Study on Deep Unsupervised Learning Optimization Algorithm Based on Cloud Computing

Hui Yan[1]，Ping Yu[1]，Duo Long*[2]
[1]Changchun Institute of Technology, Changchun, 130012, China
[2]Jilin Province Soft Science Research Institute,Changchun, China
*corresponding author's email: longduo7113@126.com

***Abstract*—Big data has already occupied a lot in the information society. The application of big data to intelligent agriculture is the core development direction for maximizing the utilization of agricultural data information, and the deep learning method can more effectively extract abstract information from big data and convert it into useful knowledge, thus supporting the development of intelligent agriculture from different dimensions. In this paper, a CNN-RNN model is constructed based on cloud computing technology, and the parallel neural network model divided by training set is adopted to design the batch gradient descent algorithm based on deep unsupervised learning and BP algorithm based on Map-Reduce. An experiment verifies the feasibility of deep unsupervised learning neural network based on cloud computing and verifies that the optimize algorithm proposed in this paper can better increase the training efficiency of neural network.**

Keywords- Neural network; deep learning; unsupervised learning; cloud computing; big data

I. INTRODUCTION

With the rapid development of IT and communication technologies such as the Internet, Internet of Things, cloud computing and triple play, the rapid growth of data has become a serious challenge and valuable opportunity for many industries. The era of big data is coming. Agricultural big data makes intelligent agriculture more intelligent. Valuable information can be obtained quickly by analyzing agricultural big data to scientifically manage and regulate the agricultural environment, agricultural products, agricultural industrial chain and agricultural market and to achieve the purpose of optimizing the allocation of agricultural resources. The research and development of simulation sand table education system based on agricultural big data is one of the important application study directions of intelligent agriculture. The project structure diagram of the intelligent agriculture education system is shown in Figure 1.

This paper focuses on the core module of the back-end part of the intelligent agriculture education system project structure - the unsupervised learning neural network module, and studies how to use the effective artificial intelligence technology based on multi-computer cluster for big data analysis. The deep learning method is the key to opening this door. Unsupervised learning in deep learning algorithms can be adopted in more extensive fields than supervised learning. Therefore, it is of practical significance to study the deep unsupervised learning optimization algorithm based on cloud computing.

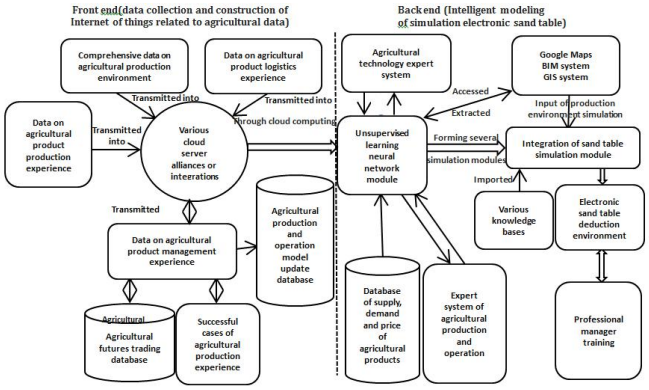

Figure.1 Project structure of the smart agriculture(Intelligent Agriculture) education system

II. CONSTRUCTION OF THE CNN-RNN MODEL

In the "big data" era, the proportion of image data is increasing. To better realize data processing, this paper constructs the CNN-RNN model based on the advantages of Convolutional Neural Network (CNN) in large-scale image feature representation, classification and extraction and changes of recurrent neural network (RNN) in time series. Because it is difficult to adopt RNN to achieve long-distance learning due to gradient divergence of it, Long-Short Term Memory (LSTM）is introduced.

The features of the image are extracted by CNN, and the coding and decoding network is constructed by LSTM. For the input features, the timing features are extracted at the encoding end and transmitted to the decoding end for decoding. In addition to learning historical information, RNN and LSTM can also be designed as two-way structures, i.e. bidirectional RNN, bidirectional LSTM, which is a good choice for speculating and completing information [1].

III. PARALLEL COMPUTING

Parallel Computing is an operational mechanism, including parallelization such as cloud computing cluster dealing with big data.

A. Neural network parallelization model

In image recognition or speech recognition applications, the computational complexity of the deep neural network model is very large, and there is a certain data correlation between layers of the model. Therefore, how to divide the task amount and computing resources is an important issue for the design of CPU or GPU cluster acceleration framework. When the scale of the trained model is relatively large, the training of the model can be accelerated by the data parallelism method [2]. Based on comparison of Several types of neural network parallelization modes, it is finally

978-1-7281-1307-4/19/$31.00 ©2019 IEEE
DOI 10.1109/ICITBS.2019.00168

determined that the parallelization of neural networks is realized by dividing the training set (data parallel).

B. Learning mechanism

In the deep learning algorithm, pre-training through unsupervised learning is adopted first based on each layer [3]. Deep network unsupervised learning is implemented through the gradient descent algorithm. The main task of deep network learning is to solve the parameter p of the loss function and to optimize the gradient descent algorithm for big data.

Loss function F

$$F = \frac{1}{2M}\sum_{i=1}^{M}\left(y^{(i)} - a^{L}\left(x^{(i)},p\right)\right)^{2} \quad (1)$$

Solving the parameter p in equation (1) is equivalent to solving the following optimization problem:

$$\min arg \frac{1}{2M}\sum_{i=1}^{M}\left(y^{(i)} - a^{L}\left(x^{(i)},p\right)\right)^{2} \quad (2)$$

And the optimization problem (2) is solved commonly by using the batch gradient descent algorithm.

The iterative formula of batch gradient descent algorithm [3] is:

$$p_{ij}^{l} := p_{ij}^{l} - \frac{\eta}{M}\sum_{k=1}^{M}\left(y^{(k)} - a^{L}\left(x^{(k)},p\right)\right)\frac{\partial a^{L}}{\partial p_{ij}^{l}} \quad (3)$$

C. Training mechanism

The batch training mechanism is adopted based on the training mechanism of neural network with Back-Propagation (BP) algorithm. The cumulative weight change is a very effective training method for the batch training mechanism of neural network. In the batch model training mechanism, calculation error is obtained by the following error calculation formula:

$$E_{sum} = \frac{1}{2}\sum_{q=1}^{q}\sum_{k=1}^{l}\left(d_{k}^{q} - o_{k}^{q}\right)^{2} \quad (4)$$

In the batch training mechanism, the weight updates are accumulated after submitting the entire training set sample. The batch training mechanism algorithm can be divided into three phases, namely the prefix phase, error BP phase, and weight update phase [4-6].

IV. CLOUD COMPUTING TECHNOLOGY

Cloud Computing is a computing model. Its main function is to realize distributed computing of big data, parallel operation of a large number of operations, storage of massive data of the network, virtualization sharing of network resources, resource scheduling and load balancing and data redundancy storage based on hot backup [7]. Hadoop is a cloud computing application software library for big data processing. It uses a Map-Reduce programming model and other mechanisms to perform distributed computing frameworks on large data sets to be processed by computer clusters. In the Map-Reduce computing mode, data calculation process is divided into two phases Map and Reduce.

A.Batch gradient descent algorithm based on Map-Reduce for deep unsupervised learning

The batch gradient descent algorithm is processed in Map-Reduce as follows:

1)Map-Reduce takes the input training sample data $\left[\langle X^{(i)},y^{(i)}\rangle\right]$ as the parameter for Map-Reduce processing module, and the Map-Reduce processing module divides the input data into M blocks, $\langle X^{(i)},y^{(i)}\rangle (i = 1,2,...,M)$ and distribute the data of each block to corresponding mapper;

2)Mapper executes $\sum_{i}\left(y^{(i)} - a^{L}\left(x^{(i)},p\right)\right)^{2}$ to $\left[\langle X^{(i)},y^{(i)}\rangle\right]$ and map the corresponding calculation result to $T^{(j)} (j = 1,2,...,M)$, and the calculation result is transmitted from the extension to a master host through the network;

3)Master host distributes the merged data transmitted by the extension to the reducer, and the reducer performs $\frac{1}{2N}\sum_{j=1}^{M} T^{(j)}$ on the merged data $T^{(j)} (j = 1,2,...,M)$ according to the function definition and Output of results to be reduced, that is, the output data $\frac{1}{2N}\sum_{j=1}^{M}\left(y^{(i)} - a^{L}\left(x^{(i)},p\right)\right)^{2}$.

B.BP algorithm based on Map-Reduce

The BP algorithm is divided into two parts on Hadoop . In the Map phase, the samples are segmented to train the network according to the batch model training mechanism, and after a certain number of iterations, the output weights are obtained. In the Reduce phase, Reduce collects all the outputs of each Map to obtain new weights, and decides whether to perform iterations according to the weight adjustment amount. Below are specific ideas and pseudocode.

1) Map function

In the Map function, the weight is read from HDFS to initialize the network. The samples are segmented and a certain number of network trainings is performed to achieve certain conditions. The output of the Map is a set of <Long, Writable> key-value pairs. Reduce takes the key-value pairs and processes them, defining an interface class - WeightWritable, which allows the Map function to operate through a writable interface. The class WeightWritable built according to the Hadoop serialization standard can save the change value of all weights. The training conclusions for each batch or round are sent to Reduce for synthesis, and the output value key of the Long Writable class is set to zero.

2) Reduce function

In the Reduce function, a set of <Long Writable, WeightWritable> key-value pairs output by the Map function is accepted and a set of <Long Writalbe, Int Writable> key-value pairs is output. Reduce counts the weight of each Map and finds the average of all weights as the new weight. In WeightWritable, an accumulation function and a division function are implemented to add and divide the weight matrix to obtain an arithmetic average weight. Before training, Reduce reads the weight from HDFS and compares two weight records. When the difference of weights is lower than the required standard, the output int value is set to false (represented by 0), otherwise it is set to True (represented by 1), and the weighted WeightWritable is adopted to update the weight in HDFS as the initial weight of the next iteration.

3) Combine function

The role of the Combine function is to merge the local output of the Map to greatly reduce the data input/output (I/O) time. It inputs a set of Map output < Int Writable,

Weight Writable > key-value pairs and outputs content being same with the output type of the Map.

4) Driving function

The content input to the Driving function is an array of strings containing two strings, a path for the training sample and an output path for Hadoop operation. This function first determines whether there is a weight file on HDFS, and if not, generates a WeightWritable type file that is given a smaller random value (-0.5~0.5) as the initial weight of the network. Then the Driving function creates a Hadoop job based on the training sample path and output path of parameters, and controls the job operation iteration according to the return value of the Reduce function.

V. EXPERIMENTAL RESULTS AND ANALYSIS

A.Algorithm convergence speed and accuracy test

The data set used in the experiment was derived from Breast Cancer Wisconsin (Original) Data Set. In the experiment, the training samples were divided into multiple parts and assigned to multiple Mappers for training and the Driving function is adopted to test network classification accuracy. The visualization results are shown in Figure 2 and Figure 3:

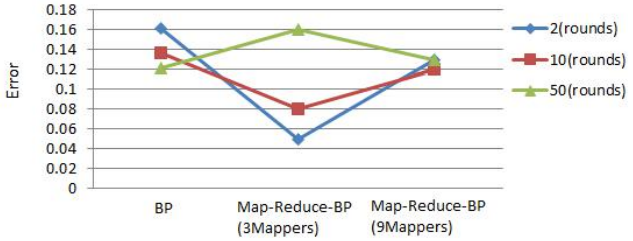

Figure 2. Error comparison between serial BP and Map-Reduce-BP

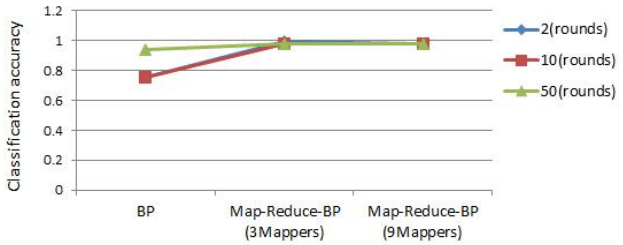

Figure 3. Classification accuracy comparison between serial BP and Map-Reduce-BP

B. Experimental results

The experiment shows the advantages of cloud computing clusters in big data processing and it shows that the bigger the data, the more obvious the advantages. It can be seen from the experimental results that the application of BP algorithm based on Map-Reduce can increase the training efficiency of neural network, which proves the feasibility of deep unsupervised learning neural network based on clouding computing. It is better than existing algorithms with better acceleration ratio, faster convergence and less iteration.

VI. CONCLUSIONS

This paper proposed a Deep Unsupervised Learning Optimization Algorithm Based on Cloud Computing. Compared with the conventional algorithms, BP algorithm based on Map-Reduce can increase the training efficiency of neural network,and it can proves the feasibility of deep unsupervised learning neural network based on clouding computing.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding of this study by Platform of Jilin Province Science and Technology (S&T) Department “Jilin Province S&T Innovation Center for Physical Simulation and Security of Water resources and Electric Power Engineering”([2018]271) and 2018 Jilin Province S&T Development Plan Technology Research Project" Development of Simulation Sandboard of Education System Based on Intelligent Mining on Agricultural big data ".

REFERENCES

[1] Yuan Bingqing;Neural Network and Deep Learning Basics[J]. Radio Wave Guard,2018(05):32-33.

[2] Dean J,Corrado G S,Monga R,et al. Large scale distributed deep networks[C]//Proc of the 25th International Conference on Neural Information Processing Systems.[S.I.]: Curran Associates Inc,2012:1223-1231.

[3] Yong Peng Yue, Research on deep unsupervised learning algorithm[D]. Southwest Petroleum University, 2015.

[4] Yan Hui,Hu Haiyan. Research and realization of ISIC-CDIO teaching experimental system based on RFID technology of web of things[J].Journal of Bionanoscience.2013.12(7):696-702.

[5] Jin Xu,Yafeng Yin,Hong Man,Feature Selection Based on Sparse Imputation,International Joint Conference on Neural Networks (IJCNN)2012.

[6] HaiJun Zhang , Parallel implementation and learning method of neural network based on Cloud Computing [D]. South China University of Technology, 2015.

[7] Yan Hui,long Duo.3D scanner-based corn seed modeling[J]. Applied Engineering in Agriculture，Vol.32(2). 2016.3:181-188.