ISSN:0975-9646

Secure Implementation of Artificial Neural Networks over Cloud

Pallavi Mhatre , Prachi Pimple , Surabhi Shikarkhane ,Poi Tamrakar

Computer Department, Pune University
ISB&M School Of Technology,Mulshi,Pune,Maharashtra,India,411042

***Abstract*: In today's scenario, there are systems proposed for organizing collaborative back propagation neural network learning over coupling of individual data sets of numerous parties. It is being done to enhance the precision of the learning result. Each of participating party wants to secure or hide their corresponding data sets from other parties. The current schemes which are supporting this kind of learning lack efficiency, due to use of either partitioning or considering only two parties at a time to do the learning. We instead are manipulating the data sets by means of scaling and transformation. We then are training these manipulated data sets using back propagation neural network on cloud forming patterns. Thus increasing the efficiency by minimizing the use of memory to store data sets, that is we are only saving the pattern and discarding the data, using cloud computing. We are also providing security to all data sets by means of scaling and transformation.**

***Key Words:* Collaborative, back propagation, artificial neural network, learning, cloud computing, manipulating**

1. INTRODUCTION

BACK-PROPAGATION using artificial neural networks (ANN) is a productive and potential way of conducting collaborative learning. The precision of back propagation neural networks is dependent on the volume and quality of data used for learning. Collaborative learning enhances the learning preciseness by integrating more data sets into the learning process, when compared to learning only with local data set. All the parties engaging can carry out learning on their own data sets as well as on other parties' data sets. Collaborative learning has become more approachable due to computing architectures like cloud computing.

One vital issue, in spite of potential advantages, is the security of data privacy for each participant. The collaborating parties, especially the ones with distinct trust discipline, may not wish to disclose their data to other participating party. To embrace the collaborative learning internet-wide, it is crucial to cater a solution for the parties lacking mutual trust to be allowed to carry out joint neural network learning without disclosing their private data sets. The solution shall also be efficient and scalable adequate to support random number of participants, having randomly arranged data sets.

The above issues can be solved by using secure multiparty computation (SMC). But SMC follows extremely complex computations and communication protocols. The reason of this complexity being the size of the circuitry used in SMC which makes it impossible to implement the two-party case. There are certain challenges that need to be satisfied, for providing a practical solution for secure implementation of ANN, simultaneously: Security should be provided to the data sets of respected participants and also to the intermediate results that are generated during the back-propagation neural network (BPNN) learning process, secure computation of various operations such as addition scalar product is required for the BPNN; the cost of computation and communication required for each participant should be minimal, for ensuring the practicality of the proposed system; the data sets, of different parties, used for collaborative training should not be partitioned in any way.

2. LITERATURE SURVEY

Many schemes related to privacy preserving BPNN learning have been proposed. Schlitter proposed a privacy preserving BPNN learning in which multiple parties perform collaborative BPNN learning with security being provided to their data. However, it works only on horizontally partitioned data. This scheme cannot provide security to intermediate results generated during learning. Chen and Zhong proposed a scheme of implementing secure BPNN, but it works only for two parties. It provides strong security to the data sets including intermediate results. Although, only vertical partitioning is supported by this scheme. As a solution to this problem, Bansal and others proposed a scheme for randomly partitioned data by enhancing it. However, this scheme was also proposed only for two parties.

Our proposed scheme tries to solve all these problems. It secures the data sets of each party that is engaged in carrying out BPNN. It reduces the Communication cost and computation cost. It increases the efficiency of computation. Our proposed scheme considers overall security of data and productivity. It also takes into consideration that the data is being unexposed to the cloud. It generates near most accurate results. It is cost effective, scalable and efficient. Hence, our scheme is more productive, secure, scalable and efficient to implement secure BPNN over cloud.

3. Algorithms

3.1 ANN:

An artificial neural network is inspired from our biological neural networks. It is a computational model and comprises of interconnected group of artificial neurons. It is a self-learning network in which neurons work by processing information through different levels of abstraction. The aim is to transform the inputs into significant outputs. It includes prediction of output values and can be used for pattern recognition.

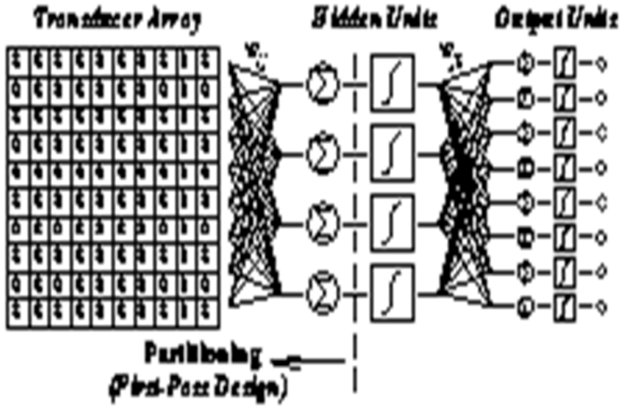


Fig.1.ANN

3.2 BPNN:

The back propagation algorithm uses feed forward ANN algorithm. Meaning the artificial neurons are organized in layers which send their signals in forward direction. When the errors are computed they are propagated backwards to get a specific output. The neurons provide inputs to the network in the input layer. The output is provided by neurons in an output layer. The number of hidden layers may vary from one to many. It uses supervised learning that is examples of inputs and outputs are provided to compute the error. This error is then reduced by back propagation.

The back propagation algorithm includes two phases: propagation and weight update.

Phase 1: Propagation:

Following steps are followed in each propagation:

1. Propagating the training pattern's input in forward direction through ANN for generating the output activations of propagations.

2. Propagating the propagation's output activations in backward direction through ANN by using the training pattern target for generating the deltas of all output and hidden neurons.

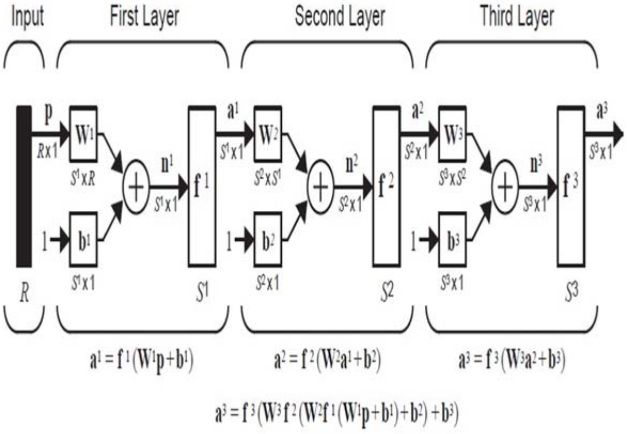


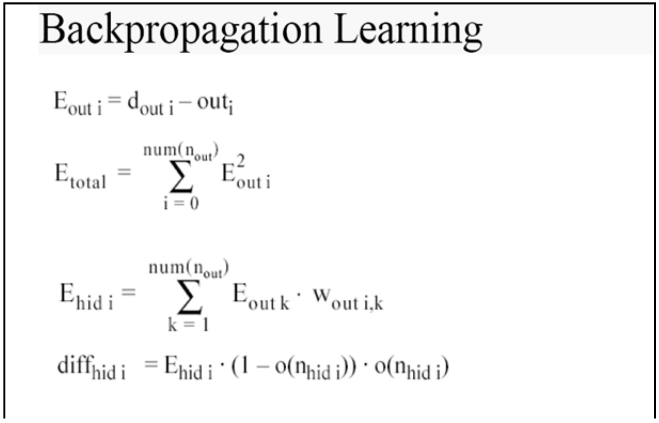
Fig.2. Equations of BPNN

Phase 2: Weight Update:

Following steps are followed for each weight-synapse:

1. Multiply output delta with input activation to obtain the gradient of the weight.

2. Subtract a ratio or percentage of the gradient from the weight. Repeat phase 1 and 2 until the performance of the network is satisfactory.



Backpropagation Learning

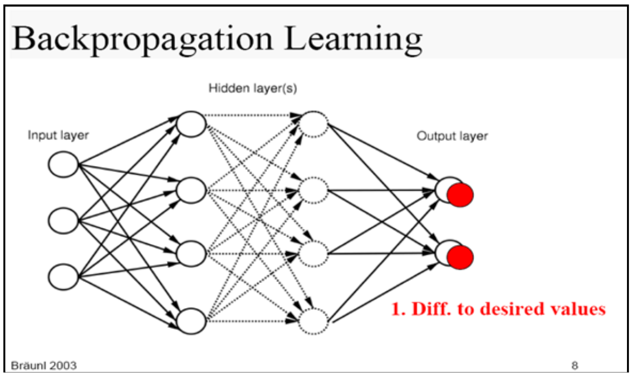
$$E_{out\ i} = d_{out\ i} - out_i$$

$$E_{total} = \sum_{i=0}^{num(n_{out})} E_{out\ i}^2$$

$$E_{hid\ i} = \sum_{k=1}^{num(n_{out})} E_{out\ k} \cdot w_{out\ i,k}$$

$$diff_{hid\ i} = E_{hid\ i} \cdot (1 - o(n_{hid\ i})) \cdot o(n_{hid\ i})$$

Fig.3. BPNN learning algorithm



Backpropagation Learning

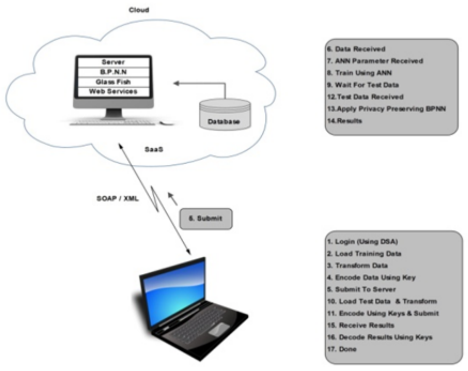
Fig.4. How BPNN works

4. PROPOSED SYSTEM

Problem Statement: The aim of our proposed scheme is to provide multiple parties to collaborate and perform BPNN learning securely. To secure the data sets of respected parties. To minimize the cost required for communication between parties and the computations carried out on the cloud. The scheme to be scalable and efficient is one of the purpose.

In our proposed system, most of the tasks are implemented by cloud. The admin after authentication, transforms, that is encodes the data set by performing transformation and scaling operations on it. The admin then transfers the data set on to the cloud. The clients authenticate themselves, generate the test data, transform the test data and encode it, and then send it to the cloud, after which the control is transferred to the cloud. The test data is encoded with same key as used by the admin. The clients then decode the results retrieved from the cloud. The keys used to encode the data are the same which are also used to decode the result retrieved from the cloud on the client side.

The cloud, on receiving the data sets from the admin, trains it by using back propagation algorithm of artificial neural networks and forms a pattern. It then discards the original data set and stores the pattern only. Thus making it unfathomable to modify the data and increasing the efficiency of security provided to the data. When it receives queries from clients it performs ANN on it and checks if any match to the queries is found in the stored data. It does so by performing back propagation on ANN that is the error received from ANN is back propagated to obtain specified approximate results. If approximated results are obtained they are propagated to the clients as response to their queries.



The parties involved in this learning decide on the keys to transform their respective data sets. The keys are not known to the cloud. The data stored on the cloud seems to be meaningful but only collaborating parties are aware of the alteration. And so the data is protected from the cloud as it is not aware of the actual meaning of the data. Hence the cloud has no clue of the transformations carried out on the data before transferring it to the cloud. For the cloud the data stored becomes useless that it will be useless to use this data for any malicious activity. Thus making our scheme further more effective for carrying out collaborative learning.

5. CONCLUSION

In this paper, we are providing security to the data stored on cloud as well as to the data sets belonging to different parties that are participating in collaborative BPNN learning. The data is secured by means of transformation and scaling. On cloud only the pattern is being stored, on which ANN is carried out, the original data sets are discarded as soon as the patterns are created.

The data sets are manipulated in such a way that only participating parties are aware about it. That is the unauthorised parties are not aware of the data manipulation. The data sets seem to be meaningful and original to the unauthorised parties, but actually it is meaningless and useless.

REFERENCES

[1] Jiawei Yuan and Shucheng Yu, *"Privacy Preserving Back-Propagation Neural Network Learning Made Practical with Cloud Computing"*, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 25, NO. 1, JANUARY 2014.

[2] Tingting Chen and Sheng Zhong,*" Privacy Preserving Back-Propagation Neural Network Learning"* The State University of New York at Buffalo, Buffalo, NY 14260, U.S.A.,2009.

[3] A. Bansal, T. Chen, and S. Zhong, *"Privacy Preserving Back-Propagation Neural Network Learning over Arbitrarily Partitioned Data Neural Computing Applications"*, vol. 20, no.1, pp. 143-150, Feb. 2011

[4] HIPPA, National Standards to Protect the Privacy of Personal Health Information, http://www.hhs.gov/ocr/hipaa/finalreg.html.

[5] R. Grossman and Y. Gu, "Data Mining Using High Performance Data Clouds: Experimental Studies Using Sector and Sphere,"

[6] Proc.14th ACM Int'l Conf. Knowledge Discovery and Data Mining (SIGKDD '08), pp. 920-927, 2008

[7] Vaidya, J. & C. Clifton. (2002). Privacy Preserving Association Rule Mining in Vertically Partitioned Data, in *Proc. of SIGKDD'02*, 639-644.

[8] Yang, Z., Zhong, S., & Wright. R. (2005). Privacy-preserving classification of customer data without loss of accuracy. In *Proc. 5th SIAM.*

[9] Nico Schlitter, "A Protocol for Privacy Preserving Neural Network Learning on Horizontally Partitioned Data" *International Conference on Data Mining (SDM),* Otto-von-Guericke-University Magdeburg, Germany, 2008

[10] Vaidya, J. & C. Clifton. (2002). Privacy preserving association rule Mining in vertically partitioned data, in The *Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 639-644.

[11] R. Grossman and Y. Gu, "Data Mining Using High Performance Data Clouds: Experimental Studies Using Sector and Sphere," Proc. 14th ACM Int'l Conf. Knowledge Discovery and Data Mining (SIGKDD '08), pp. 920-927, 2008.

[12] R.L. Grossman, "The Case for Cloud Computing," IT Professional, vol. 11, no. 2, pp. 23-27, Mar. 2009.

[13] R. Law, "Back-Propagation Learning in Improving the Accuracy of Neural Network-Based Tourism Demand Forecasting," Tourism Management, vol. 21, no. 4, pp. 331-340, 2000.