The Training Mode of University Experimental Talents Based on Big Data

Yunhai Zhao

Shaanxi Radio and Television University, xi'an 710119

Email: quchangbu69m@163.com

Keywords: Colleges and universities; Big data; Education management; Communicated by: (to be filled by the JWE editorial); Experimental talents; Training model

Abstract: The development of big data is inseparable from the education of college experimental talents. In the new stage of higher education management development, we can find value through big data. Big data enables universities to adopt smarter ways to inspire and generate new wisdom. There is a fundamental difference between big data in university education management and big data in business fields. The big data of colleges and universities take the relationship as a breakthrough, and finally seek a special relationship. The use of big data, cloud computing and the Internet of Things technology is an important tool and foundation for higher education to improve school running efficiency and promote higher education management from incremental development to high-quality development. At present, there are not many systematic research results on the combination of big data and university education management in academia, the depth and breadth are not enough, and specific empirical research is lacking. This article researches and discusses the training of college experimental talents and big data technologies, hoping to help related practitioners.

1. Introduction

Mankind has entered a new era: the information age. The rapid development of information technology with the core of network communication technology and multimedia technology have aroused profound changes in many fields and accelerated the progress of information society[1]. At the same time, with the acceleration of the global economic integration process, the degree of information has become an important symbol to measure the international competitiveness and modernization level of a country and region, as well as an important symbol to measure the overall national strength and the level of economic growth[2]. With the development of information technology, people pay more and more attention to information technology. As the forefront of educational informatization, the informationization of higher education has had an unprecedented impact on teaching in colleges and universities. The traditional teaching and learning environment of campus, classroom and library have changed dramatically. The time and space of education and teaching have also changed greatly[3].

With the popularization of higher education, the connotation of education management is constantly expanding and expanding, which leads to the complexity of education management[4]. How can we improve the quality, efficiency and level of education management and adapt to the needs of education management in the information age? I think we must speed up the pace of information technology in education management[5].

The digital campus is the physical manifestation of the education information. Under the guidance of the theory, the information technology of computer, network and multimedia is integrated with the education and teaching, and a new educational system is formed[6]. "Digital Campus" is a complex information system. Its construction can be divided into four parts: reconstructing the concept of education, building network hardware, building application system centered on teaching application, and building information resource management platform software system. The construction of digital

campus makes the original single mode of university education management informatization face the challenge of integration in the campus card.

Under the new situation, higher education management faces many new challenges. Networking, informatization and standardization are inevitable choices for the reform of educational management in colleges and universities. In order to survive and develop for a long time, colleges and universities should not only have a high level of teaching but also a high level of management[7]. It is undeniable that many universities have made some achievements in the construction of educational information management. However, due to technical limitations and personnel literacy constraints, there are still many problems to be solved in the process of educational management informationization[8].

2. Large Data Technology

2.1 Research on large data distributed storage technology and system

In the era of large data, the data scale of many industry application systems is expanding, which leads to the traditional data storage management methods and systems cannot effectively cope with the growing demand of large data storage management. In order to effectively solve the problem of storage and management of large data, a variety of large-scale distributed data storage systems have been launched worldwide. The most widely used distributed file system is HadoopHDFS. However, because HDFS and other mainstream file systems are data storage systems based on traditional hard disk, the reading and writing operation under large-scale data scenes is a big performance bottleneck, and it is difficult to meet the needs of more and more real-time or quasi real-time large data analysis and processing applications[9].

Allxio is a unified large data storage system that can support hierarchical large data distributed storage management based on MEM-SSD-HDD (memory solid disk hard disk). The early name of Alluxio was Tachyon, which came from the AMP Laboratory of University of California at Berkeley. The main idea of its early design is to use distributed memory to store and manage data, thereby speeding up the reading and writing speed of large data storage systems. However, although the memory allocation capacity of the current big data processing server node is already smaller than that of island (64GB, even 128GB or 256GB). However, compared with the actual application of TB or even PB level data, it is far from enough to store and manage large-scale data entirely based on memory. Therefore, a hierarchical large data distributed storage architecture based on MEM-SSD-HDD is needed to handle memory-based cache scheduling for stored data.

In this hierarchical storage scenario, we need to study memory cache scheduling strategies and techniques based on large-scale hierarchical storage structure. In a hierarchical large data storage system, storage devices are organized hierarchically, and the higher the storage capacity is, the closer it is to CPU. Similar system architectures have been used in CPU cache for decades. As shown in Figure 1, the more frequently visited data will be stored in the storage medium with higher performance. Data such as memory or SS instead of data will be moved to the HDD layer with low access speed.

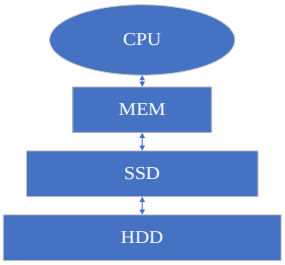

Figure 1 Schematic diagram of hierarchical storage structure of CPU- [MEM-SSD-HDD].

2.2 Mainstream large data parallel computing system

Big data parallel computing system is in the core of big data processing technology. They can access large-scale data and process the underlying distributed file systems (such as HDFS) to provide parallel computing services for large data applications on the upper level.

Compared to the traditional single computer computing system, large data and catty computing systems (such as Hadoop, Spark) usually use distributed architecture, which makes them have large scale and catty processing capability, better system fault tolerance and scalability. However, in the process of design and implementation, the mainstream large data parallel computing systems usually focus only on the common problems of a certain type of large data applications. These systems have a space for performance optimization when dealing with large data analysis applications with other characteristics. For example, Hadoop is mainly designed for off-line batch data applications, and it is more focused on high throughput rather than low latency. Therefore, Hadoop has low performance in dealing with large data flow computation problems.

2.3 Parallel algorithm for large data analysis

The basic algorithms of large data analysis (such as machine learning and data mining algorithms) play an important role in the industry large data analysis application and intelligent service, and they are the key technologies of many large data analysis applications landing. However, many traditional machine learning and data mining algorithms have many technical challenges in dealing with big data. Machine learning and data mining algorithms can work effectively when data sets are small. But when the data scale is hundreds of TB scale or PB scale, the time overhead of the traditional serialization algorithm is very long, which makes the algorithm not work in the actual scene. Therefore, in addition to finding new algorithms, such as low computational complexity and reducing data scale, an important method is to study parallel algorithms for smart large data machine learning and data analysis.

There is no standard and unified method for parallel design of big data machine learning and data analysis algorithms. Instead, we should carry out specific parallel optimization design based on specific algorithm. The parallel design of simple machine learning and data analysis algorithms is relatively easy, but the parallel design of complex machine learning and data mining algorithms is relatively difficult.

The neural network structure of the three-level perceptron is shown in Figure 2. It contains an input layer, a hidden layer and an output layer. Neurons in the same layer are not connected to each other, while neurons between two adjacent layers are connected to each other. Backward propagation algorithm using gradient descent technique is one of the most commonly used supervised training algorithms for training multilayer feedforward neural networks. The backward propagation algorithm consists of three stages: forward stage, backward stage, and generation stage.

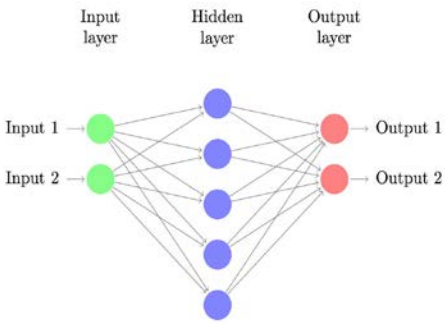

Figure 2 Schematic diagram of hierarchical storage structure of CPU- [MEM-SSD-HDD].

3. The Innovation of University Education Management in the Era of Big Data

The teaching materials developed in the traditional educational model are usually some teachers developing a more subjective research project through their own teaching experience materials, but there are also great limitations.

The response of the national education effect is unable to complete the true statistical data, so the traditional mode is restricted by the development. The big data age is changing this situation thoroughly. First of all, in the era of big data, we can deal with the existing materials very quickly through network investigation and statistics[10]. In this way, we can find out the merits and demerits of teaching materials. In the shortest time, these advantages and disadvantages are often more objective, and there is not much subjective consciousness. In fact, big data compiled by textbooks is only a side effect. The large data of university education management, not only for the statistical error often entered by the students, or for the influence of every teacher's education statistics, the whole educational management thought is changed in the data technology[11].

3.1. The problems of managing information in large data environment

While the information revolution can bring benefits to people's lives, it cannot be denied that it will also cause some negative effects and cannot be eradicated. All problems concerning information must not be solved all at once, nor can information be a flood. The profound influence of technology on education management and the negative influence of education on technology are all questions we should ponder deeply.

Nowadays, information technology brings very large and very easily accessible information, which makes many people no longer keen on investigation. Some managers download the rules and regulations of other institutions directly from the Internet while ignoring the actual survey, which is very common in the implementation of the rules and regulations of education. Teachers, students, courses, school roll, teaching materials, teaching, teaching websites and other information form educational management information[12]. With the support of modern information technology, especially the information system in education management, it can be open and interactive. Under complex procedures, the weaknesses and omissions of information and educational management systems make information extremely likely to be taken out at will. The problem of copying and intercepting are very common in the process of storage and transmission, causing information leakage and hidden security risks. Progress in the information age has given mankind too much information, but it is also a hindrance compared to too little information. Managers rely largely on information technology, they will lose the ability to independently explore problems, and will be divorced from reality.

3.2. The path of information management in Colleges and Universities under big data environment

University education management system needs to be reformed under information technology. In today's information age, the environment of schools is becoming more complicated and more diverse. This requires that the management mode of schools should be diversified and individualized.

Establishing the concept of data opening, data sharing, data transboundary and data cooperation is the premise for the healthy development of big data education and management in colleges and universities. In the application of IT planning and management, colleges and universities should highlight the high integration of people, people and resources, and develop a unified and ubiquitous platform, which can simplify the management task and make it easier for students to accept. The development of large data education management in colleges and universities involves system construction, platform building, management mode, and the construction of talent team. And the clear working principle is the prerequisite and guarantee for its successful development. The top-level design has the characteristics of long-term, strategic and scientific. The scientific big data development plan,

the perfect large data development mechanism and the democratic governance model are the important reasons for the success of the large data education management of University of Maryland, which is of great significance to the big data education management in our universities[13].

4. Conclusions and Future Work

With the advent of the era of large data and the development of large data technology, the development and innovation of educational circles are actively promoted, and better learning patterns and life patterns are provided to teachers and students in Colleges and universities. At the same time, a more excellent educational platform will be provided to educators, so that the goal of communication and sharing among educators can be realized.

Therefore, universities should use big data reasonably in all teaching links. In addition, in the era of large data, colleges and universities need to promote the reform of information based on the network platform, to further innovate management information and to strengthen the construction of network security. Through these measures, we can promote the reform of management informatization in Colleges and universities and better apply big data technology in the process of education and teaching.

References

[1] Hossam Hakeem.Layered Software Patterns for Data Analysis in Big Data Environment[J]. International Journal of Automation and Computing,2017,14(06):650-660.

[2] Call for Papers Special Issue on Cloud Computing and Big Data Applications[J]. Tsinghua Science and Technology,2017,22(03):342.

[3] Zhiqiong Wang,Junchang Xin,Hongxu Yang,Shuo Tian,Ge Yu,Chenren Xu,Yudong Yao.Distributed and Weighted Extreme Learning Machine for Imbalanced Big Data Learning[J].Tsinghua Science and Technology,2017,22(02):160-173.

[4] ZHANG Yaoxue,REN Ju,LIU Jiagang,XU Chugui,GUO Hui,LIU Yaping.A Survey on Emerging Computing Paradigms for Big Data[J].Chinese Journal of Electronics,2017,26(01):1-12.

[5] CHEN Jianping,XIANG Jie,HU Qiao,YANG Wei,LAI Zili,HU Bin,WEI Wei.Quantitative Geoscience and Geological Big Data Development:A Review[J].Acta Geologica Sinica(English Edition),2016,90(04):1490-1515.

[6] Xinqing Gao,Fan Yang,Chao Shang,Dexian Huang.A review of control loop monitoring and diagnosis:Prospects of controller maintenance in big data era[J].Chinese Journal of Chemical Engineering,2016,24(08):952-962.

[7] Huadong Guo,Lizhe Wang,Dong Liang.Big Earth Data from space: a new engine for Earth science[J].Science Bulletin,2016,61(07):505-513.

[8] Jari Yli-Hietanen,Antti Ylip??,Olli Yli-Harja.Cancer research in the era of next-generation sequencing and big data calls for intelligent modeling[J].Chinese Journal of Cancer,2015,34(10):423-426.

[9] LI DeRen,CAO JianJun,YAO Yuan.Big data in smart cities[J].Science China(Information Sciences),2015,58(10):179-190.

[10] Call for Papers Special Issue on Big Data Computing and Communications[J]. Tsinghua Science and Technology,2015,20(02):220.

[11] Tao Xu,Dongsheng Wang,Guodong Liu.Banian: A Cross-Platform Interactive Query System for Structured Big Data[J].Tsinghua Science and Technology,2015,20(01):62-71.

[12] Zhen Chen,Yuhao Wen,Junwei Cao,Wenxun Zheng,Jiahui Chang,Yinjun Wu,Ge Ma,Mourad Hakmaoui,Guodong Peng.A Survey of Bitmap Index Compression Algorithms for Big Data[J].Tsinghua Science and Technology,2015,20(01):100-115.

[13] Changjun Jiang,Zhijun Ding,Junli Wang,Chungang Yan.Big data resource service platform for the internet financial industry[J].Chinese Science Bulletin,2014,59(35):5051-5058.