

Exploration of Data Mining Courses for Non-Computer Majors

Chunmei Li

Qinghai University
Department of Computer Technology and Applications
Xining, China
e-mail: li_chm0422@sina.com

Yuming Wang

Qinghai University
Department of Computer Technology and Applications
Xining, China
e-mail: wangym_crystal@126.com

Zhiqiang Liu

Department of Computer Technology and Applications
Xining, China
e-mail: wsliufree@sohu.com

Lu Wang

Department of Computer Technology and Applications
Xining, China
e-mail: 465232991@qq.com

Abstract—This paper analysis the necessity of setting up the courses -data mining engineering for Non-computer majors from aspects such as the field of big data, formats, channels, media industry competition, comprehensive talents training, cross disciplinary under the era of big data. Introduces investigation effect of the course of data mining for non-computer majors in our university. Pointed out that it is the demand of the times to cultivate comprehensive talents with scientific analysis ability in the era of big data

Keywords-Big Data; Non-Computer Major; Data Mining Course

I. INTRODUCTION

For a half of century, with the full incorporation of computer technology into social life .The explosion of information has accumulated to the point where it has begun to cause change. Not only has it flooded the world with more information than ever before, but also its rate of growth is also accelerating. The disciplines of information explosion, such as astronomy and genetics, have created the concept of "big data."

In this context, the internet (social networking, search, ecommerce), mobile internet, internet of things (sensors, smart earth), GPS, medical imaging, security monitoring, finance (banking, stock market, insurance) Call, SMS) are crazy to produce data. Data storage unit by the B, MB, GB, TB, PB, EB, ZB, YB, BB, are growing with 1024 times speed of linear. There are countless daily data.

In May 2011, a report released by the U.S. consultancy McKinsey & Company on "Big Data: The Next Frontier in Innovation, Competition and Productivity" opened the door to big data industry. The report argues that big data is a dataset of sizes that go beyond the capabilities of traditional database software tools to access, store, manage, and analyze. International data companies summarize the four characteristics of big data: mass, diversification, rapid, value[1]. With the development of computer technology, various types of data terminals that focus on different types and contents of different formats in various fields generate a large amount of data and are influencing and changing

people's lifestyles, which are influencing and changing the mode of business survival, the country's decision-making strategy and international exchange and deployment. Big data affects everything.

II. THE NEED FOR NON-COMPUTER MAJORS TO OPEN DATA MINING COURSE IN BIG DATA ENVIRONMENT.

A. Big data is generating more and more fields

Big data is no longer a patent in astronomy and biology, and today's big data comes from all walks of life. Every day business generates a large number of commercial data [2], and data formats are not uniform ,such as electronic document data, used to store customer's basic information, transaction information, profitability analysis, another example business approval documents, rules and regulations Meeting minutes and records, etc.; and the medical community generates a rich medical database every day[3], such as the patient's medical records, B-ultrasound information, imaging information, ECG information, electronic health records, electronic medical records, medical images, etc. Meteorological data generated by meteorological data [4], meteorological data include not only conventional observations and experimental data such as terrestrial, upper-air and oceanographic data, statistical data, forecasting and other analytical data, but also a large number of abnormal Such as satellite and radar, automatic detection of station data, etc. In the aspect of water and electricity [5], generating a large number of daily power data, In the aspect of ecological continues to produce all kinds of ecological resources data[6], such as watermarks, weather and other data. The press generates a large amount of news and information every day, such as traffic information, sports information, film and television industry Smell, etc., including personal mobile phones are produced every day a lot of prestige circle of friends information, blog information, chat messages, phone call information and so on.

B. The channels and formats for generating big data are becoming more and more widespread

Big data produces different industry, different equipment, different channels, resulting in different data formats. The generation of data have a great relationship with the environment generating data and hardware and software. Such as video data, voice data, text data, pure digital data, sensor data, network data, analog data, digital data, nominal data, binary data and so on.

C. The media push and speech guidance under big data has become more and more

In the context of big data, huge amounts of information are generated every day, so that some people analyze the data every day, so as to guide and influence people's ideas and disseminate and push them through the media such as computers, mobile phones and televisions, guiding and exerting subtle influence every moment People's thinking. Smartphones are an essential communication tool for all adults, and a large part of the push of various types of news media is based on mobile terminals, such as various types of mobile news pushing applications such as Tencent News.

D. Big data industry competition more and more powerful

The big data industry competition is no longer the traditional competitive model, who grasp the market, who understand the client, who has advanced scientific analysis methods, who adopt scientific management and decision-making, who is the winner. As a result, competition in the industry is getting more and more intense, which is fairer from the perspective of the survival of the fittest.

The demand for comprehensive talents under big data is becoming more and more urgent

In the context of big data, it is not enough for the development of the industry to rely solely on the knowledge of the industry. It is necessary to master the method of analyzing the domain knowledge. That is to master the data mining method, which is the foundation of the industry. Therefore, it is urgent to grasp the demand of comprehensive talents in field data processing methods.

The development of interdisciplinary under big data is promising

In the context of big data, the cross between disciplines becomes the characteristics of the times. The study of interdisciplinary research is one of the hot research topics at present. The study of interdisciplinary research has become the demand of the big data age. Therefore, the development prospect of interdisciplinary is very extensive.

III. THE ATTEMPT OF CURRICULUM REFORM IN OUT SCHOOL

In order to cultivate a comprehensive range of interdisciplinary talents, we first set up the "Data Processing and Analysis" (Data Mining) course for some of the non-computer major and we replaced the existing "Software Technology Foundation" public course with "Data Processing and Analysis "(Data Mining) to enable them to master the professional knowledge of the premise, learn to

use scientific idea and scientific method to analyze the problems in the field, in the future to make greater contribution to the industry.

Doing a comprehensive publicity and promoting the coming of big data era, so that everyone deeply understand the importance of learning data mining

We are advertising the big data era has quietly come through advertising on the huge billboards and promoting the applications of data mining in all walks of life, so that students have a profound understanding of the era of big data and learn to use scientific methods to analyze and deal with professional issues in the industry. This is in the long-term based on the industry. In addition to advertising on the huge billboards, we also use lectures, preaching on the theme of the beauty of big data to promote the advent of the era of big data and learn the importance of data mining methods.

The first time to set up data mining course for part of the engineering major of non-computer major on the whole school

For the first time, we set up "Data Processing and Analysis"(Data Mining)course for 16 classes of sophomores in electrical engineering major of water conservancy and electric power school, civil engineering major of civil engineering school and automation major of chemical engineering school. We explain the application and the theory in the classroom, using weka software environment to do data analysis experiments in the experimental class. Finally, let the way of lecture and discussion to summarize the course. After a semester of learning, we found that learning attitude of the course of students is very positive, exceeding our expectations.

To point the surface, affecting other non-computer major students interested in data mining courses

In the future, we plan to adopt the methods of online courses such as micro-course and Mooc and face the whole school to set up "Data Processing and Analysis"(Data Mining) course. Pointing to area and influencing other non-computer major students to actively learn the knowledge of data mining to train comprehensive talents for university to work hard.

IV. ACHIEVEMENTS THAT WE GET

Through the mobile platform, face-to-face communication, questionnaires and other means to know about our completed work and the achievements, there are a lot of benefits.

Spread rate of WeChat is very high, students in the discussion on this course in the mobile WeChat platform, or circle of friends is very fierce, so we achieved our goal.

Advertising and preaching make teachers and students know big data. By means of random sampling on campus, we know about students are no longer unfamiliar with big data through the above various publicity and have a certain understanding about the importance of data mining.

We randomly selected students in the school to conduct a questionnaire survey on "lectures, advertisements, and courses on theme of the beauty of big data." We randomly distributed 1000 questionnaires and collected 800 useful

questionnaires. From which we can learn the change of from students never heard of big data to the needs of setting up data analysis courses, the survey results are as follows:

About the question "Do you know the concepts of big data, data mining, etc.? A: never hear about B: understanding". We learned that the change of from students never heard of big data to understand big data in the following Figure 1.(Whether knowing the concepts of big data, data mining, etc.), only 30 of 800 answered "Never hear about" and 770 answered "Get It."

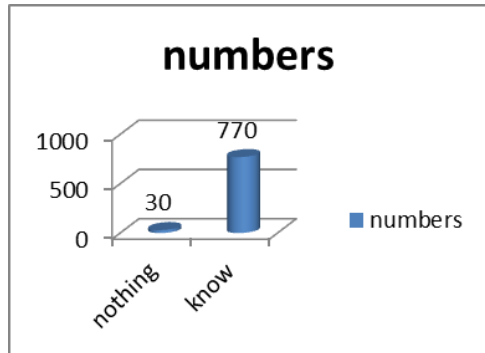


Figure 1. Whether you heard of such concepts as big data, data mining

The revenue of questionnaires is the most. About we set up the "Data Processing and Analysis" course, we conducted a questionnaire and found that the results were very good. We give out 900 questionnaires and 100% of them were recycled.

Aiming at "Your attitude toward setting up this course? A:No necessity, B: necessity ".And The survey results of students' attitudes toward the need to open data mining are shown in Figure 2 below (Whether it is necessary to set up a data mining class).

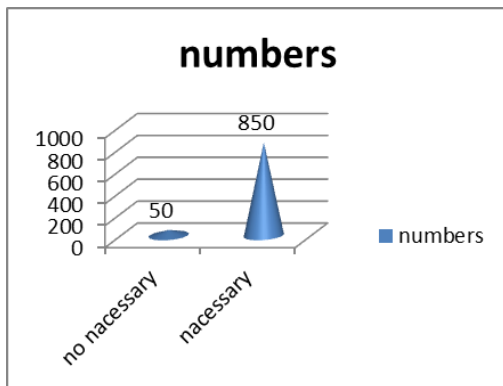


Figure 2. Whether it is necessary to open a data mining course

The results of the survey on the satisfaction of non-computer major engineering students who have set up data mining courses are shown in Figure 3 below (whether they are satisfied with the effectiveness of this course). It can be seen that 89% of the students are satisfied, and only 11% of the students are dissatisfied. So the effect is good.

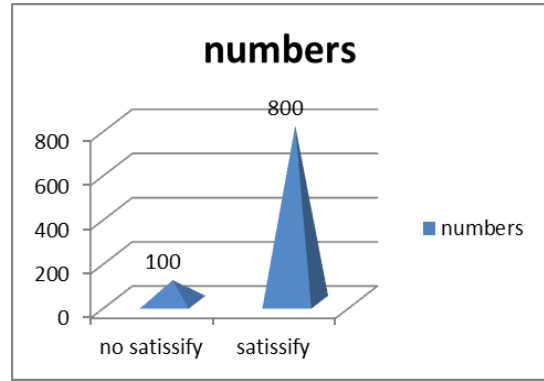


Figure 3. Satisfaction with the effectiveness of this course

V. CONCLUSION

In the era of big data, it is necessary to learn how to analyze problems with scientific methods, solve problems and see deeper knowledge from phenomena so as to cope with the ever-changing subject issues. Everyone learns to understand phenomena, analyze problems and solve problems with scientific methods and means, and interdisciplinarity has become a demand of the times and has become one of the important means of training comprehensive talents.

ACKNOWLEDGMENT

Project: Curriculum construction project of Qinghai University, Fund number: KCFL-16-3-2, qdyk-170110

REFERENCES

- [1] Guo Xiaoke. *Big Data*[M]. Tsinghua University Press, 2013:5.
- [2] Lu Jin-song ,Wang Zhi-cheng, Wang Qin-hui, Xu Quan, Research on the Unstructured Data of CommercialBank Auditing in Big Data Environments, *Soft Science*, 2017.1, 205 (31):141-144
- [3] Lu Yi, Huang Zheng-xing, Yu Si-wei, Duan Hui-long., Research situation and expectation of clinical healthcare big data [j], *Chinese Medical Equipment Journal*, 2017.3, 38 (3):112-115
- [4] Nie Fengying, The research of the big data resources technical services collaboration-A Case Study of Meteorological Data [j], *Informatization Research*, 2016.2, 42 (1):6-11
- [5] Chen Junfei, Dneg Menghua, Wang Huimin, A review of water resources big data [J], *ADVANCES IN WATER SCIENCE*, 2017.7, 28(4):622-631
- [6] Yang Rukun, Practice and Innovation of Promoting Environmental Information and Ecological Monitoring Network Construction in Qinghai Province [j], *Environmental Protection*, 2017.12, 20-22