Research on the Application of AI in the Field of Education
Big Data Mining

Wenzheng Zhao\textsuperscript{1*}, Zaiyun Wang\textsuperscript{1} and Lianqing Zhu\textsuperscript{1}

\textsuperscript{1}Department of Electronics and Communication, Shandong College of Information Technology, Shandong, China, 261061

*Corresponding author e-mail: zhaowz@sdcit.edu.cn

Abstract. With the global summit on AI and education Big data (hereinafter referred to as BD), the deep integration of AI and education will bring more opportunities. Among them, deep learning is a hot issue in the field of AI, which will also become the key to the development of education. The integration of AI and intelligent education has become the inevitable trend of the development of education informatization. With the rise of education BD, we can analyze and accurately predict education-related problems through BD, which is a new topic in the era of AI. By integrating with machine learning, we can meet the needs of BD analysis and prediction in education. Therefore, AI has many applications in the field of education big data mining (hereinafter referred to as BDM). First of all, this paper analyzes the important role of AI in the field of education BDM. Then, this paper puts forward the relevant process. Finally, some suggestions are put forward.

Keywords: AI, Intelligent Education, BDM

1. Introduction
In recent years, with the full penetration of AI, the field of education is facing unprecedented challenges. UNESCO is the world's most authoritative large-scale international educational and cultural organization, which has always attached importance to the forward-looking research of education, science and culture, especially "AI + education" [1]. In 2016, UNESCO released the preliminary report on robot Ethics (Draft), which specifically discusses the manufacture and use of intelligent robots. Therefore, great changes have taken place in the field of Education under the influence and penetration of AI. Through education BDM, we can calculate and obtain accurate and clear forms of educational, psychological and social knowledge, which are often implicit [2].

At present, a large number of educational AI systems are applied in schools, which integrates educational AI and educational data mining (EDM) technology. Through machine learning algorithms, we can track student behavior data, which will predict their learning performance to support personalized learning [3]. At the same time, with intelligent education leading the development of education informatization, intelligent education drives the innovation and development of education and teaching, which has become an inevitable trend in the information age. Personalized learning is the core element of smart education, which will better support and promote the development of
personalized learning [4]. At present, the practical application of personalized learning mainly focuses on self-paced learning, individualized guidance and self-adaptive learning content.

2. The important role of AI in the field of education BDM

2.1. Realize individualized education

When education enters the era of popularization, the state gradually improves the accessibility of education and the fairness of educational opportunities, which ignores the individual needs of students. Therefore, the school gradually adopts standardized teaching methods. However, in the era of BD, the education field can meet different individual needs, which will achieve large-scale personalized. Through the integration of AI and BDM technology, we can promote education equity to a greater extent. Large scale open online courses bring learners to education [5]. Therefore, learners' learning behavior and process on MOOC platform can be collected and recorded, which will generate massive online learning data. Through the integration of AI and BDM technology, we can achieve personalized education, which can meet the individual needs of learners. Therefore, through education data mining, we can master the learning situation and learning analysis technology of learners, which is the key element to achieve large-scale personalized education.

2.2. Teaching intervention based on accurate prediction

Through large-scale open online courses, learners' learning behaviors will be recorded in the database of the system in real time, including video course operation, exercise, test, group discussion and forum interaction. A small amount of learning behavior data seems to be disordered. However, when the data is accumulated to a certain amount, the group behavior reflected by the data will show some order and law, which will reflect the development trend of group users. When the future trend presents potential problems, we can timely attract the attention of teachers and education managers, which can take early intervention measures and prevent the evolution of problems. At the same time, the teaching intervention based on BD prediction results will be based on the effectiveness of learning, including the overall level, development trend and individual differences [6]. BD prediction is good at capturing learners' online activity trajectory. By analyzing and mastering the law of learning behavior, we can predict the potential learning consequences, which will achieve the best teaching effect.

2.3. Improving learning support based on visual data

Data visualization technology has the characteristics of visualization, relevance, artistry and interactivity, which can intuitively analyze multiple attributes or variables of objects or events, and can classify, sort, combine and display according to the values of different dimensions. Data visualization technology is a very important link in the process of BD analysis and processing, which is the treasure map of BD. Through data visualization technology, we can achieve data collection, management, processing and analysis in a short time, which will present the information and value contained in BD in a simple form. Through the visual representation of data, we can extract and extract various attributes and variables from the corresponding information units, which will interpret the data information in a visual way, including graphic image processing, computer vision, user interface, etc. Data visualization focuses on visual expression, interaction mode and user's psychological perception, which can simplify and concretely process complicated, fuzzy and abstract information. Through the form of visual expression, we can promote users to grasp and internalize the connotation of information.

3. Educational BDM process based on AI

3.1. BDM process

Education BD is complex and diverse, which can be obtained from the activities of different objects in various educational environments. The workflow of education BDM can be divided into education BD
acquisition, preprocessing, data mining, visualization, etc., as shown in Figure 1. In the process of educational data mining, machine learning mainly plays a role in data mining and interpretation, which can realize the functions that traditional education is difficult to complete manually. Through the automatic analysis of data, we can find new unknown knowledge and patterns. In the process of BD acquisition and preprocessing, the difficulty of data acquisition and preprocessing is gradually increasing. Complex unstructured data can be collected in various ways, such as Internet of things perception technology, video monitoring technology, intelligent recording and broadcasting technology, online learning and management platform technology. For different service objects, we need to preprocess the original data, which will be for the purpose of data generation and classification.

![Figure 1. BDM process.](image1)

![Figure 2. Education BDM process based on AI.](image2)
3.2. Education BDM process based on AI

Through deep learning, we can supervise the learning of educational data. By dividing the education data into training set and verification set, we can optimize the model parameters, which will improve the accuracy. Through clustering analysis, we can use unsupervised learning to identify model rules and anomalies. Through semi supervised learning, we can reduce the cost and improve the accuracy. By preprocessing the education BD, we can transform the education BD set into variables, which is the input of the deep learning model. Image data is mainly processed by some software and tools. Through imread function, we can transform the image into matrix vector, which can be used as the input of deep learning model. Speech data is extracted by acoustic features, which can transform each frame waveform into a multi-dimensional vector as input. The output vector of deep learning model is to transform the output vector of education BDM for the purpose and service object. We should select deep learning model according to the purpose and service object of education BDM, as shown in Figure 2.

4. Suggestions on the application of BD in education in China

4.1. Construction of educational informatization application platform

After years of education information construction, a large scale of data resources have been formed in the field of education. There are still a series of problems, such as lack of data openness and integration. Among them, education informatization application platform is mainly constructed by schools, education management departments at all levels and education training institutions, which is caused by access to educational data resources. With the deepening of the application of information technology, we need to constantly innovate teaching means, teaching environment, education mode, which will produce huge educational data in teaching activities. Therefore, we must build an education informatization application platform that conforms to the characteristics and concepts of the BD era, which requires effective means to improve the openness and integration of educational data and improve the availability of educational data resources.

4.2. Strengthen the intelligent function of cloud platform

BD is the foundation of intelligent platform. It is not only a technology, but also a way of thinking. With the continuous improvement of the openness of educational data and the increasing of unstructured educational data, the raw materials of educational data mining are becoming more and more abundant. Therefore, the value of education BD will become increasingly prominent. The collection of educational data should follow the principle of "comprehensiveness". We can collect more data as much as possible through sensing, positioning, logging and other technologies. In the application of educational data, we should follow the principle of "value maximization". Data mining and learning analysis technology is a bridge connecting education BD and smart education. Through the innovation of data processing and presentation technology, we can analyze the educational data from multiple perspectives and paths, which will fully tap the hidden value of educational data. The analysis methods include association analysis, cluster analysis, regression analysis, decision tree analysis, neural network analysis and so on. Therefore, we must strengthen the intelligent function of cloud platform, which will better improve the application of AI in education.

4.3. Policy guarantee

Government departments can refer to the field of communication to issue operator licenses, which will better develop access standards for educational data operators. By referring to the regulations of the Ministry of education, the national development and Reform Commission and the Ministry of Commerce, the government departments can formally supervise the education data industry. At the same time, the government should issue relevant documents for data standards in the field of education, which can formulate relevant measures for BD governance. We need to specify the basic collection standards and quality standards of educational data, which will better set the standards,
thresholds and qualifications for the use of educational data. If there are non-governmental organizations to study, we need to meet the corresponding conditions and obtain the certification qualification in order to obtain access to educational data.

5. Conclusion
Through the realization of intelligent education, deep learning technology has an irreplaceable important position. Machine learning can effectively help intelligent education, which can play an important role in educational data mining. With the continuous development of machine learning technology, the integration of education will continue to promote educational innovation. At the same time, we need to actively protect personal education data, which can fully protect the privacy of students and teachers.

References
[1] Jiang Xin, Ma haiqun. Evaluation method and practice of open government data. Modern intelligence: 2016, 36 (09): 22-26.
[2] Li Jiajia. New perspective of information management: open data [J]. Intelligence theory and practice: 2010, 33 (07): 35-39.
[3] Liang Jun, Chai Yumei, Yuan Huibin, et al. Micro blog sentiment analysis based on deep learning [J]. Chinese Journal of information technology, 2014, 8 (5): 155-161.
[4] Mou Zhijia. Cracking personalized learning path under the support of learner data portrait - the value of learning computing [J]. Journal of distance education, 2016 (6): 11-19.
[5] Xu Lin. opportunities and challenges: China's dual situation of state governance in the era of BD [J]. Social scientist, 2015 (5): 13-18.
[6] Zhu Zhiting. New development of Smart Education: from flipped classroom to smart classroom and smart learning space [J]. Open education research, 2016 (1): 18-26.