Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Massive AI based cloud environment for smart online education with data mining

Ying Pei[1]*, Gang Li[2]
1.Xi'an Siyuan University, Xi'an Shaanxi, 710038, China
2.Xi'an Thermal Power Research Institute Co., Ltd , Xi'an Shaanxi, 710043, China
Peiyingxiansiyuan@mail.com

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432201

***Abstract*— Under the background of the deep integration of Internet technology and artificial intelligence technology with the field of education, the traditional teacher centered teaching mode is facing a historic change. How to guide students to learn and communicate actively, and explore and improve the new student-centered teaching mode, has become the key problem to be solved. In the mixed cloud environment, the data stream is disturbed, and the error of mining association data is large. Aiming at the problem of poor anti-interference of scattered point cloud adaptive compression mining algorithm, this paper analyzes the data mining technology in cloud environment. Firstly, the time series analysis model of big data information flow in hybrid cloud environment is constructed to analyze the data structure, and then the high-dimensional phase space of data information flow in hybrid cloud environment is reconstructed. In the reconstructed phase space, the association rules are extracted, and the extracted features are used as pheromones to guide data location mining, so as to improve the data mining algorithm.**

Keywords*— *Online Education; Artificial Intelligence; Deep Learning; Cloud Computing; Data Mining

I. INTRODUCTION

Information technology, such as Internet, mobile communication, cloud computing, big data and artificial intelligence, has a revolutionary impact on the field of education. China's education informatization has experienced the digital era represented by audio-visual education and the network era represented by distance education and online education. At present, it is moving towards a new stage of intelligent education characterized by intelligence, ubiquity, personalization, openness and collaboration. July, 2017 In June, the State Council issued the "development plan for the new generation of artificial intelligence", which pointed out that "using intelligent technology to accelerate the reform of personnel training mode and teaching methods, build a new education system including intelligent learning and interactive learning, carry out the construction of intelligent campus, and promote the application of artificial intelligence in the whole process of teaching, management and resource construction", marking the beginning of China's intelligent education The beginning of generation [1-5]. In April 2018, the Ministry of Education issued the education informatization 2.0 The action plan puts forward the requirement of "based on emerging technologies such as artificial intelligence, big data and Internet of things, relying on all kinds of intelligent devices and networks, actively carrying out innovative research and demonstration of intelligent education, and promoting the reform of education mode and ecological reconstruction under the support of new technologies", which further promotes the renewal of education concept, the reform of teaching mode, and the reconstruction of education system, It accelerates the transformation and upgrading of education informatization to intelligent education. At present, there is no strict definition of smart education at home and abroad. Singapore has pointed out the characteristics of smart education in the iN2015 plan, namely ubiquitous learning, interactive digital learning resources, and intelligent learning experience adapting to different learning styles [6-10].

Natural language processing (NLP) is an important branch of artificial intelligence. It is a bridge for computer to understand human language and realize natural and fluent human-computer interaction. With the support of NLP technology, intelligent question answering robot has attracted more and more researchers' attention because it conforms to human natural interaction habits and can effectively deal with imprecise information interaction. Especially with the development of knowledge graph technology and the promotion of NLP technology, the intelligent level of question answering robot is significantly improved. With the support of knowledge map, the intelligent question answering robot transforms the questions raised by the questioner in natural language into the sentences that can query the knowledge map by NLP technology, and then uses the structured data in the knowledge map to complete the knowledge query, and then feeds back the query knowledge to the questioner as the answers, so as to provide accurate and concise answers for the questioner. In the field of higher education, the application of intelligent question answering machine has covered engineering, English, medicine, psychology, education and other disciplines, fully reflecting the learner centered learning concept of personalized learning, cooperative learning, incentive learning and so on [11-15].

In the past, the enlightenment education is a face-to-face teaching for children. This kind of teaching form has high price, fixed teaching time and place, and often causes children to be unable to keep up with the teaching progress for various reasons. Online learning platform effectively solves this situation, children can be accompanied by their parents anytime and anywhere to receive enlightenment education, and the tuition is cheap, can learn again and again. Japan is the first to develop an online learning platform for enlightenment education, using browser / server structure. There are three login identities in the platform: children, parents and administrators. Administrators set up the enlightenment

education courses, parents select the courses, and children realize sensory interaction according to the course content. This online learning platform has perfect functions and obvious effect of enlightenment education, but it needs to use the platform's exclusive electronic facilities and wired network, so children can only learn at home or in the classroom. This paper analyzes the framework of intelligent online education, and studies the network education mode based on artificial intelligence and data mining technology.

II. The Proposed Methodology

A. Design of Big Data Clustering Mining Algorithm based on Swarm Intelligence Algorithm

Computational intelligence is an important branch of artificial intelligence. Due to its heuristic and stochastic characteristics, computational intelligence is very suitable for solving large-scale optimization problems. The traditional optimization algorithm mostly adopts the centralized design idea, which mainly considers the convergence and convergence speed of the algorithm. When the problem to be solved is very complex or large scale, especially the problem faced by big data, the traditional centralized optimization algorithm can not deal with or the calculation is very time-consuming. The increasing scale and complexity of data brings new challenges to traditional computational intelligence algorithms. Therefore, it is necessary to study distributed optimization algorithms to solve the problems of big data optimization.

The ant colony algorithm in swarm intelligence algorithm is used to treat the cluster center as the food source of ants, the whole data set as the ants looking for food, and the clustering mining process as the process of ants looking for food source [16-20]. The data set is as follows:

$$Q=\{Q\mid q_{i1},\ldots q_{in}\}ni=1,2,\ldots m \tag{1}$$

The pheromone between all data is set to a constant, and the pheromone update formula between different data objects is set. By updating the formula, the data objects are merged into the cluster center to generate a data set which is merged into the Q domain. The specific calculation formula is as follows:

$$\Delta\lambda_{ij}=\frac{A}{l(i,j)+B} \tag{2}$$

$$\Delta\lambda_{ij}(t+1)=\kappa\lambda_{ij}(t)+\Delta\lambda_{ij} \tag{3}$$

$$C_j=\frac{1}{N}\sum_{i=1}^{N}Q_i \tag{4}$$

In ant colony algorithm, the initial data is usually randomly selected, which is prone to uneven data distribution, which affects the overall performance of the algorithm [21-24]. Therefore, the method based on density and maximum minimum distance is used to update the new data in clustering.

B. Big Data Analysis based on Swarm Intelligence

Swarm intelligence is an important method for massive, high-dimensional and dynamic big data analysis. Distributed implementation based on swarm intelligence algorithm is a branch of big data distributed optimization. Because distributed computing environment can accelerate the search process of optimization algorithm, it plays an important role in big data optimization. At present, the big data analysis methods based on swarm intelligence are mainly based on particle swarm optimization algorithm. In addition, there are also big data analysis methods based on ant colony algorithm, cuckoo algorithm, firefly algorithm, cat swarm algorithm and so on.

Similar to particle swarm optimization, other swarm intelligence algorithms can be used for big data analysis. At present, it mainly focuses on the distributed implementation of swarm intelligence algorithm in MapReduce programming framework, which is used to improve the efficiency of the algorithm and speed up the solution speed. When the amount of data has a certain scale, the convergence speed of swarm intelligence algorithm based on MapReduce is better than that of non distributed implementation.

Ant colony algorithm is a kind of parallel algorithm in essence, which has strong global search ability. Due to the slow convergence speed of ant colony optimization (ACO) and easy to fall into a standstill state, Cheng et al. Proposed a dynamic positive and negative feedback ACO algorithm, which adopts the strategy of internal positive feedback and negative feedback between groups, and is implemented on haloop, a framework constructed by iterative MapReduce model. Wu Hao et al. N081 introduced divide and conquer strategy and simulated annealing algorithm into ant colony algorithm, and proposed an ant colony algorithm based on MapReduce, which improved the ability of ant colony algorithm to deal with big data. Ma Wenlong proposed an improved ant colony algorithm to solve the dynamic service composition optimization problem of cloud manufacturing, that is, the selection mechanism of optimal path list and roulette was introduced into the ant colony algorithm.

Swarm intelligence algorithm is very suitable for big data distributed processing because of its natural distribution characteristics. Especially with the emergence of a number of new swarm intelligence algorithms such as quantum particle swarm optimization algorithm, the optimization performance is getting better and better. Therefore, the big data analysis method based on swarm intelligence algorithm will get more and more research and application.

C. Construction of Online Education Platform based on Artificial Intelligence Robot

A question answering robot is an intelligent system that uses natural language to answer the questions input by the questioner. After the questioner raises a question in the form of natural language, the intelligent question-answering robot gives a short and accurate answer to the questioner in the form of natural language through three steps: question analysis, information retrieval and answer extraction.

The teacher first constructs the curriculum knowledge map through the knowledge map management system, organically organizes the curriculum knowledge system and teaching resources based on the curriculum knowledge map, and assists the answer generation and learning resource recommendation

of the intelligent question-answering robot. Students ask questions to the intelligent question-answering robot and get the feedback answers from the intelligent question-answering robot to realize autonomous learning. Through the analysis of Q&A logs, the system will recommend appropriate learning resources to students. At the same time, according to the learning situation of students, online examination topics are selected to urge students to consolidate the learning content. Students ask questions to the intelligent question-answering robot, and then learn independently after obtaining answers and learning resources. Finally, they complete the online exam to feedback the learning situation, so as to form a closed loop of students' self-study and realize the improvement of students' independent learning and learning quality. On the other hand, through the analysis of Q&A logs, students' learning can be tracked from the macro and micro levels, and the weak links in students' learning can be found. Combined with the online examination system, the objective evaluation of the teaching quality can be realized, and the evaluation results can be fed back to the teachers to help the teachers find and correct the deficiencies in the teaching process, and give targeted guidance to the students, so as to realize the continuous improvement of the teaching quality.

III. EXPERIMENT

The data set used in the experiment is stored in the MYSQL database, and the data used is from a large website. The data format of the data set is ISBN, and the data in the MYSQL database is called. The data display interface is shown in the figure below.

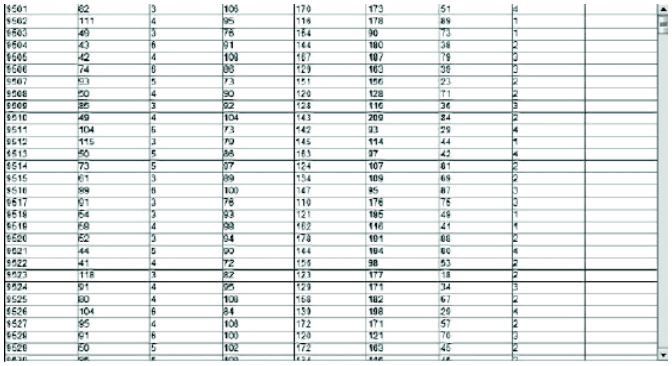

Fig. 1. Experimental data display interface

A comparative experiment is designed to verify the parallel efficiency of the designed big data clustering algorithm and the traditional big data clustering algorithm under the same data condition.

In the established experimental environment, the data set prepared above is used to test the parallel efficiency of different clustering mining algorithms under the same experimental environment. Three groups of data were randomly selected to display their experimental results, as shown in Fig. 2.

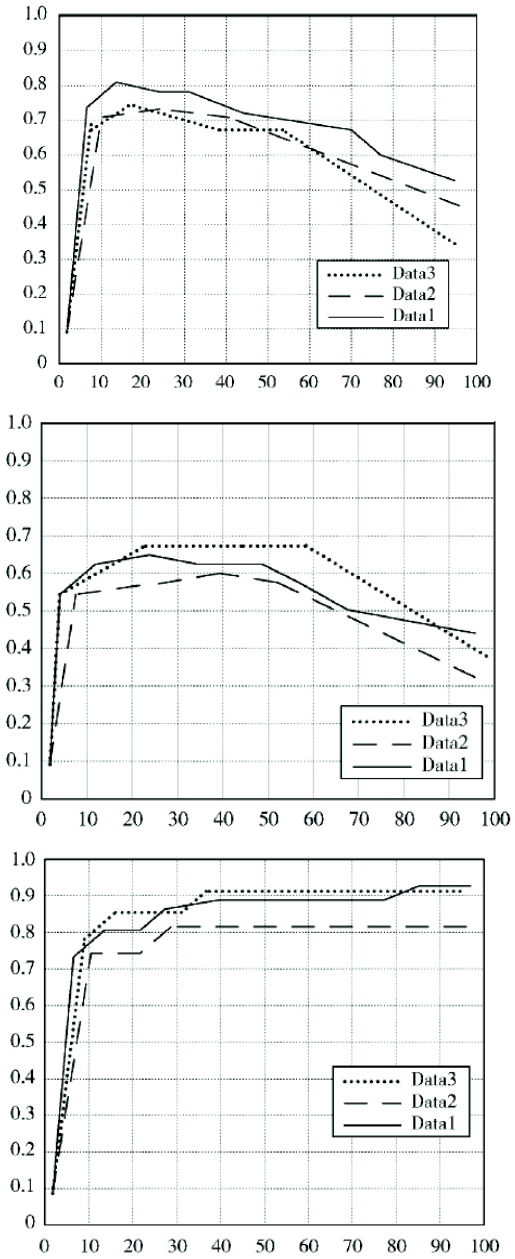

Fig. 2. Experimental results of different mining algorithms

IV. CONCLUSION

In order to improve the data orientation and mining ability in the mixed cloud environment, a data flow mining algorithm based on phase space reconstruction of data information flow and association rule feature extraction in the mixed cloud environment was proposed. The time series analysis model of big data information flow in the mixed cloud environment was built, and association rule features were extracted in the reconstructed phase space. The extracted features were used as pheromones to guide data location mining, and the improvement of data mining algorithm was realized. This paper constructs a set of intelligent education platform and explores an online teaching model based on intelligent question-answering robot. Based on the intelligent question-answering system, students can ask questions, discuss and

study with each other, thus stimulating the ability of students to study independently. By participating in the Q&A, teachers can timely understand the learning situation of students, grasp the depth of students' understanding of the problems, find out the difficult problems, analyze the deficiencies in the teaching process, so as to adjust the teaching content and teaching methods in time. With the help of this intelligent question-answering system, it is helpful to form a new teacher-student cooperative teaching mode, and realize the transformation from teacher teaching to teacher guidance, students learn from each other and cooperate in learning.

V. ACKNOWLEDGEMENT

Project of Shaanxi Higher Education Society in 2019 "Research on the construction of university innovation and entrepreneurship ecosystem based on makerspace" （XGH19230）.

REFERENCES

[1] Zheng Qinghua, Dong Bo, Qian buyue, et al. Research status and development trend of Smart Education [J]. Computer research and development, 2019, 56 (01): 213-228

[2] Notice of the State Council on printing and distributing the development plan of new generation artificial intelligence (GF [2017] No. 35) [Z]. 2017-07-20

[3] Ministry of education. Notice on printing and Distributing "education informatization 2.0 action plan" (JJ [2018] No. 6) [Z]. 2018-04-18

[4] Innovation, integration, internationalization: Report by the iN2015 Steering Committee[J].2018.

[5] Yang Bing, Yin Jiaqi, Yang Min, et al. Current situation and development: Reflection on Intelligent Question Answering robot promoting learning [J]. China audio visual education, 2018383 (12): 36-43

[6] Huang Hengqi, Yu Juan, Liao Xiao, et al. Review of knowledge mapping [J]. Computer system applications, 2019,28 (06): 1-12

[7] Huang Yan. Design and development of mobile learning resources based on smart phones [J]. Software guide, 2016, 15 (8): 99 - 101

[8] Wang Zhanmin. Web based English online learning system [J]. Computer system application, 2016, 25 (7): 96 - 100

[9] Zhao Xueming, Wang Gang. Research on interactive mobile learning platform based on HTML5 [J]. Modern educational technology, 2016, 26 (9): 106 - 112

[10] Wang Jianhu, Wu Wenxin. Comparative study of mobile learning app software generation platform [J]. Network new media technology, 2016, 5 (5): 16 - 21

[11] Ji Guangyong, Zhang Rui. Research on development technology of responsive online education platform [J]. Software, 2016, 37 (7): 138 - 141

[12] Hu Ying, Huang Yong, Chen Baoling. User experience design and implementation of educational software based on intelligent mobile terminal [J]. Computing technology and automation, 2016, 35 (3): 132 - 136

[13] Li Boyang, Han Shujie, Zhang Xiaorong. Hybrid teaching mode based on online learning and intelligent evaluation system [J]. Maritime education research, 2016, 33 (1): 47 - 50

[14] Wang Ting, Jin Tian, Li Shuai, et al. On the development of software simulation online learning platform [J]. Electronic Science and technology, 2016, 3 (4): 457 - 459

[15] Luo Chen, Han Jiabao, Luo Dapeng. Semi autonomous online learning target detection system [J]. Modern electronic technology, 2016, 39 (9): 121 - 125

[16] Shi Bo, he Chu, Zhuo Tong, et al. Topic interaction model based on local community discovery in MOOC Teaching [J]. Computer application research, 2015 (6): 1724 - 1727

[17] Yu Lou Cheng. Research on tracking methods of video moving objects in sports field [J]. Television technology, 2018, 42 (9): 74 - 79

[18] Wang Junpeng, Hou Xiaomao. Design of sports video target tracking algorithm based on optimized particle filter [J]. Machine tools and hydraulics, 2018, 46 (6): 164 - 169

[19] Liu Haoyang. Analysis and Prospect of sports coaching system based on artificial intelligence [J]. Journal of Beijing Sport University, 2018, 41 (4): 55 - 60

[20] Pei Songwen, Yang Baoguo, Gu Chunhua. Research on video stream classification based on fused 3D convolutional neural network [J]. Minicomputer system, 2018, 39 (10): 140 - 144

[21] Zhi Hongxin, Yu Hongtao, Li shaomei. Video classification based on two-level coding fusion of spatial-temporal depth features [J]. Computer application research, 2018, 35 (3): 926 - 929

[22] Dai Guanglin, Xu Mingmin, Dong Tianyang. Research on vehicle classification method of traffic video based on spatial pyramid visual bag of words model [J]. Journal of Zhejiang University of technology, 2016, 44 (3): 247 - 253

[23] Xu Tongde, Zhao Zhijun, Gao Junwen. High precision data mining algorithm for multi-layer cascade minority clustering [J]. Control engineering, 2018, 25 (5): 829 - 834

[24] Wen Jing, Cao Yan, Zhang Lin, et al. Research on clustering analysis algorithm based on dual genetic algorithm [J]. Computer engineering and science, 2017, 39 (12): 2320 - 2325