



Construction of Scientific Research Training System and Investigate Its Efficiency on Teaching Reform in the Medical Big Data Mining Course

Yuan-Nong Ye^{1,2}(✉), Min Wang², and Meng-Ya Huang²

¹ Key Laboratory of Infectious Immune and Antibody Engineering of Guizhou Province, School of Big Health, Guizhou Medical University, Guiyang, China

yyn@gmc.edu.cn

² Department of Medical Informatics, School of Big Health, Guizhou Medical University, Guiyang, China

Abstract. The importance of medical big data mining in scientific research cannot be ignored. In order to cultivate undergraduates' medical scientific research literacy, improve their scientific statistical thinking and ability to solve practical problems in scientific research, it is necessary to combine scientific research training courses in undergraduate medical big data mining teaching. In this study, we developed an online scientific research training system using Mysql and Django. It adopts the theoretical knowledge of software engineering, carries on the demand analysis to the system, the module design, the object-oriented realization code writing, and finally completes a functional drug storehouse management system. And hence we use the system to investigate the teaching efficiency of the teaching reform based on scientific research training. And the machine learning and statistical methods were using to analyse the teaching reform efficiency in comparison with traditional classroom teaching. Through the comparison of the two groups, it is found that the average scores of the two groups are 64.90 and 77.13, respectively. Through the t-test analysis, the P value is 8.20×10^{-6} , far less than 0.01, indicating that there is a significant difference in the distribution of the scores of the two groups, indicating that under the online scientific research training system, students have better learning results. On the basis of analysing the present situation of medical big data mining teaching, this paper puts forward a new teaching mode of medical big data mining combined with scientific research training, and puts forward specific ideas and practice methods on the reform of teaching content, teaching mode and examination mode based on practical teaching experience. It is hoped that by strengthening the scientific research training for undergraduates, their scientific research consciousness and innovation ability can be improved.

Keywords: medical big data mining · scientific research training · undergraduates · instructional design and technology

1 Introduction

Data mining refers to the process of searching hidden information through algorithms from a large amount of data [10]. Data mining is usually associated with computer science and is accomplished through data mining, online analytical processing, information retrieval, machine learning, expert systems (relying on past rules of thumb), and pattern recognition. Medical big data mining is a subject that applies data mining method to medical practice [9]. Due to the nonrepresentational concept of medical big data mining, complex derivation of mathematical formulas, strong logic and other reasons, students with medical background have weak mathematical foundation, so they feel particularly difficult and boring to learn the medical big data mining course [2]. At present, the teaching of medical big data mining course mostly adopts the traditional teaching method and teaching content of partial mathematical theory, which leads to the derailment of teaching and scientific research, which is also one of the reasons for the unsatisfactory teaching effect [4].

Therefore, this paper proposes to adopt the theoretical knowledge of software engineering, carries on the demand analysis to the system, the module design, the object-oriented realization code writing, and finally completes a functional drug storehouse management system. And we use this system to carry out a new teaching mode of medical big data mining combined with scientific research training to guide undergraduates to

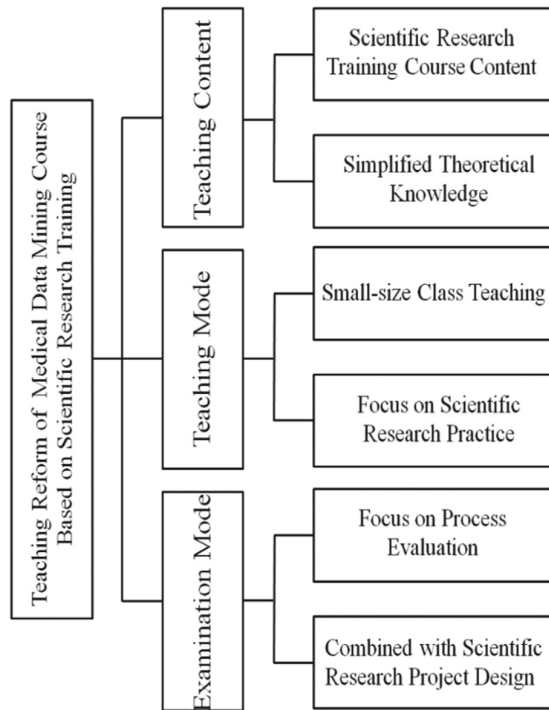


Fig. 1. Frame chart of teaching mode of medical big data mining combined with scientific research training course.

master basic scientific research methods and learn to design scientific research task [1]. The undergraduates study medical big data mining during scientific research practice, which train undergraduates to use medical big data mining knowledge to solve practical scientific research problems [7, 11, 12].

The new teaching mode of medical big data mining combined with scientific research training mainly involves the following three aspects: teaching content, teaching mode and examination mode [7]. The reform framework involved in this teaching model is shown in Fig. 1.

2 Methods and Materials

2.1 Building the Online Scientific Research Platform Training System

In order to better let students participate in scientific research, we developed an online scientific research topic management system. Students can participate in scientific research projects on the platform in their spare time and apply what they have learned in class to project practice [13]. The platform records the whole process of students' scientific research. This system should be met the following business description:

- (1) The system must be able to record the progress and results of each student's scientific research project.
- (2) The system should be able to achieve the basic topic release.
- (3) The system should have strict permission control to ensure security.
- (4) The system must strictly control the login times, so as to avoid illegal elements using login vulnerabilities to attack.
- (5) The system should have a simple and friendly interface with fast response speed.

2.2 Bring in the Content of Scientific Research Training Courses

Scientific research training courses mainly focus on the basic knowledge, basic methods and basic skills of scientific research. It aims at giving undergraduates who lack practical experience in scientific research a basic understanding of scientific research work [5]. Moreover, it helps the undergraduates master the basic knowledge and skills of scientific research work.

The teaching content mainly includes research topic selection, research project application, research plan formulation, experimental research, research results analysis and summary and paper writing, etc. In scientific research training course in the teaching process, we found that although the students can remember the relevant content on the books, but for the each link lack of practice and understanding of scientific research. Students generally reported that the course was of no practical value. Therefore, we propose to combine scientific research training with medical big data mining courses for teaching.

2.3 Simplify Theoretical Knowledge and Aggrandize Operation of Data Mining Software

Most students feedback that the formula derivation of medical big data mining course is complicated and the theoretical knowledge is non-objective, which makes it very boring to learn. Therefore, students' enthusiasm for this course is weakened, leading to the weak willingness and ability of independent learning.

Under the new teaching mode, teachers should simplify the theoretical knowledge of medical big data mining and dilute the explanation and derivation of formulas. And it increases the proportion of computer operation of data mining software such as Kaggle and Weka in the course [3, 8]. It pay attention to the interpretation of data mining results and use data mining methods to solve scientific research problems, stimulate students' interest in learning medical big data mining courses, enhance their confidence in learning this course.

2.4 Small-Size Class Teaching Simplify Theoretical Knowledge and Aggrandize Operation of Data Mining Software

Large-size class teaching is a kind of “demonstration” teaching. The instructor takes into account the wide range of students in the class, so the course is set at a medium level of difficulty [6]. After class, students need to do a lot of review to learn a subject well. In addition, interaction between teachers and students is limited in large class environments. Students are also falling far short of their potential. In developed countries, small class sizes have become common practice. The Chinese Universities, such as Peking University and China University of Petroleum are also promoting small-size class. Therefore, this study proposes a teaching method of Medical big data mining based on small-size class teaching mode.

In the teaching mode of small-size class, the teacher can take each student into account, and the interaction between teachers and students is more frequent. When the teacher asks questions, each student has the opportunity to interact and students can better respond to the teacher. And hence, it drives students actively into teaching activities, so that the whole classroom really active. And then, it would fully activate the active initiative of students to learn. In addition, the number of students in small class teaching is small, so teachers can better understand the advantages and disadvantages of each student and learning status, help students improve and develop themselves, and help students learn the course well.

2.5 Focus on Scientific Research Practice

Taking theoretical teaching as the leading, and paying attention to scientific research practice training, we should establish a teaching mode that can improve students' theoretical knowledge and scientific research thinking ability. Teachers guide students step by step from data acquisition, scientific research questions, data sorting, problem solving, etc., to stimulate students' interest in scientific research. Table 1 provides an example of guiding students to explore scientific research questions. Students are required to

Table 1. Examples of statistical practice combined with scientific research training.

Item	Description
Problem	How many types of autism?
Method	K-means clustering
Data mining	Data cleaning, data normalization, dimensionality reduction, feature selection, clustering
Conclusion	It can be divided into three or five categories

download behavioural data of 2226 subjects from an open database abroad, including autistic patients and normal controls from different research centres.

After a semester of teaching for Medical Information Engineering students in medical college, we found that 100% of the students could raise questions, but only 38.45% of the students raised questions of basic scientific value; 48.98% of the students could choose the correct data mining method to analyse the questions; 80.78% of the students can correctly use data mining software to analyse and analyse data mining results. This indicates that it is difficult for undergraduates to combine scientific research problems with data mining methods. Therefore, in the teaching of medical big data mining, it is necessary to strengthen the scientific research practice of undergraduate students, and cultivate their scientific research consciousness and innovation ability.

In scientific research and practice of teaching mode, students interested in the research of small problems, to excavate the undergraduate scientific research potential, to make the students master the theoretical knowledge while, can apply this knowledge to solve practical problems in scientific research, to learn is boring, learn something and use it, breaking the rote learning machine learning model. This kind of training cultivates the students' scientific research thinking ability and enables them to master certain scientific research skills.

2.6 Combined with Scientific Research Project Design

Medical big data mining is an important tool in medical science research, so the key of assessment is how students use this tool well. After the students have learned the content of the scientific research training part, they will be guided to design the scientific research topic in groups, and gradually improve the topic, including data mining methods and analysis of data mining results at each stage, until a complete scientific research topic is formed at the end of the course.

2.7 Combined with the Computer Operation Examination

Since engaged in the teaching of medical big data mining, medical students generally feedback that they like computer operation very much, because in the process of computer operation, they can personally feel the use of theoretical knowledge. In this course assessment, can be combined with computer practice operation examination. For example: provide a large number of medical data to students, according to the data, students

set their own scientific research problems, and use appropriate data mining methods to solve the problem, and analyse the data mining results in detail.

3 Results and Discussion

The system is now divided into landing module; landing module is divided into direct landing module and registration module. The system structure includes back-end support, front-end development and surface application. Below is the interface of the system.

In order to study the effect of scientific research training on students' innovation ability and course learning ability, the data used in this article is there cords exported from the online scientific research platform training system for further analysis. We selected 32 students as the small-class teaching of scientific research training (Treatment group), and the remaining 62 students as the large-class teaching of conventional teaching mode (Control group). We took the final exam results as the analysis and comparison.

According to the above score statistics and distribution chart (Table 2, Fig. 2 and Fig. 3), it can be seen that there are students of different levels in both groups. The average scores of the two groups are 77.13 and 64.90 respectively. According to T test

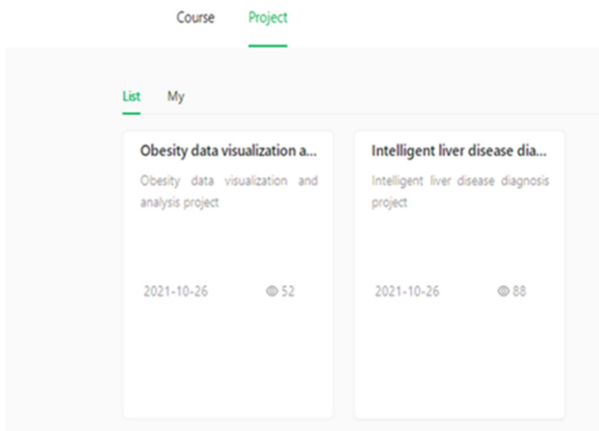


Fig. 2. The interface of the scientific research training system (Screenshots from the scientific research training system, which is developed by Yuan-Nong Ye/<http://10.100.98.13/strain>).

Table 2. The exam score reticulation.

Group	Treatment	Description
Maximal score	96	95
Minimal score	69	14
Average score	77.13	64.90
Standard deviation	7.17	14.96

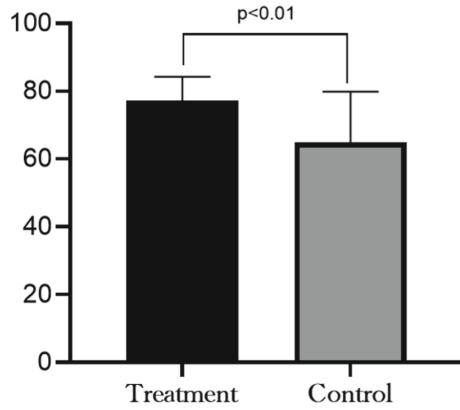


Fig. 3. Frame chart of difference comparison.

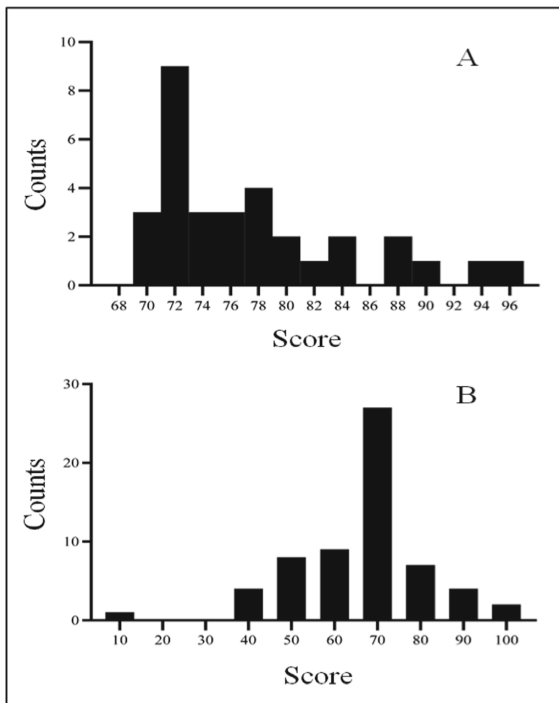


Fig. 4. Grade distribution map.

analysis, the P value is far less than 0.01, indicating a significant difference in the score distribution of the two groups. In the small-class teaching of scientific research training, students' learning results are better (Fig. 3).

4 Conclusions

For medical information engineering major, big data mining is a highly practical course, which is of great significance in experimental teaching. In the teaching design schedule, to carry out our practice education, scientific research, education, culture, education, health, education, the school idea, further deepening the reform of education teaching, strengthen the cultivation of practice ability, we constantly practice teaching reform and explore different ways of experiment course influence on student's practice ability, including use the experimental model, such as "scientific way" Practice teaching should be strengthened and comprehensive and innovative experiments should be added to improve students' ability to comprehensively apply programming knowledge to solve practical problems.

In summary, after students' feedback and self-summary, the teaching reform has achieved certain results, but there are still shortcomings. It is hoped that the teaching links will be continuously improved and refined in the subsequent teaching process, so as to further improve the teaching quality and effect of medical big data mining course.

Acknowledgements. The authors are grateful to anonymous reviewers for their valuable suggestions and comments, which have led to the improvement of this paper. And we thank the funds of Teaching Reform Research Project of Education Steering Committee of Electronic and Information Majors of Higher Education, Ministry of Education (2021-JG-21); Association of Fundamental Computing Education in Chinese Universities Research Project (2021-AFCEC-188); Guizhou Medical University Teaching Content and Curriculum System Reform Project (JG2021074, 0620302050136); Ministry of Education industry-University Cooperative Education Program (202101018013); Curriculum Ideology and Politics Construction Project of Guizhou Medical University (Biostatistical analysis, No. 39).

References

1. Ageed, Z. S., et al., 2021. Comprehensive survey of big data mining approaches in cloud systems. *Qubahan Academic Journal* 1.2 (2021): 29-38.
2. Bai, M., et al., 2019. Current status, challenges, and future directions of university laboratory safety in China. *Journal of Loss Prevention in the Process Industries* 74 (2022): 104671.
3. Bojer, C. S. and Jens, P. M., 2021. Kaggle forecasting competitions: An overlooked learning opportunity. *International Journal of Forecasting* 37.2 (2021): 587-603.
4. Chun T., Ylona, M. B. and Karen F., 2019. Grounded theory research: A design framework for novice researchers. *SAGE open medicine* 7 (2019): 2050312118822927.
5. Ferri, F., Patrizia G. and Tiziana G., 2020. Online learning and emergency remote teaching: Opportunities and challenges in emergency situations. *Societies* 10.4 (2020): 86.
6. Li, X. and Li, M., 2021. The Effect of Large and Small Class Size in Compulsory Education in China. *Rangsit Journal of Educational Studies* 8.1 (2021): 46-58.
7. Li, Y. and He M., 2022. Study on Machine Learning Applications in Ideological and Political Education under the Background of Big Data. *Scientific Programming* 2022 (2022).
8. Saleh, B., et al., 2020. Analysis of Weka Data Mining Techniques for Heart Disease Prediction System. *International journal of medical reviews* 7.1 (2020): 15-24.
9. Song, C.W., Hoill, J., and Kyungyong, C., 2019. Development of a medical big-data mining process using topic modeling. *Cluster Computing* 22.1 (2019): 1949-1958.

10. Tarawneh, M. and Ossama, E., 2019. Hybrid approach for heart disease prediction using data mining techniques. International Conference on Emerging Internetworking, Data & Web Technologies. Springer, Cham, 2019.
11. Tarmizi, S. S., et al., 2019. A review on student attrition in higher education using big data analytics and data mining techniques. International Journal of Modern Education and Computer Science 10.8 (2019): 1.
12. Yi, L. and Zhai, X., 2018. Review and prospect of modern education using big data. Procedia Computer Science 129 (2018): 341-347.
13. Zheng, S., 2018. Research on the Application of Big Data in the Field of Online Education. DEStech Transactions on Computer Science and Engineering iciti (2018).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

