Big Data Technology, Logistics Engineering and General Education Courses

Jian-Hua ZHANG\textsuperscript{1,a}, Zhe ZHANG\textsuperscript{2,b}, Hui-Juan ZHAO\textsuperscript{1,c}, Yi-Jing XU\textsuperscript{1,d}

\textsuperscript{1}General Education Center, Beijing Normal University Zhuhai Campus, China
\textsuperscript{2}Shenyang Yunyou Network Science and Technology Limited Company, China
\textsuperscript{a}zhuhaijianhua@aliyun.com, \textsuperscript{b}engel2002@msn.com, \textsuperscript{c}bnuzhj2008@sina.com, \textsuperscript{d}845443202@qq.com

Keywords: Big Data Technology, Logistics Engineering Course, General Education Course.

Abstract. This paper first introduces the "Big Data Processing and Data Mining" this course history in domestic universities; Then it is pointed out that in today's teaching reform, the contents of the teaching should be changed to meet the needs of social development. According teaching practice and experience, the author proposes that set the "Big Data Processing" and "Data Mining" the necessity of these two courses, and different between General Education and Logistics Engineering in colleges and universities. As a teacher teaching the courses for many years, this academic point of view, it is useful as reference.

Introduction

In July 2014, this explosive news shocked us: "IBM voted $100 million to support the domestic university and college for big data and analysis of talent training" In the IBM U-100 cooperation program, IBM will donate a series of data and analysis software and provide professional knowledge, support the 100 Chinese colleges and universities to train the next generation of data scientists. From that time up to now, in less than two years, the domestic colleges and universities in the field of teaching and research of data analysis, has made rapid progress.

First, many colleges and universities in different teaching methods, set and opened related about "Big Data Processing and Data Mining, the course. For example: Beihang University, in 2014, in the form of a series of lectures, for all the students to set and open this course. Beijing Normal University, Zuhai, in September 2014, in the way of general education courses for the whole school students to set and opened this elective courses. It should be said that the two universities, at first in the domestic college and university, had set up this undergraduate education curriculum in. At the same time, in 2014, Nanjing University proposed a "computer professional system capacity of the curriculum system," the overall idea. In 2015, held a meeting, that is about some colleges and universities in China base on studying of the basic courses of computer application in Universities, the "data processing" teaching content, textbook compiling and publishing, as well as the teaching way of seminars. However, the progress in the field of big data research, it is difficult to imagine the speed of the development of people. National economic construction and development, are catching up with the pace of technological progress, social demand has made more demands on our education sector.

The Latest Development of Big Data Technology

We are in the era of data explosion. Big data is getting hot in recent years. It is not the quantity of the data that is revolutionary. It is that we can now do something with it that's revolutionary. The power of big data solutions is they can process large and complex data sets very fast, generate better and faster insights than conventional methods. A big data solution suite can consist of several critical components, from the hardware layer like storage, compute and network, to data processing engine, to analytics layer where business insights are generated using improved statistical & computational algorithms and data visualization. Among all, the data processing engine is one most critical player. It is not overstating that the data processing engine for big data is like CPU for a computer or brain for a

213
human being. And in the past ten years, our research, mainly concentrated in the software framework for distributed processing of large amounts of data, as the Hadoop, and concentrated for distributed computing software on large amounts of data, as the MapReduce. In the last year, in 2015, computer workers’ research is the hottest topic, that is "studying of Spark" and “Spark application”.

Spark was initially started for the purpose of creating a big data processing and computing framework, when Matei Zaharia was doing his Ph.D. research at UC Berkeley AMPLab in 2009. Different from the traditional data processing framework, Spark's in-memory primitives provide performance up to 100 times faster for certain applications. By allowing user programs to load data into a cluster's memory and query it repeatedly, Spark is well-suited for big data and machine learning use cases. Spark is becoming one best adopted among all big data modules. Spark is now evolving the Hadoop and big data ecosystem to better support the end-to-end big data analytics needs. Learning Spark and its internals will not just help improve the processing speed for big data, but also help developers and data scientists create analytics applications with more ease. With big data solutions like Spark, we expect to see significant improvement with business insights which will help expedite the decision making—like we've never seen before, from enterprise, healthcare, transportation, and retail.

How to Adapt This Change of General Education Teaching in Universities

In recent years, about General Education and teaching, the domestic and foreign experts and scholars had published a lot of articles and works. On the principles of General Education courses, we have also published some articles. Now, we are more concerned about the teaching content of the big data course, whether it is necessary to carry out the corresponding update or re set. For example as Beijing Normal University. Zuhai, their "large data processing" courses, the content should include two parts: the first part is the data processing, selected the Big Data Process, the second part is the Data Mining. The first part of the content of Big data processing, is about the big data parallel processing framework, "Hadoop", and the second part is about big data distributed calculation software framework, "MapReduce". The second part of "data mining" telling the analysis software for the big data, "Mahout". So, in the future the content of this course should be what? We believe that the first part of the teaching content should be retained, the second part of the teaching content, can be added to the new "Spark" software and its application. At the same time, large data analysis software "Mahout" as the expansion of knowledge, introduce to the students.

Now, in the university to carry out general education, has been widely recognized. Various colleges and universities have different experiences. Each school preselected course are more than hundred courses, if including the different content MOOC courses and open online courses, there are hundreds of courses. And its content can be said to cover the content of modern science and technology to primary even the primary and middle school education content. In the face of this "mixed content, structured chaos, different level" phenomenon, we should be to discuss the following questions: what are the training goal of higher education. The higher education is our main work, but "Science Popularization Education" and "Mass education", it's just our additional works. To reduce the level of higher education in the university is a great waste of social resources, but also is the fall back of modern education. Systematic and Standardized that is making the quality and status of the course to be improved.

In order to faster promotion of large data processing technology, adapt to the actual situation of the current society is in urgent need of the talent, since 2015, Beijing Normal University Zhuhai Campus, already set "Big Data Processing and Data Mining" as a General education course.

Teaching Reform of Logistics Engineering and Management Specialty in Universities

The application of the big data processing, has the extremely broad prospects in all walks of life. In order to meet the needs of national construction, each college should be based on its own characteristics, to carry out a large data technology research and teaching work. Taking logistics engineering and management major as an example, it is very necessary to set up the courses about
related to data analysis. Now the reality is that only a few universities have opened courses related to data analysis. Only a handful of colleges and universities have opened courses about big data analysis. Therefore, we believe that the future of teaching and research works should be like this: The "Big Data Processing and Data Mining" course, should divides into two courses, and respectively named "Big Data Processing" and "Data Mining". The "Data Mining" course content should be the data analysis software "Mahout" and its application. In this course, the main is about the big data analysis algorithm and its application. Of course, this data analysis software isn't perfect, for example, it can only be run on a single computer. However, as the main purpose of learning algorithm, it is still capable of the work of the. It should be pointed out that: in the course of teaching, it should be more examples of its application.

Summary

The development of modern science and technology is faster. As a college educator, we have the responsibility to take on the responsibility of our country. According to the needs of the national economy and construction, to pay more efforts for the country to cultivate more and useful talents.

Acknowledgement

This paper is the research result of the teaching reform project of Guangdong Province, and it is supported by this project. The item number is GDJG20141261.

References

[1] Information on http://baike.baidu.com/link?url=r1vAIPjRVIVtf8cyNgXFcAiwYziTfDtDyc3tKEWGmhEJAOny94Z9z_1gJuCUU3-P6blDz5zkbt1i7o1kfZH38-5FKhU8CpcnBwYlsNyrzMBJsXYSIXAFIInnnwc2NsU0-oR-Yz51vMhvH0XnQ11hvFHzRm1VVwsheRnM8fqgOX8G4yKkEqCq_3RqKbb0o18KxXnJrhqHSDp1pwFztaqeRa
[2] Information on http://baike.baidu.com/link?url=B3QN8bCpcHIWrgl391eJ0diIR4VdtyTbwVxrzB_91TZoV4YiCiV8a7IQMiP20W1kyExxm1smGA ZZW18fh7q