Application of Big Data in Online Education in China

Lingling Dong[1, a], Jiachen Han [1, b] and Xiaohui Zhang[2, c]

[1]School of Distance and Continuing Education, Dalian University of Technology, Dalian, China

[2]school of computer and information technology, Liaoning Normal University, Dalian, China

[a] network1@dlut.edu.cn, [b] hanjiachen@mail.dlut.edu.cn, [c] zxh928@163.com.

Keywords: Online education, big data, application.

Abstract. Big data is a new type of technology derived from the development of information technology, bringing people new ideas as well as knowledge and leading them into a new era, the age of big data. Despite the new challenges it puts forward, this new era also provides new opportunities to online education. This paper explains the definition, features and technical framework of big data, analyzes the model of big data in online education, discusses problems occurred during the application of big data in online education, and proposes methods and solutions to solve these problems.

1. Introduction

Online education is a method that facilitates the spread of knowledge and the development of fast learning through technologies related to information and internet. Online education platforms store enormous complex data. Taking a platform of MOOC (Massive Open Online Courses) as an example, over 160,000 people from more than 190 countries signed up for the course of Intro to Artificial Intelligence lectured by professors from Stanford University on the platform of "Udacity" (https://www.udacity.com) in 2011. The number of users on this platform has exceeded 10 million by the October of 2014. On "Xuetangx" (http://www.xuetangx.com), another MOOC platform built by Tsinghua University, over 11,000 students from more than 150 countries are learning the course of Principles of Electric Circuits [1]. In the light of enormous complex data from different courses and users, research on application of big data in online education is of great significance.

2. Definition, features and technical framework of big data

Big data usually refers to data sets with the size beyond the capability of commonly used software tools for capturing, curating, managing, and processing data within a tolerable elapsed time [2]. In 2012, Gartner updated its definition as follows: "Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization." [3] Features of big data are listed in Table 1 [4].

Table 1. Features of big data.

Features	Content
High Volume	The storage volume of data is so huge that it can reach the level of EB (Exabyte), ZB (Zettabyte) or even YB (Yottabyte).
High Variety	Data are various in type with a full coverage of several forms of semi-structured, structured, etc., demonstrating a panoramic view of educational behaviors.
Low Value	Though the value of data can be high, its density is low. The potential value of data depends largely on the innovative methods and how they are used.
High Velocity	Processing velocity is high because massive results can be drawn from data analysis within a certain period of time at second level.

In the light of these features, the system of big data technology mainly focuses on the collection, transmission, processing and application of structured, semi-structured and unstructured data. It contains several stages, including collection and pretreatment of big data, storage and management of

big data, computing mode and system of big data, analysis and mining of big data, etc. Technical framework of big data is shown in Fig. 1 [5].
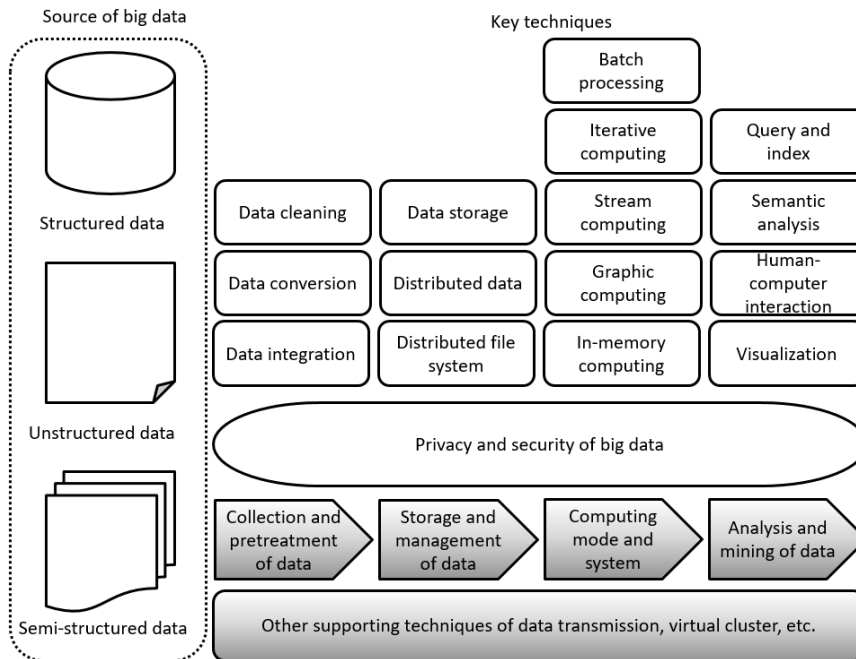

Figure 1. Technical framework of big data.

3. Analysis of the model of big data in online education

The model of big data in online education mainly contains three layers. They are application service, data processing and hardware facilities, as illustrated in Fig. 2 [6].
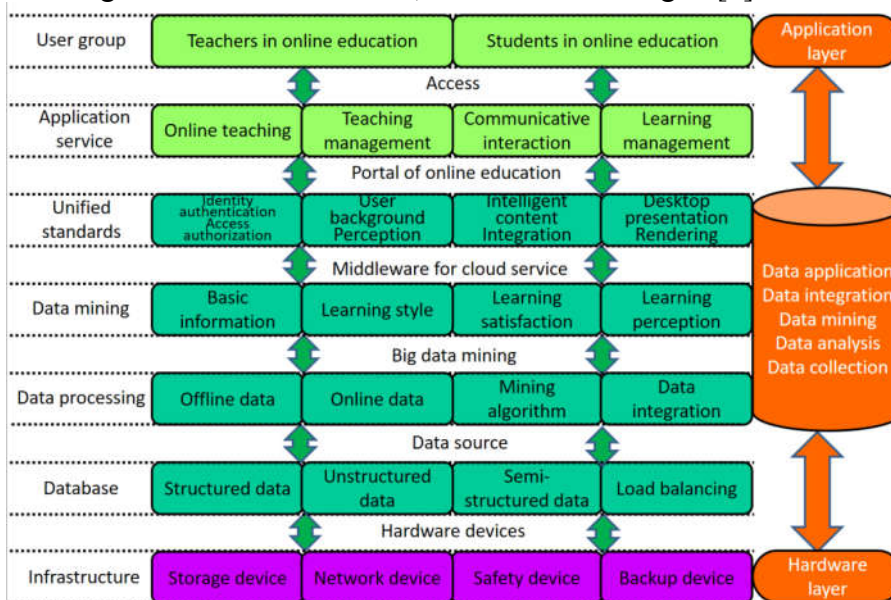

Figure 2. Model of big data in online education.

3.1 Application Service Layer.

Two user groups in online education are teachers and students. Application service layer mainly offers these two user groups with four types of service, i.e. online teaching, teaching management, communicative interaction and learning management. Different users have different user interfaces, access rights and services. In order to meet users' demands, this layer reconstructs relevant information and allocates all the information into different services. The integration of data is carried out in the background by data processing layer, which is unknown to users.

3.2 Data Processing Layer.

Data processing layer consists of database, data processing, data mining and standardized processing. Among them, data mining is the core. Data of students' learning behaviors are collected through big data technology. The procedure of data mining manages to dig out patterns of these data and establish a set of teaching models. By studying these models, students' learning experience and quality can be continuously improved.

3.3 Hardware Facilities Layer.

Hardware facilities layer contains storage device, network device, safety device and backup device. With the technology of cloud computing, centralized management of hardware facilities is achieved, energy consumption of hardware system is decreased, and reliability of the system and utilization of hardware devices are increased.

4. Problems occurred during the application of big data in online education

4.1 Problem of Processing and Converting.

Due to the feature of high variety for data and information, it is difficult to filter useful information within a limited time. In addition, if the information cannot be scientifically and effectively classified and managed, the safety and integrity of the information cannot be guaranteed since a large amount of information may inevitably contain illegal information or viruses.

4.2 Problem of Data Storage.

The application of big data in online education brings many advantages, but the storage space of online education platforms is not always able to hold all the information, of which the amount increases every day. During procedures of storage and exporting, excessive calculation may lead to a dead halt.

4.3 Problem of Information Security.

Against the backdrop of big data, many online education platforms lack measures for encryption and sound databases to store users' information. When computers are processing information obtained from the internet, personal information given by users during registration may be stolen by lawless people with the help of illegal information and viruses.

5. Improving methods for big data in online education

5.1 Adding Offline Operating Units.

For the purpose to store the information, of which the amount increases every day, necessary offline operating units need to be added. In this way, classification of information can be improved, filtration can be done according to the timeline, working efficiency of the backstage can be increased and management of online education platforms can be perfected.

5.2 Setting Web Blocking System.

As enormous data and information have many varieties, web blocking system can be set by using relevant technology. This system can automatically identify users' information delivered to the backstage. If the information is defined as illegal information or viruses, users are not allowed to enter the online education platform. In that sense, only effective information can be stored promptly and harmful information is blocked outside, guaranteeing the rapid and stable operation of the platform.

5.3 Improving Preventive Consciousness.

Apart from hardware technology, to solve the problem of information security also requires the strengthening of preventive consciousness of all the parties. Only by improving their preventive consciousness can students effectively avoid the disadvantages brought by the technology of big data. And only by strengthening self-management and formulating protection regulations for information security can online education platforms and institutions of higher education spread knowledge to those in need, thus achieving the goal of lifelong education [7].

6. Conclusion

The coming age of big data is an inevitable trend of the development of our society, though the application of big data in online education is still at its primary and exploring stage. With the deepening of people's understanding of big data, more scholars will be willing to adopt big data technology in online education and conduct relevant research to solve existing problems. Big data technology in online education will be more mature and refined, thus facilitating the sustainable and comprehensive development of online education.

References

[1]. L. L. Dong, Impact of the development of MOOC on the course of Advanced Maths, Proceedings of the 15th Forum on 21st Century Continuing Education. (2015) 247-249.

[2]. C. Snijders, U. Matzat, and U. Reips, "Big Data": Big gaps of knowledge in the field of Internet, International Journal of Internet Science. 7 (1) (2012) 1–5.

[3]. A. de Mauro, M. Greco, and M. Grimaldi, A formal definition of big data based on its essential features, Library Review. 65 (2016) 122–135.

[4]. Q. D. Xing, J. Jiao, and Z. H. Du, Application research of online education interaction based on cloud computing and big data, Modern Educational Technology. 4 (2014) 88-95.

[5]. H. B. Xu, Analysis of the application of big data in online education, Digital Technology & Application. 12 (2015) 85-86.

[6]. Q. D. Xing, J. Jiao, and Z. H. Du, Application research of cloud computing and big data in online education interaction, Journal of Information Resources Management. 3 (2013) 22-28.

[7]. Y. Li and F. L. Chen, Analysis on large data technology applied in network education, Wireless Internet Technology. 7 (2016) 41-42.