PRIVACY PRESERVING BACK-PROPAGATION NEURAL NETWORK IN CLOUD COMPUTING

Dr.P.Kumar M.sc., M.Tech., Ph.d, R.Muthu Vijay Deepak M.Tech.,

Abstract— **we find the problem of Privacy Preserving Back Propagation Algorithm for a Vertically Partitioned Dataset. To improve the learning, Enhanced data is more important to find the exact privacy concern of each data holder by extending the privacy preservation suggested to original learning algorithms. In this paper, we try to improve preserving the privacy in an important multilayer neural networks and learning model. We present a privacy preserving multiparty distributed algorithm of back propagation which allows a neural network to be trained without requiring either party to reveal her data to the others. We gave more correctness and security analysis of our algorithms. The effectiveness of our algorithms is checked and verified by experiments on various real world data sets. We address this open problem by incorporating the computing power of the cloud computing. The main idea of our paper can be summarized as follows: each participant first encrypts her/his private data with the system public key and then uploads the cipher texts to the cloud; cloud servers then execute most of the operations pertaining to the learning process over the cipher texts and return the encrypted results to the participants.**

***Index Terms*— Privacy reserving, Learning, Neural Network, Back-Propagation, Cloud computing, Computation Outsource**

I. INTRODUCTION

Back-propagation is an effective method for learning neural networks and has been widely used in various applications. The perfect accuracy of the learning result is despite other facts, that the learning is highly affected the solution to provide the volume of high quality data used for learning. As compared the volume of quality data to learning with only local data set, combined learning algorithm that improves the learning accuracy by incorporating more data sets into the learning process the participating data set address the problem not only on their own data sets, but also on others data sets. Now a day's remarkable growth of new computing infrastructures such as Cloud Computing, it has been more convenient and use over to ever users across the Internet, who may not even know each other, to conduct collaborative learning through the shared infrastructure. Despite the future benefits, one crucial issue connected with the Internet-wide collaborative neural network learning is the protection of data privacy for each participant. In particular, the participants may use the different trust domains may not want to disclose their private data sets, which may contain privacy or proprietary information to anybody else. In computer applications such as healthcare, disclosure of sensitive data, e.g., protected health information, is not only a privacy issue but of legal concerns according to the privacy rules such as Health Insurance Probability and Accountability Act(HIPAA). In order to embrace the Internet wide collaborative learning, it is imperative to provide a solution that allows the participants, who lack mutual trust, to conduct neural network learning jointly without disclosing their respective private data sets. Preferably, the solution shall be efficient and scalable enough to support an arbitrary number of participants, each possessing arbitrarily partitioned data sets.

A. Back-Propagation Neural Network Learning

Back-Propagation neural network learning algorithm is mainly composed of two stages:

1. Feed forward
2. Error back - propagation.

In the Feed Forward Stage, values at each layer are calculated using the weights, the sigmoid function, and the values at the previous layer. In the Back - Propagation stage, the algorithm checks whether the error between output values and target values is within the threshold.

Back-propagation is an effective method for learning neural networks and has been widely used in various applications. The accuracy of the learning result, despite other facts, is highly affected by the volume of high quality data used for learning. As compared to learning with only local data set, collaborative learning improves the learning accuracy by incorporating more data sets into the learning process the participating parties carry out learning not only on their own data sets, but also on others data sets. With the recent remarkable growth of new computing infrastructures such as Cloud Computing, it has been more convenient than ever for users across the Internet, some of us don't know each other they to conduct joint/collaborative learning through the shared infrastructure of cloud computing.

In order to use the Internet wide collaborative learning, it is imperative to provide a solution that allows the participants to correct the values, who lack mutual trust, to conduct neural network learning jointly without disclosing their respective private data sets. Preferably, the solution shall be efficient and scalable enough to support an arbitrary number of participants, each possessing arbitrarily partitioned data sets.

***Manuscript received June, 2014*.**

***Dr.P.Kumar M.Sc., M.Tech.,** Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli , India, Mobile No:+91-9943365978*

***Mr.R.Muthu Vijay Deepak B.Tech..,** Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli , India, Mobile No:+91-9790433106*

B. Neural Network

Neural networks have seen an explosion of interest over the last few years and are being successfully applied across an extraordinary range of problem domains, in areas as diverse as finance, medicine, engineering, geology and physics.

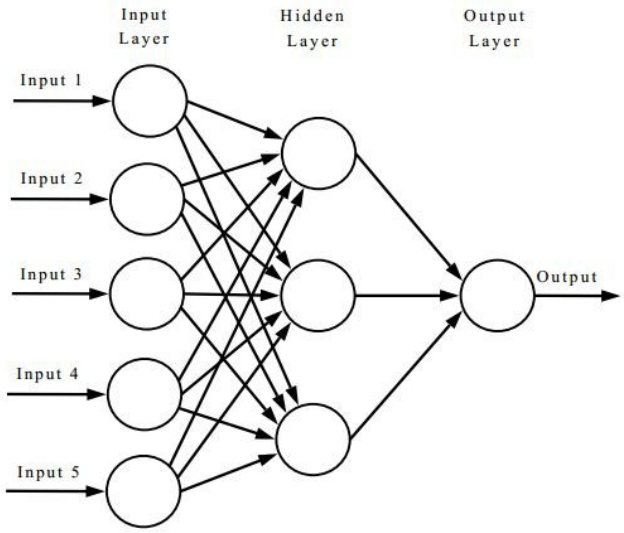

Fig 1.1: Neural Network

If problem structure is well analyzed, traditional computers could still outperform Neural Network but in cases where problem has not been analyzed in details, Neural Network could be used to learn from large set of examples. Neural Network can handle errors better than traditional computers programs.

II. SURVEY ON PRIVACY PRESERVING BACK PROPOGATION

A. Privacy Preserving Neural Network Learning On Horizontally Partitioned Data

Barni, M., Orlandi, C., & Piva, A. (2006)[6] Presents a new approach for privacy preserving neural network training. Several studies have been devoted to privacy preserving supervised model learning, but little work has been done to extend neural network learning with a privacy preserving protocol. Neural networks are popular for many applications, among else those calling for a robust learning algorithm. In this study, we elaborate on privacy preserving classification as well as regression with neural networks on horizontally partitioned data. We consider a scenario of more than two parties that are semi-honest but curious. We extend the neural network classification algorithm with protocols for secure sum and secure matrix addition. The extended algorithm does not fully guarantee privacy in the sense, but we show that the information revealed is not associated to a specific party and could also be derived by juxtaposing the local data and the final model.

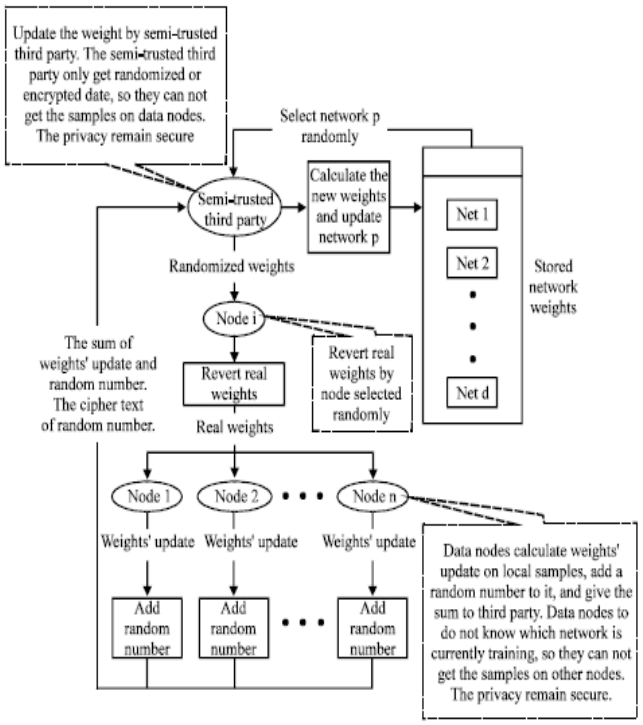

Fig 2.1: Neural Network on Horizontally portioned Data

Neural Networks are nature inspired computation models that are widely used for regression and classification tasks. A neural network consists of nodes and weighted edges. In general, we distinguish feed forward and recurrent neural networks. In a feed forward network, the information is transmitted only in one direction, from the input nodes, through the hidden nodes (if any) to the output nodes. A feed forward network has therefore no cycles or loops. In recurrent neural networks, connections between nodes form a directed cycle.
Therefore, Information might also be transmitted backwards. An example of a feed forward network is shown in Fig. 1. Each node, also called neuron, produces an output by applying the internal activation function A on the weighted neuron inputs. The output is weighted by the edge weight and the updated value is forwarded to the next layer. By this, an input vector traverses from the input layer through the network and produces an output in the output layer. In other words, the function performed by the network is determined by the nodes' activation function and the internal network weights.

B. Privacy Preserving Back-Propagation Neural Network Learning Over Arbitrarily Partitioned Data

Neural Networks have been an active research area for decades. However, Ankur Bansal Tingting Chen Sheng Zhong (2010) [4] are privacy bothers many when the training dataset for the neural networks is distributed between two parties, which is quite common nowadays. Existing cryptographic approaches such as secure scalar product protocol provide a secure way for neural network learning when the training dataset is vertically partitioned. In this paper we present a privacy preserving algorithm for the neural network learning when the dataset is arbitrarily partitioned between the two parties. We show that our algorithm is very secure and leaks no knowledge (except the final weights learned by both parties) about other party's

data. We demonstrate the efficiency of our algorithm by experiments on real world data. Privacy preserving neural network learning With the invention of new technologies, whether it is data mining, in databases or in any networks, resolving privacy problems has become very important. Because all sorts of data is collected from many sources, the field of machine learning is equally growing and so are the concerns regarding the privacy. Data providers for machine learning are not willing to train the neural network with their data at the expense of privacy and even if they do participate in the training they might either remove some information from their data or can provide false information.

C. *Challenges in Smc*

Theoretically, secure multi-party computation (SMC) can be used to solve problems of this kind. But the extremely high computation and communication complexity of SMC, due to the circuit size, usually makes it far from practical even in the two-party case. In order to provide practical solutions for privacy preserving back-propagation neural (BPN) network learning, three main challenges need to be met simultaneously:

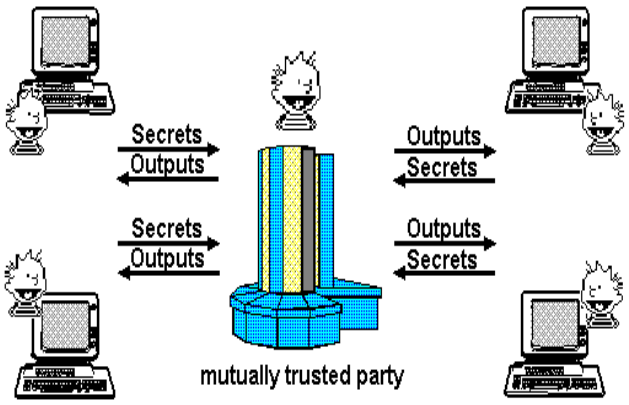

Fig 2.2: Secure Multi-Party Computation

1) To protect each participant's private dataset and intermediate results generated during the BPN network learning process, it requires secure computation of various
Operations, e.g. addition, scalar product and the nonlinear sigmoid function, which are needed by the BPN network algorithm
2) To ensure the practicality of the proposed solution, the computation/communication
cost introduced to each participant shall be affordable. In order to accommodate a large range of collaborative learning, the proposed solution shall consider system scalability. In particular, it shall be able to support an arbitrary number of participants without introducing tremendous computation/communication costs to each participant.
3) For collaborative training, the training data sets may be owned by different parties and
Partitioned in arbitrary ways rather than a single way of partition.

III. PROBLEM DEFINITION

A. *Direct Rule Protection (Method1):*

In this algorithm α-discriminatory rules are used and finding the subset of the database. Among the records of the database one should change those with lowest impact on other rules. Then the records with minimum impact are selected for change with aim of scoring well in terms of utility measures proposed. We call this procedure is called impact minimization.

B. *Direct Rule Protection (Method2):*

In this algorithm is to be used to find the subset of the database and perform impact minimization same as in algorithm 1 but α-discriminatory rules and data transformation are different.

C. *Direct Rule Protection and Rule Generalization (Method3):*

Algorithm 3 should be run to combine rule generalization and one of the two direct rule protection methods. α- discrimination rule –show rile generalization should be performed after determining the records that should be changed for impact minimization.

D. *Indirect Discrimination Prevention Algorithm:*

Direct rule protection for IRP is to be provided from which as algorithm implementers method for IRP can be early derived.

E. *Direct and Indirect Prevention Algorithm:*

The algorithm starts with redlining rules. From each redlining rules more than one indirect α-discriminatory rule might be generated. Transformation is performed until both the direct and indirect rule protection requirements are satisfied.

F. *Problem to handle*

Discrimination can be viewed as the act of unfairly treating people on the basis of their belonging to a specific group. Data in decision records are typically highly dimensional: as a consequence, a huge number of possible contexts may, or may not, be the theater for discrimination. Consider the case of gender discrimination in credit approval: although an analyst may observe that no discrimination occurs in general, it may turn out that older women obtain car loans only rarely. Many small or large niches that conceal discrimination may exist, and therefore all possible specific situations should be considered as candidates, consisting of all possible combinations of variables and variable values: personal data, demographics, social, economic and cultural indicators, etc.

The complexity is indirect discrimination: the feature that may be the object of discrimination, e.g., the race or ethnicity, is not directly recorded in the data. Nevertheless, racial discrimination may be hidden in the data, for instance in the case where a redlining practices is adopted: people living in a certain neighborhood are frequently denied credit, but from demographic data we can learn that most people living in that neighborhood belong to the same ethnic minority.

IV. OUR PROPOSED SCHEME

A. Extraction of Knowledge:

It allows for automatic and routine collection of large amounts of data. Those data are often used to train association/classification rules in view of making automated decisions, like attribute identification and selection, outlier removal, data normalization and numerical discretization, visual data analysis, hidden relationships discovery, and a diabetes prediction model construction.etc. The background knowledge might be accessible from publicly available data or might be obtained from the original data set itself because of the existence of nondiscriminatory attributes that are highly correlated with the sensitive ones in the original data set. Knowledge Extraction is the creation of knowledge from structured and unstructured databases. The resulting knowledge needs to be in a machine-readable and machine-interpretable format and must represent knowledge in a manner that facilitates inference. It requires either the reuse of existing formal knowledge or the generation of a schema based on the source data.

B. Classification Rules Protection:

Classification rule is a procedure in which the elements of the patients set are each assigned to one of the classes. Direct and Indirect discriminatory rules are converted to legitimate classification rules. It is used to remove the unwanted values from the dataset. Classification rules are trained on given data for the prediction of class labels of unknown data samples. The extracted information can then be used for the classification of the content of large textual bases. In this module, we present two examples of information that can be automatically extracted from text collections: probabilistic associations of key-words and prototypical document instances.

C. Measure Discrimination:

Discrimination discovery methods consider each rule individually for measuring discrimination without considering other rules or the relation between them. Discrimination prevention, the other major antidiscrimination aim in data mining, consists of inducing patterns that do not lead to discriminatory decisions even if the original training data sets are biased. Three approaches are conceivable

Preprocessing
In processing
Post Processing

In measure discrimination we use some measures to identify the direct discrimination efficiently. Two types of measures are there: Direct discrimination measures and indirect discrimination measures. In direct discrimination measure one of the measure is used as extended lift. In indirect discrimination measure some theorems are used. It is used to identifying the redlining rules. A PD rule could probably lead to discriminatory decisions. Therefore, some measures are needed to quantify the direct discrimination potential. A PND rule could lead to discriminatory decisions in combination with some background knowledge to quantify the indirect discrimination potential.

D. Direct Rule Protection and Generalization:

Discrimination prevention based on preprocessing. They attempt to detect discrimination in the original data only for one discriminatory item and based on a single measure. This approach cannot guarantee that the transformed data set is really discrimination free, because it is known that discriminatory behaviors can often be hidden behind several discriminatory items, and even behind combinations of them. They only consider direct discrimination. No measure to evaluate how much discrimination has been removed and how much information loss has been incurred. Overcome this problem we use data transformation method as rule Protection and rule Generalization .It is based on both direct and indirect discrimination.

1) Rule Protection:

There are two methods that could be applied for direct rule protection. One method changes the discriminatory item set in some records (e.g., gender changed from male to female in the records with granted credits) and the other method changes the class item in some records (e.g., from grant credit to deny credit in the records with male gender). Similar data transformation methods could be applied to obtain direct rule protection with respect to other measures.

2) Rule Generalization:

Rule generalization is another data transformation method for direct discrimination prevention. It is based on the fact that if each discriminatory rule and the database of decision rule was an instance of at least one non redlining, the data set would be free of direct discrimination. In rule generalization, we consider the relation between rules instead of discrimination measures. Data transformation with minimum information loss should be applied in such a way that each discriminatory rule either becomes protective or an instance of a non redlining PND rule. We call the first procedure direct rule protection and the second one rule generalization.

E. Discrimination Prevention:

Direct and indirect discrimination prevention can be described in terms of two phases:

Discrimination Measurement
Data Transformation

Direct and indirect discrimination discovery includes identifying α-discriminatory rules and redlining rules. To this end, first, based on predetermined discriminatory items in DB, frequent classification rules in FR are divided in two groups: PD and PND rules .Second, direct discrimination is measured by identifying discriminatory rules among the PD rules using a direct discrimination measure and a discriminatory threshold. Third, indirect discrimination is measured by identifying redlining rules among the PND rules combined with background knowledge, using an indirect discriminatory measure, and a discriminatory threshold. Let MR be the database of direct discriminatory rules obtained with the above process. In addition, let RR be the database of redlining rules and their respective indirect discriminatory rules obtained with the above process.

Data transformation method, such as Direct and Indirect Discrimination Prevention Algorithms are used for simultaneous direct and indirect discrimination prevention. The algorithm starts with redlining rules. Data transformation is performed until both the direct and the indirect rule protection requirements are satisfied. Here direct rule protection algorithm and direct and indirect discrimination prevention algorithm is used. The investigational outcome reported demonstrates that the proposed techniques are quite successful in both goals of removing discrimination and preserving data quality.

F. Security Model

In this security model we assume the existence of a trusted authority that was all trusted by the parties. System secret key will not participate the model of key generation and issuing. Whenever necessary we learned the each party's private data. When investigation is needed in case malicious party that may be in found in the private data, intentionally interrupts the system in data sets.

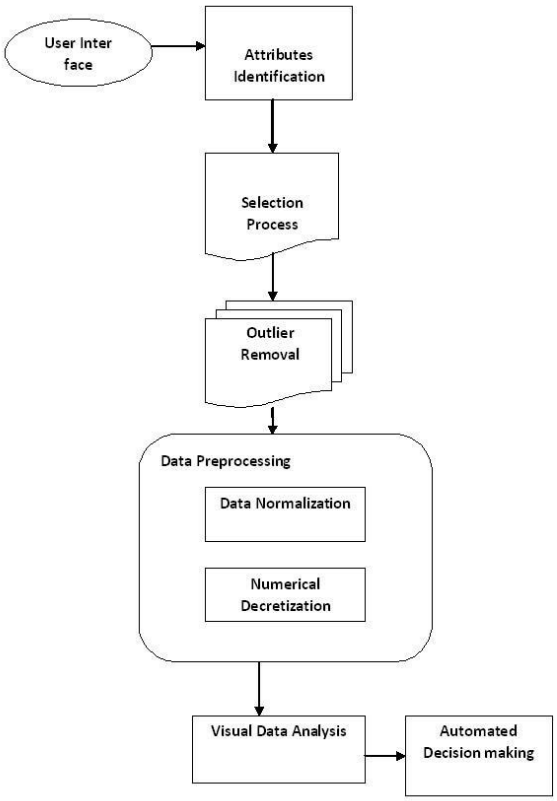

Fig 4.1: Security Model of BPN

User interface of the algorithm is easily to interact to that of process of system infrastructure. First process of the algorithm is checked the identification of user, it will collect the attributes of user to identify who is the user of this interface. Second process is selection process it will select the attributes to cipher text to encrypt the process making for user. The removal of outlier is simply to use the cipher text to encrypt process it will be remove the non-processing data. Third process is Data preprocessing is this measures of discrimination which has been data transmission with minimum loss of information that be applied in each way should be such a way that each discriminatory rule either becomes protective or an instance of a non redlining PND rule. In a Numerical decretization in terms of computation cost and communication cost and compare it with the existing techniques.

We see this fig 4.2 it predicate the logic of security model that was assuming in the level of security prospection of detailed shown from the diagram.

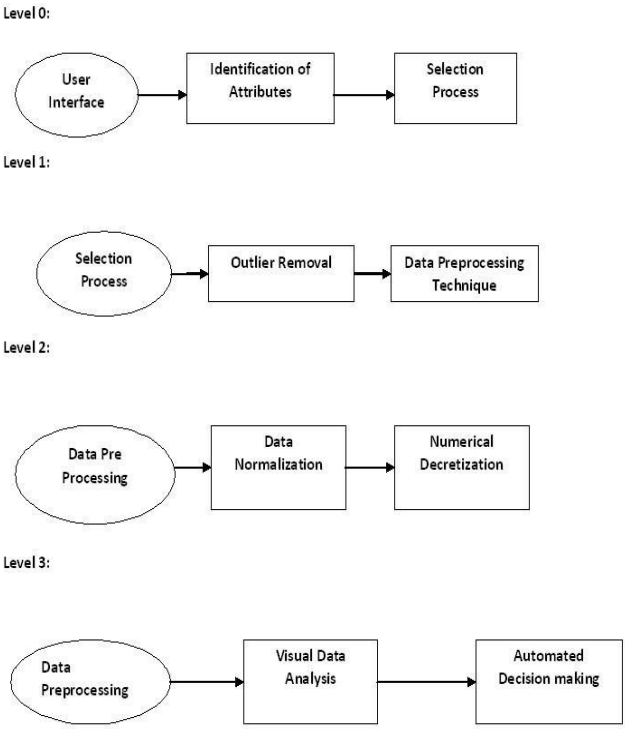

Fig 4.2: Data flow diagram of security model

System key participate the public key pre-processing that generate and issued to the user. For multi-party privacy preserving BPN network learning with fixed neural network architecture, there are two factors - the number of party and the size of dataset - that mainly affect the system performance according to the experimental results of existing schemes. When investigation is needed in case malicious party that may be in found in the private data, intentionally interrupts the system in data sets. The participating parties do not fully trust each other. Therefore, they do not want to disclose their respective private data (except for the final weights learned by the network) to any other parties than TA. The cloud is not fully trusted by the participating parties either, i.e., the cloud is not allowed to learn about the sensitive information, such as original data sets and intermediate data. In this paper, we follow the curious-but-honest model. The trusted parties that encrypted the file easily and much more elaborate the data.

In real life, parties such as the government agents or organization alliances can be the TA. Although the existence of TA is helpful, we leave the completely distributed solution as a future work.

G. Secure scalar product with cloud

Propose an algorithm that allows multiple parties to perform secure scalar product and homomorphic addition operations on cipher texts using cloud computing. Specifically, each party encrypts her/his data with the system public key and uploads the cipher texts to the cloud. The

cloud servers compute the sum of original messages based on their cipher texts. If the original messages are vectors, the cloud computes the scalar product of the vectors. During this process, the cloud does not need to decrypt nor learn about the original messages. The final result of the sum or scalar product is returned to all the parties in cipher text. Decrypting the results needs the participation of all the parties using Pollard's lambda method

Message decryption in the BGN algorithm involves solving the discrete log using Pollard's lambda method. On a single contemporary computer, for example, the Pollard's lambda method is able to decrypt numbers of up to 30-40 bits within a reasonable time slot (e.g., in minutes or hours). Decryption of larger numbers is usually believed less practical in terms of the time complexity. In practice, however, it is hard to guarantee that the final results (numbers) are always small enough for the Pollard's lambda method to efficiently decrypt.

H. Privacy preserving for cloud environment

In this process data owner generates the key for the user all are using the PRE-Key generating function that distributes all the users. They generates the re-encryption key for each user using the PRE-ReKeyGen function; this key was enable the cloud to convert the original key to encrypt the whole data under the data owner public key to another encrypted format under the user public key without knowing the data.

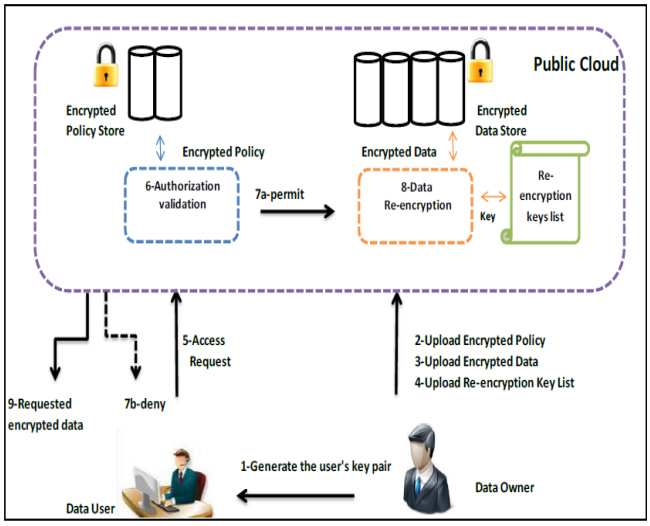

Fig 4.3: Privacy preserving of cloud computing

User access is we request access to specific dada to client for public key to decrypt the whole encrypt data client using Pol-Enc function. When the cloud receives the encrypted request; it searches the policy of authentication to validate the original data. If the user has permissions to access the whole data and performs the next process is decrypt the requested data using the authentication public key belongs to the user in PRE-ReEnc function. In case the user was not authorized the cloud service that denies the access and informs the user.

I. Accuracy analysis

We analyze the accuracy loss in our privacy-preserving BPN network learning scheme and compare it with existing schemes. Recall that in our proposed scheme the only place that introduces accuracy loss is the approximation of the activation function. We utilize the Maclaurin series expansion to approximate the function, whose accuracy can be adjusted by modifying the number of series terms according to the system requirement. Similar method of approximation with Maclaurin series expansion is also used in though it just supports the two-party setting.

we can reduce the accuracy loss by about 1% which outperforms the existing schemes. This is because introduce accuracy losses not only in the approximation of the activation function, but also during the mapping of real numbers in sigmod function to fixed-point representations in every step of Feed Forward Stage and Back-Propagation Stage. Differently, our proposed scheme omits this limitation and thus can be efficiency performed on the sigmod function without any accuracy loss during the secure computation process. Compared to the non-privacy-preserving BPN network learning algorithm, our scheme introduces about only 1.3%-2% more error rate in 9 series terms setting. Which are acceptable in practical use and can be further improved by adding more series terms.

V. CONCLUSION

In this paper, we build on our previous work by expanding our data to include multiple product categories and multiple textual features such as different readability metrics, information about the reviewer history in the use of Back Propagation Neural Network, different features of reviewer disclosure and so on. The present paper is unique in looking at how lexical, grammatical, semantic, and stylistic levels in the text of reviews affect product sales and the perceived helpfulness of these reviews.

VI. FUTURE ENHANCEMENT:

Future work can look at real demand data. Our sample is also restricted in that our analysis focuses on the sales at one e-commerce retailer. The actual magnitude of the impact of textual information on sales may be different for a different retailer. Additional work in other on-line contexts will be needed to evaluate whether review text information has similar explanatory power that are similar to those we have obtained. Despite these limitations, we hope our paper motivates future research in this area. . In the future, we also to extend it with the capability to manage a Cloud environment with multiple data centers. Thus, we will apply a decentralization approach whereby the proposed system is installed on each data center.

REFERENCES

1. The health insurance portability and accountability act of privacy and security rules. url: http://www.hhs.gov/ocr/privacy.

2. National standards to protect the privacy of personal health information url:http://www.hhs.gov/ocr/hipaa/finalreg.html.
3. M. Abramowitz and I. A. Stegun. Handbook of Mathematical Functions: Graphs, and Mathematical Tables. Dover books on mathematics. Dover, New York, 1964.
4. A. Bansal, T. Chen, and S. Zhong. Privacy preserving backpropagation neural network learning over arbitrarily partitioned data. Neural Comput., Feb. 2011.
5. 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.
6. Barni. M., Orlandi, C., & Piva, A. (2006). A Privacy-Preserving Protocol for Neural-Network-Based Computation Trans
7. L. Cun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Handwritten digit recognition with a back-propagation network. In Advances in Neural Information Processing Systems, pages 396–404. Morgan Kaufmann, 1990.
8. S. D. C. di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, and P. Samarati. Over-encryption: management of access control evolution on outsourced data. In Proceedings of the 33rd international conference on Very large data bases, VLDB '07, pages 123– 134. VLDB Endowment, 2007.
9. T. El Gamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In Proceedings of CRYPTO 84 on Advances in cryptology, pages 10–18, New York, NY, USA, 1985.
10. S. E. Fahlman. Faster-learning variations on Back-propagation: An empirical study, pages 38–51. Morgan Kaufmann, 1988.
11. K. Flouri, B. Beferull-lozano, and P. Tsakalides. Training a svmbased classifier in distributed sensor networks. In Proceedings of 14nd European Signal Processing Conference, pages 1–5, 2006.
12. A. Frank and A. Asuncion. UCI machine learning repository, 2010.
13. R. Grossman and Y. Gu. Data mining using high performance data clouds: experimental studies using sector and sphere. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '08, pages 920–927, NewYork, NY, USA, 2008.
14. R. L. Grossman. The case for cloud computing. IT Professional, 11(2):23–27, Mar. 2009.
15. A. Inc. Amazon Elastic Compute Cloud (Amazon EC2). Amazon Inc., http://aws.amazon.com/ec2/#pricing, 2008.
16. R. Law. Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. Tourism Management, 21(4):331–340, 2000.
17. A. J. Menezes, P. C. V. Oorschot, S. A. Vanstone, and R. L. Rivest. Handbook of applied cryptography, 1997.
18. D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Parallel distributed processing: explorations in the microstructure of recognition, vol. 1. chapter Learning internal representations by error propagation, pages 318–362. MIT Press, Cambridge, MA, USA, 1986.
19. [19] 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.
20. S. Stolfo, A. L. P. S. Tselepis, A. L. Prodromidis, S. Tselepis, W. Lee, D. W. Fan, and P. K. Chan. Jam: Java agents for meta-learning over distributed databases. In In Proc. 3rd Intl. Conf. Knowledge Discovery and Data Mining, pages 74–81. AAAI Press, 1997.
21.] Archak, N., Ghose, A., and Ipeirotis, P. G. Show me the money! Deriving the pricing power of product features by mining consumer reviews. In Proceedings of the Twelveth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2007) (2007), pp. 56-65.
22. Breiman, L. Random forests. Machine Learning 45, 1 (Oct. 2001), 5-32.
23. Brown, J. J., and Reingen, P. H. Social ties and word-of-mouth referral behavior. Journal of Consumer Research 14, 3 (Dec. 1987), 350-362.
24. Burges, C. J. C. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 2 (June 1998), 121-167.
25. Caruana, R., Karampatziakis, N., and Yessenalina, A. An empirical evaluation of supervised learning in high dimensions. In Proceedings of the 25th International Conference on Machine Learning (ICML 2008) (2008).

Mr.R.Muthu Vijay Deepak B.Tech.., M.Tech *Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, (e-mail:deepak.r674@gamil.com). Tirunelveli , India, Mobile No:+91-9790433106*

Dr.P.Kumar M.Sc., M.Tech., Ph.d.., *Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, (e-mail: kumarcite@gamil.com). Tirunelveli , India, Mobile No:+91-9943365978*