and students, but also improves the work efficiency.

REFERENCES

1. Andrea A., Ferri F., Grifoni P., Guzzo T. (2015). Approaches, tools and applications for sentiment analysis implementation. International Journal of Computer Applications, 125, pp. 26–33.
2. Bi S , Ho C K , Zhang R, (2015). Wireless Powered Communication: Opportunities And Challenges. Ieee Communications Magazine, 53(4), pp. 117–125.
3. Bi Anqi, Wang Shitong, (2016). Transfer Affinity Propagation Clustering Algorithm Based On Kullback-leiber Distance. Jeit, 38(8), pp. 2076–2084.
4. Bayindir R., Colak I., Fulli G. (2016). Smart grid technologies and applications. Renewable & Sustainable Energy Reviews, 66, pp. 499–516.
5. Cattivelli F S, Sayed A H., (2011). Distributed Detection Over Adaptive Networks Using Diffusion Adaptation. Ieee Transactions On Signal Processing, 59(5), pp. 1917–1932.
6. Chitra K, Jeevarani B., (2013). Study On Basically Available, Scalable And Eventually Consistent Nosql Databases. International Journal Of Advanced Research In Computer Science And Software Engineering, 3(7), pp. 1356–1360.
7. Castelli M., Manzoni L., Vanneschi L. (2017). An expert system for extracting knowledge from customers' reviews: The case of amazon.com, inc. Expert Systems with Applications, 84, pp. 117–126.
8. Djenouri Y., Bendjoudi A., Mehdi M. (2015). Gpu-based bees swarm optimization for association rules mining. The Journal of Supercomputing, 71 (4), pp. 1318–1344.
9. Ebesu T., Fang Y. (2017). Neural semantic personalized ranking for item cold-start recommendation. Information Retrieval Journal, 20 (2), pp. 109–131.
10. Fang H., Zhang J., Bao Y. (2013). Towards effective online review systems in the chinese context: A cross-cultural empirical study. Electronic Commerce Research and Applications, 12, pp. 208–220.
11. Jiang Y Z, Chung F L, Wang S T, Et Al. (2015). Collaborative Fuzzy Clustering From Multiple Weighted Views. Ieee Transactions On Cybernetics, 45(4), pp. 688–701.
12. Leiva J., Palacios A., Aguado J. (2016). Smart metering trends, implications and necessities: a policy review. Renewable & Sustainable Energy Reviews, 55, pp. 227–233.
13. Ling Conggang, Wang Hongzhang, (2016). Optimization Research On Differential Evolution Algorithm And Its Application In Clustering Analysis. Modern Electronic Technology, 39 (13), pp. 103–107.
14. Li Mu-dong, Zhao Hui, Weng Xing-wei, Han Tong, (2016). Differential Evolution Based On Optimal Gaussian Random Walk And Individual Selection Strategies. Control And Decision, 31(08), pp. 1379–1386.
15. Mohan B, Govardhan A, (2013). Online Aggregation Using Mapreduce In Mongodb. International Journal Of Advanced Research In Computer Science And Software Engineering, 3(9), pp. 1157–1165.
16. Ulukus S, Yener A, Erkip E, Et Al, (2015). Energy Harvesting Wireless Communications:a Review Of Recent Advances. Ieee Journal On Selected Areas In Communications, 2015,33(3), pp. 360–381.
17. Zhang Y, Jiang C, Song L, Et Al., (2017). Incentive Mechanism For Mobile Crowdsourcing Using An Optimized Tournament Model. Ieee Journal On Selected Areas In Communications, 35(4), pp. 880–892.