Research on Online Education Based on Big Data

Haifeng Li* and Hongrui Sun
School of Management Science and Engineering, Dalian Jiao Tong University, Dalian, China

*Corresponding author email: lihaifeng@djtu.edu.cn

Abstract. With the continuous development and innovation of global information technology, all walks of life are developing toward the direction of deep integration with information technology, improving the efficiency and quality of the industry level has become the general trend, education is no exception. With the analysis and summaries of big data’s application in online education at home and by means of literature review, the analysis clarifies the specific situation of the application of big data in China's network education. Finally, I put forward several solutions to the shortcomings of big data’s application in online education in China and predict the future development trend in order to make an all-around classification and analysis towards big data’s application in online education, so it can provide reference for better big data’s application in education.

Keywords: Big data; Online education; Problems; Solution proposals.

1. Introduction
With the comprehensive coverage of the network and the rapid development of large-capacity storage technology, data is growing and accumulating at an unprecedented speed. The advent of the digital era makes complicated data particularly important, and the development of big data promotes the rapid development of all walks of life. Especially in the online education industry, such as classroom flipping, moocs, micro-classes, etc., "large internet open interactive courses " based on new technologies have become a hot topic in the social and educational fields.

In the field of online education, due to the use of big data, from participants behavior analysis to the data adjustment, teaching supervision of teachers, teaching management and so on have more convenient way to do, the teachers will be liberated from many tedious teaching management work and engage in the more creative students guidance, which effectively promotes the scientific decision-making of education, improves the efficiency of education management, the innovation of classroom teaching mode, brings the network education "great knowledge", "big bang" and "great development".

2. Generation of Big Data
Big data comes into being on the basis of long-term data research and application, especially based on the in-depth application of mobile Internet, cloud computing, Internet of Things and other technologies. It is an inevitable outcome of the development of information technology in today's era. Big Data or massive Data refers to the amount of Data involved that is too large to be captured, managed, processed and collated in a reasonable amount of time by current mainstream software tools to help enterprises make more positive business decisions. At present, big data is no longer only used to describe the storage of large amounts of data, but also shows the high speed of processing massive data, especially it can help people objectively find the valuable information truth hidden behind
massive data. The application of big data in the field of online education makes it possible to deeply mine educational data and comprehensively grasp the development status of education. The characteristics of big data mainly include: (1) Volume: it mainly refers to the large storage capacity of data, and the unit of measurement even rises to the level of EB, ZB, YB and above; (2) Velocity: it mainly refers to the extremely fast speed of data analysis, which can generally provide results of mass data analysis within the time range of seconds; (3) Variety: it refers to a variety of data types, including website records, web logs, picture information, animation, video, geographical location and other information; (4) Value: It means that the focus of big data is to discover valuable information hidden behind massive data.

2.1.2 The characteristics of big data. "Digitalization of Everything." From macro to micro, from objective to subjective, from concrete to abstract, both direct and indirect activities and motivations are recorded in a comprehensive and real-time manner and become digital information[1].

3. Analysis of the Application Status of Big Data Network Education

As for the current situation of network education application of big data, based on consulting the famous literature data database in China, we mainly adopt literature research method to carry out the research, and the analysis is shown as follows.

With the help of the paper search on CNKI, the application research of big data in the field of online education can be understood. In the "advanced search" of the CNKI, we input "big data" in the "subject", "network education" and "online education" are included, we can find that as of June 14, 2020, a total of 12680 records associated with it, the year the paper was published dates back to 2012, and they have been substantial increasing year by year. Through the analysis of important literature, big data in the network application research in the field of online education is roughly divided into the following three stages.

3.1. The Early Budding Stage (2009-2012)
The term "big data" became popular in 2009, but it wasn't until 2012 that scholars and experts began to formally study it, and there were only 5-6 related papers in this area. In July 2012, The Plan issued by the State Council explicitly proposed to support the development and industrialization of massive data storage and processing technologies, thus opening the era of big data. In the same year, Professor Junping Qiu from Wuhan University proposed the definition of big data in a conference: First, big data is a relative concept, and there is no strict standard to define how large the data is. Second, big data not only refers to a class of objects, but also represents a technical capability, which can collect, store, and analyze data on a large scale[2]. These studies and policies have made big data an important trend and laid a foundation for the development of Big data in China.

3.2. Mid-term Development Stage (2013-2016)
In August 2013, several Opinions on Promoting Information Consumption and Expanding Domestic Demand issued by the State Council emphasized the policy of "Building the big data industry chain, promoting the effective grafting of innovation and industry chain", which pointed out the direction of China's big data development. Due to the national policy support and the development of information technology, the application research of big data in online education has mushrooming rapidly, which is directly reflected in the doubling of the number and quality of research papers. The number of papers in this field has increased from a few to more than 100, and the total number of papers in the future is even higher.

In 2015, "Data Science and Big data Technology" was added to the undergraduate major, and some practical explorations have been started, such as diversified assessment of student development, discovery of the reasons behind grades, and improvement of class procedures. In addition, there are also related apps, such as taking photos to search for questions and correcting papers online. As Xiaowu Wang, the deputy director of the Central Audio-Visual Education Center said, "Education is moving towards the era of big data. Whoever can discover the data will win the future survival; Whoever can mine data will win the future development; Whoever can harness data and use it to provide personalized services will win the competition of the future[3]."
3.3. Outbreak Phase (2017-present)

Since 2017, big data has penetrated into every aspect of people's lives. In November 2017, China's Big Data Talent Training System Standards were officially released. Since then, 21 big data management agencies have been set up in at least 13 provinces. Big data has also become a hot major in universities, with 293 universities applying for undergraduate majors in data science and big data. From 2017 to 2019, nearly 2,000 papers were published annually. The research scope is expanding and the level of applied research in education is more deeper than before. The main research focuses on the role of big data in the reform and innovation of online education.

4. Analysis of the Main Contents of the Application of Big Data Network Education in China

4.1. Big Data Promotes the Development of Personalized Education and Realizes "Teaching Students According to Their Aptitude"

Yannan Zhang pointed out that learners in the era of big data leave a lot of digital fragments in the process of digital learning. By analyzing these digital fragments, we will find out learners' various learning behavior patterns[4]. QiuDan Xing pointed out that the high-speed real-time processing technology of big data can provide real-time insight into the changes of learners on the online education platform, grasp the needs of learners, and improve the learning effect[5]. Through the analysis of the behavior of students, it can also find the existing problems and enable the educators to guide students in a targeted way. The English teachers of the Beijing Academy of Chinese Academy of Sciences use the intelligent synchronous teaching platform to continuously evaluate the data of students’ learning behavior and learning methods, and then provides personalized learning content and guidance for different learners, Making the teaching model change from a factory model with unified teaching standards to individualized teaching based on students' abilities.

4.2. Big Data Provides Scientific and Objective Data for Educational Evaluation

With the advancement of educational informatization, digital learning has become a normal way of learning for today's learners. With the support of learning terminals, learners produce a large number of digital learning records in various learning systems. Educational big data technology can track and pay attention to learners' learning process conditionally, providing us with the most direct, most objective and most accurate basis for educational evaluation and learning analysis. Shunping Wei pointed out that the learning analysis technology based on big data can be used to predict and optimize the learning process by storing and analyzing the process data of learners' learning situation, and provide an important basis for teaching decision-making[6].

At the same time, teachers can also use the information provided by big data to analyze their own teaching behaviors. Through the data reflected in the teaching process, teachers can find their own teaching strengths and shortcomings so as to change their own behaviors to improve the teaching quality.

4.3. Big Data Realizes Perfect Online Interactive Activities of Network Education

In order to balance the imbalance of resource allocation, the early online education in China simply copied the learning content and learning materials of traditional classroom courses to the Internet, making the Internet a channel for knowledge flow. Liu Tao pointed out that the only difference between such online education and traditional education is the digitalization of learning materials and the networking of learning interfaces. The lack of "personalized" online courses leads to a greatly reduced learning effect of online education, and the efficiency is lower than that of traditional education. This dilemma can only be improved if the big data provides the personalized online services, and this will form a new "mentoring relationship" between teachers and students. This model enables teachers to fully understand the level and learning ability of each learner, which is obviously the most effective way to learn[7]. Qiudan Xing pointed out that online education interaction can be divided into individual interaction and social interaction. The former is the interaction between learners and learning materials, while the latter is the interaction between learners and teachers or
learners themselves, social interaction plays a key role in improving the level of online education interaction[8].

5. Problems Faced by Big Data in Network Education

Although the development and utilization of big data of network education in China has begun to take shape to a certain extent, its application and promotion still face many difficulties due to its uniqueness and complexity.

5.1. The Amount of Collected Data is Small and Not Real-time

At present, a lot of education big data are not collected in real-time, but direct result-oriented data. On the other hand, compared with the current scale of online education in China, the existing data volume of third-party online education institutions is far from enough to meet the demand of in-depth mining.

5.2. Poor Data Integration Ability

Massive network education data not only needs to be mined and analyzed within the system, but also needs to be integrated across boundaries. Although the Chinese government has issued the national big data development plan for education, it is still not insufficient in terms of technology and talent reserve. We must step up the layout of educational big data research projects, vigorously develop and train a number of professional institutions and professionals, so that we can better develop data analysis technology and combine technology with practical education problems.

5.3. Data Governance Issues

In the era of the Internet, huge amounts of online education data are generated every day. China's big data online education is still in the early stages. How to cooperate with all parties to build an advanced big data governance system, create a positive and healthy online education big data industry environment, and handle multiple aspects Multi-type, multi-source educational data, while ensuring the security of data is a matter that we must comprehensively measure and consider systematically.

5.4. Data Security Lag

In recent years, students' information data is stolen frequently. On the one hand, reasons are that data protection isn't in place and managers' security awareness is weak. On the other hand, the current data processing system structure can no longer meet the requirements of the current era for security, so the improvement and upgrade of the architecture must be accelerated.

5.5. Inadequate Data Operation

Educational data is an intangible and valuable asset, and its ownership and use rights to use should be protected. In some ways, the security of online education data is as important as financial data. Data operation should encourage more government departments, enterprise subjects and relevant individuals to join in. At the same time, we should pay attention to the scope, objects and degree of openness, strengthen data supervision, and deal with objects and behaviors that endanger the security of online education data in accordance with the law.

6. Big Data in Online Education Solutions

6.1. Change the Philosophy of Education Management

Under the model of big data education, due to the adoption of new technology and the transformation of education model, education management should also be transformed. Both the educational management system and structure at the macro level and the system at the micro level, such as the final evaluation, teaching methods, the system of entering higher education and the professional title system, should be reformed to adapt to the precision, refinement, and humanization of the big data era. For example, schools should set up special functional departments to manage their education data, conduct accurate analysis of past historical data and newly generated data, and formulate teaching management strategies scientifically. In addition, they should strengthen the connection between
school education and family education, and attach importance to the role of family education. Only by changing the concept and thinking of education management can we better ensure the collection of data and promote the effective implementation of data governance.

6.2. Pay Attention to the Importance of Educational Data
On the one hand, the role of big data is fully utilized in each link of teaching activities, and the educational data is always in a state that can be processed and stored, so that the content of teaching activities can be reasonably adjusted according to the data analysis results. On the other hand, through the analysis of educational data, we can predict potential problems in teaching activities, so that new thinking and methods can be summarized in time. Although the value of educational data is high, we cannot rely on it completely, it is also crucial to focus on data error rate and cost.

6.3. Enhance the Important Role of Educational Technology
The large amount of collected data includes structured and unstructured data. Accurately finding value data is an important task, it is the basis for achieving technological transcendence. Technology transcendence is mainly reflected in two aspects: on the one hand, we should apply new technologies to fully liberate educational productivity on the basis of ensuring teaching efficiency and learning effectiveness, constantly develop new educational technology and establish sharing mechanism so that its function covers more and more fields. On the other hand, digging out old technology and actively promoting the development of educational productivity, data analysis technology is the most important. Horizontal connection, vertical upgrade, data activity and data popularity are the determinants of the development of educational productivity.

6.4. Strengthen Effective Data-based Management
The government should formulate and improve relevant laws and regulations at the institutional level to ensure the safety of educational data in legal form. On this basis, the third party supervision organization is introduced to supervise the data center, to ensure the transparency and openness of the operation of the data center, and to ensure the security of the data center under clear standards and strict requirements. At the same time, the government should increase publicity, creating a social atmosphere that respects data security, so as to make people consciously safeguard the legitimate use of data. The teaching management platform is the center of decision-making, control, and evaluation. It should also strengthen the security of data and make it used reasonably.

6.5. Strengthen the Construction of Professional Personnel
In order to obtain data in a timely and accurate manner, it is necessary to have necessary technical support and analysis personnel. The state should strengthen the training of professional and technical personnel, and offer big data-related courses in universities, so as to enable students to have relevant professional skills. Data institutions should strengthen the professional orientation, guide and strengthen the innovation of big data technology, set up reasonable data application standard system, and improve the ability of relevant workers to collect, process and analyze data, so as to cultivate data sensitivity and critical thinking. Only in this way can we effectively improve the governance of big data and promote the further integration and development of big data and online education.

7. Prospect of Future Research Trends
Through the above analysis, we find that although the application of big data has obvious advantages in promoting the development of network education and it is also the inevitable trend of future education development, most of the studies still focus on the theoretical discussion of the influence of big data, and the specific application practice is not mature. The practical application case is few. At the same time, big data is new to the whole field of education, and we also need to learn from advanced foreign ideas and research experience to explore and find in the future research practice. With the development and improvement of cloud computing, mobile Internet, Internet of Things and other technologies, it is believed that the application of big data in the field of education will be more and more extensive and in-depth, and relevant research will be more and more comprehensive, China's
education industry will continue to advance towards the direction of modernization with the support of big data technology.

Acknowledgments

Natural Science Fund of Education Department of Liaoning Province, JDL2019027; Liaoning Natural Science Foundation, 201800177; The subject of educational science planning in Liaoning Province, JG16DB054

References

[1] Big Data[OL] < http://baike.baidu.com/view/6954399.htm.>
[2] Junping Qiu, Wei Zhou. Progress and trend of information management in China [J]. Evaluation and management, 2012, 10(04):16.
[3] Xiaobo Wang. Big Data Promotes Education Reform and Innovation -- An interview with Wang Xiaowu, deputy Curator of Central Audio-Visual Education Center [J]. Information Technology Education for Primary and Secondary Schools, 2013(10):10-11.
[4] Yannan Zhang, Zhongjian Zhao. Enlightenment of thinking mode on education in the era of big data [J]. Research on educational development, 2013, 33(21):1-5.
[5] Qiudan Xing, Jing Jiao, Zhanhe Du. Research on online education interaction in cloud computing and big data environment [J]. Journal of information resource management, 2013, 3(03):22-28.
[6] Shunping Wei. Learning analysis technology: the value of mining educational data in the era of big data [J]. Modern education technology, 2013, 23(02):5-11.
[7] Tao Liu. Online education: from "move" to "mentoring" era [J]. Journal of shandong industrial technology, 2013 (7) : 106-107.