ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

Privacy Preserving Neural Network Learning Using Cloud Computing

Rashmi Nade [1], Prof.P.D.Lambhate [2]

P.G. Student, Department of Computer Engineering, JSCOE, Handewadi Road, Pune, India[1]

Associate Professor, Department of Computer Engineering, JSCOE, Handewadi Road, Pune, India[2]

ABSTRACT: In cloud computing, to increase the correctness of learning result, multiple parties might collaborate through conducting neural network learning with union of their individual information sets, and during this learning method not single party needs to disclose her/his non-public information to others. The existing methods that support this kind of approach for collaborative learning, but these are limited only for two parties as well as data partition. There is no solution for two or more parties, for collaboratively conduct the learning, with an arbitrarily partitioned data set. Here we proposed a solution to overcome the above stated limitations of existing collaborative learning methods used in cloud computing. This open problem is solved by using the power of cloud computing. In this proposed scheme, each party encrypts his/her private data locally and uploads the cipher texts into the cloud. The cloud then executes most of the operations pertaining to the learning algorithms over cipher texts without knowing the original private data. Back Propogation algorithm is used for training the neural network..

KEYWORDS: Neural Network, Privacy Preserving, Cloud computing, Back Propagation , AES algorithm.

I. INTRODUCTION

Neural Network is formed with a set of inputs and outputs units which are connected, and each connection has a weight associated with it. Neural network needs a long training period. In general a neural network is trained initially or given a large amount of data and rules about data relationship. Neural networks are well suited for continuous valued input and output, successful on a wide array of real world data. The algorithms used in neural network are naturally parallel which help to speed up the computation process.

Neural Network is the collection of nodes and edges. The edges are associated with a weight and there is an activation function associated with the network that takes weights as inputs and generates the output.

Back-propagation is one of the methods for learning the neural networks and has been widely used in various applications. Back propagation method works backward, it calculates the error between output (expected values) and calculated values and trains neural network until the results are near to expected values. The learning accuracy is mainly affected by data used for learning. Instead of learning with limited dataset collaborative learning improve the learning result. In collaborative learning more than two parties involved in learning process and they can use data of other parties, so the privacy preservation of data is very important here. For collaborative learning cloud is the best infrastructure.

Cloud is an infrastructure which provides resources and services over the internet. Cloud computing platform gives a large computing capacity, where the capacity is resizable in the cloud. By using the cloud which saves the time and application can develop and deploy faster.

Cloud Computing is a technology which uses the internet and central remote servers to maintain data and applications. Cloud computing allows users to use applications without installation and access their personal files at any computer with internet access. This technology allows for much more efficient computing by centralizing data storage, processing and bandwidth. A simple example of cloud computing is Yahoo email, Gmail.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

II. LITERATURE SURVEY

Jiawei Yuan; Shucheng Yu.[1] Presents a back propagation neural network learning algorithm is used for learning process over arbitrarily partitioned data. As we discussed above back propagation method is very efficient for training neural networks and here data is arbitrarily partitioned means here is no limitations for partitioning data like data is partitioned only horizontally or only vertically, it can be portioned either horizontally or vertically. Here a solution is proposed for multiple parties, so the protection of each user's private data is very important task, for this purpose they have implemented a privacy preserving multiparty BPN network learning algorithm. This method protects the private data of participants and also protects the in between results produced during the process of learning. To protect the learning result a secure scalar product algorithm is used.

N. Schlitter.[3] Introduces a method for privacy preserving BPN network learning, and allows two or more than two parties to jointly conduct Back Propagation Neural network learning. During this process private data of users does not revealed. As this method allows more than two parties, it needs a secure sum and secure matrix addition so for this purpose a neural network classification algorithm is used; this algorithm is extension of Rumelheart's classification algorithm.

Here are some limitations, that the solution is given only for horizontal partitioned data, in horizontal partition of data every user share the same data schema, and each party holds some records of total data sets. This method uses extension of Rumelheart algorithm which does not guarantee privacy.

This method cannot protect, in between results, generated during the BPN learning process. To protect the private data of participants, a secure computation of in between results is very important, because in between results also contains the sensitive data.

T. Chen and S. Zhong. [2] This paper presents a privacy preserving BPN network learning algorithm, but this algorithm is applicable only for two parties. This algorithm provides a strong security for data sets and during the learning process it protects the in between results. This method is having limitations like it just supports two parties. This method only supports vertically partitioned data. This method is not scalable.

A. Bansal, T. Chen, and S. Zhong. [7] Presents a improved method and gives a solution for data which is arbitrarily partitioned. In arbitrary partitioning of data, there is no order of how the data is divided. This method is just like the method used in [2], which was for vertically partition of data, and here it is for arbitrary partition of data. Here they tried to extent the method from two parties to multi party scenario but, extension of two parties to multi party introduce computation/communication complexity which is quadratic in the number of participants. In real world implementation, such a complexity represents a huge cost on each party. That's why this paper just considers the two-party scenario though it supports arbitrarily partitioned dataset.

III. PROPOSED WORK

Problem Statement: In proposed work we enable more than one party to perform neural network learning without revealing their own private data set. During this process data is uploaded to cloud. Data collected from user is arbitrarily partitioned. And the communication cost is kept less.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

System Architecture:

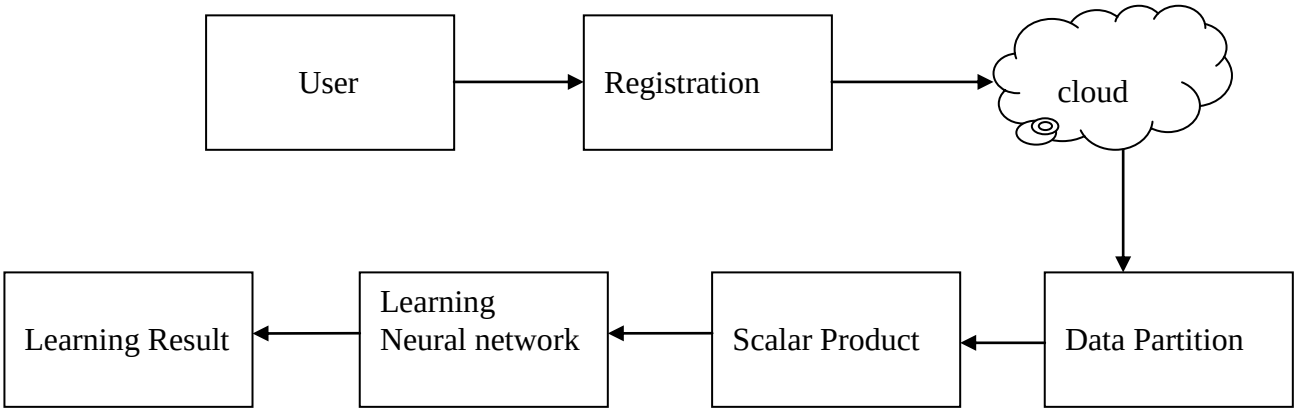

Fig. 1 System Architecture

In this scheme at first user register on cloud, Then user uploads file to cloud, When user uploads file to cloud file is encrypted with AES algorithm, which means a cipher text is uploaded to cloud. As cipher text is uploaded to cloud all learning process is done on cipher text only. So that no user can learn any other users data, and thus privacy of multiple users is preserved.

After that data from all users is collected for neural network learning process, data is arbitrarily partitioned in to all users. In arbitrarily partition there are no limitations like horizontal partition or vertical partition, data is randomly partitioned as follows:

$$\text{Data Partition} = \frac{\text{Total Collected Data}}{\text{No. of users}}$$

Partitioned data is distributed among all users, no user is having same data set and then scalar product is calculated. After that a Neural network learning process starts, for learning neural network a Back-Propagation algorithm is used. This algorithm works in two stages Feed Forward and Back Propagation.

Feedforward Stage

1. Initialize weights with small, random values
2. While stopping condition is not true – for each training pair (input/output):
   - •each input unit broadcasts its value to all hidden units
   - •each hidden unit sums its input signals & applies activation function to compute its output signal
   - •each hidden unit sends its signal to the output units
   - •each output unit sums its input signals & applies its activation function to compute its output signal.

Backpropagation stage

3. Each output computes its error term, its own weight correction term and its bias(threshold) correction term & sends it to layer below
4. Each hidden unit sums its delta inputs from above & multiplies by the derivative of its activation function; it also computes its own weight correction term and its bias correction term signal.

Adjusting the Weights

5. Each output unit updates its weights and bias

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

6. Each hidden unit updates its weights and bias – Each training cycle is called an epoch. The weights are updated in each cycle – It is not analytically possible to determine where the global minimum is. Eventually the algorithm stops in a low point, which may just be a local minimum.

IV. RESULTS AND ANALYSIS

Following figures shows the results obtained after the implementation of work. In Fig(2)we see the encryption time required for different data sets we use three data sets iris, diabetes and kr_vs_kp. In our proposed system we used AES algorithm for encryption of data. In fig(3) we see the time required for learning neural network, if the data is small then learning time is also small. When encryption and learning time is less automatically communication cost is less. In fig(4) we see the accuracy comparison with existing systems.

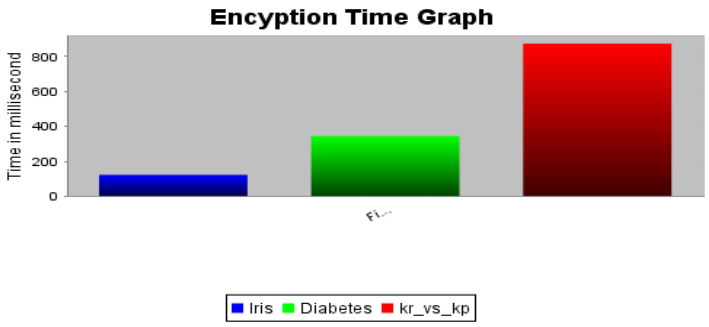

Fig 2: Encryption time for different data set

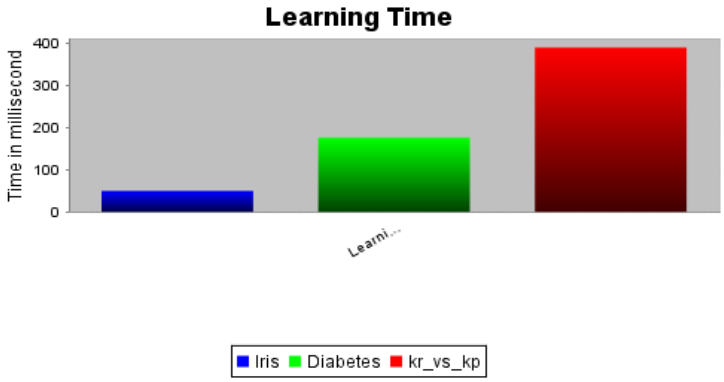

Fig 3: Learning time for different data set

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

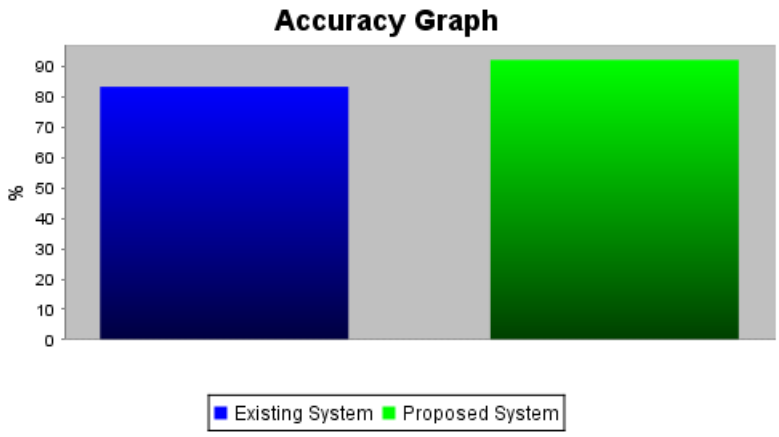

Fig 4: Accuracy Graph

V. CONCLUSION

In this paper, we proposed a secure and practical Collaborative learning method, by using BPN network learning algorithm over arbitrarily partitioned data. And a learning time for neural network is less. The communication and computation cost is less and independent to number of user. This scheme is secure, efficient and scalable.

REFERENCES

[1] Jiawei Yuan; Shucheng Yu. "Privacy Preserving Back-Propagation NeuralNetwork Learning Made Practical with Cloud Computing," *IEEE Transaction paper on parallel and distributedsystem, digital object identifier* 10.1109/TPDS.2013.18,1045-9219/13/$31.00,2013 IEEE.

[2] T. Chen and S. Zhong. "Privacy-preserving backpropagation neural network learning". *Trans. Neur. Netw.*, 20(10):1554–1564, Oct. 2009

[3] N. Schlitter. "A protocol for privacy preserving neural network learning on horizontal partitioned data". In *Proceedings of the Privacy Statistics in Databases (PSD)*, Sep. 2008.

[4] D. Boneh, E.-J. Goh, and K. Nissim. "Evaluating 2-dnf formulas on ciphertexts". In *Proceedings of the Second international conference on Theory of Cryptography*, TCC'05, pages 325–341, Berlin, Heidelberg,2005.

[5] A. Frank and A. Asuncion. "UCI machine learning repository", 2010.

[6] M. A. Inc. *Amazon Elastic Compute Cloud (Amazon EC2)*. Amazon Inc., http://aws.amazon.com/ec2/#pricing, 2008.

[7] A. Bansal, T. Chen, and S. Zhong." Privacy preserving backpropagation neural network learning over arbitrarily partitioned data". *Neural Comput. Appl.*, 20(1):143–150, Feb. 2011.

[8] karmila.staff.gunadarma.ac.id/.../files/.../TayanganBackpropa gation.pdf .