International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Review on Privacy Preserving Deep Computation Model on Cloud for Big Data Feature Learning

Varsha Govindrao Kulkarni[1], Dr. Kishor Wagh[2]

[1, 2]Department of Computer Engineering, G. H. Raisoni College of Engineering, Savitribai Phule Pune University, Pune, India

Abstract: ***Big Data Analytics and Deep Learning are two high-center of information science. Big Data has become important because of companies i.e. both public and private have been gathering huge measures of domain-specific data, which can contain helpful data about issues, for example, national intelligence, cyber security, fraud detection, marketing, and medical informatics. Deep Learning algorithms extract high-level, complex reflections as information representations through a progressive learning process. A advantage of Deep Learning is the study and learning of huge size of unsupervised information, making it a significant instrument for Big Data Analytics where crude information is to a great extent unlabeled and un-sorted. The present survey gives an idea of the previous work done by several researchers on the Big Data Analytics and Deep Learning applications.***

Keywords: Deep learning, big data.

1. Introduction

Profound learning and Big Data are two most disregarded strategies in today's mechanized world. While Big Data has been portrayed in different ways, in this way it is implied the exponential advancement and wide openness of computerized data that are overseen and broke down by making utilization of customary programming apparatuses and systems. Advanced data, in all shapes and sizes, is creating at stunning rates. For example, according to the National Security Agency, the Internet is taking care of 1,826 Petabytes of data for every day. In 2011, advanced information has grown nine circumstances in volume in just five years and by 2020, its whole on the planet will reach 35 trillion gigabytes. This impact of computerized data brings gigantic chances and transformative potential for various zones, for instance, attempts, restorative administrations industry producing, and instructive administrations. It furthermore prompts to an enthusiastic viewpoint change in our intelligent research towards data driven disclosure.

While Big Data gives the impressive potential to disquieting all parts of our general public, social event of critical data from Big Data is not a typical operation. The incomprehensible and rapidly improvement of information concealed in the striking volumes of non-routine data needs advancement of cutting edge advances and additionally interdisciplinary groups working in close joint exertion. Today, machine learning techniques, together with advances in available computational power, have come to expect a principal part in Big Data investigation and revelation of information. They are used extensively to impact the prescient force of Big Data in fields like web hunt instruments, medication, and space science. As a to an extraordinary degree dynamic subfield of machine adjusting, profound learning is seen as, together with Big Data.

Instead of most traditional learning techniques, which are considered using shallow-organized learning model, significant learning suggests machine learning frameworks that use administered/unsupervised methods to therefore learn various leveled representations in significant outlines for grouping. Persuaded by organic observations on human personality systems for get ready of common signs, profound learning in has pulled in much thought from the scholarly group starting late as a result of its best in class execution in many research territories, for instance, acknowledgment, and PC vision. Profound learning has moreover been successfully utilized as a part of industry things that adventure the unfathomable volume of mechanized data.

Section II gives the Literature review for Big Data Analytics and Deep Learning applications

2. Literature Review

In paper [1] authors, developed a privacy preserving deep learning model for big data feature learning by making use of the computing power of the cloud. The developed system uses the BGV encryption scheme to support the secure computation operations of the high-order back-propagation algorithm efficiently for deep computation model training on the cloud. In our scheme, only the encryption operations and the decryption operations are performed by the client while all the computation tasks are performed on the cloud.

In paper [2] authors developed a privacy preserving back propagation algorithm depending on the BGV encryption technique on cloud. One property of the designed algorithm is to apply the BGV encryption system to the back-propagation algorithm for preventing disclose of private data with cloud computing. Furthermore, the developed algorithm improved the efficiency of massive data feature learning by incorporating the strong power of the cloud computing. In paper [3] authors propose a novel access control model combining Role-based Access Control (RBAC) model, symmetric encryption, and cipher text attribute-based encryption (CP-ABE) to support fine grained access control for big data outsourced in cloud storage systems. We also demonstrate the efficiency and performance of our proposed scheme through the implementation.

In paper [4] authors developed an efficient and privacy preserving single-layer perceptron model, known as PSLP.

The designed PSLP is characterized by employing a Homomorphic paillier cryptosystem, which enables the outsourced medical data are processed on cloud in cipher text without leaking the sensitive medical information. Analysis shows the PSLP really achieves the training target and gets the disease model.

In [5] authors describes some projects and propose that next-generation computing systems for big data machine learning and graph analytics need innovative designs in both hardware and software that provide a good match between big data algorithms and the underlying computing and storage resources. After studying the various project authors concluded, next-generation computing systems for big data machine learning and graph analytics shall take full advantage of hardware accelerators and non-volatile memory, and deliver high-performance computing and storage services to big data applications.

The use of information and communications technologies (ICT) to gain objective presents an opportunity for the development of smart cities, where city management and citizens are given access to a wealth of real time information about the urban environment upon which to base decisions, actions and future planning.

In paper [6] authors developed a framework for the realization of smart cities through the Internet of Things (IoT). The framework uses the complete urban information system, from the sensory level and networking support structure through to data management and Cloud based integration of respective systems and services, and forms a transformational part of the existing cyber-physical system.

Paper [7] analyzes the performance of cloud computing services for scientific computing workloads. Authors quantify the presence in real scientific computing workloads of Many-Task Computing (MTC) users, that is, of users who employ loosely coupled applications comprising many tasks to achieve their scientific goals. Then, they perform an empirical evaluation of the performance of four commercial cloud computing services including Amazon EC2, which is the largest commercial cloud. At the end they compare through trace-based simulation the performance characteristics and cost models of clouds and other scientific computing platforms, for general and MTC-based scientific computing workloads. In paper [8] authors claims to propose that they have developed the first secure and practical multi-party BPN network learning scheme over arbitrarily partitioned data. In developed approach, the parties encrypt their arbitrarily partitioned data and upload the cipher texts to the cloud. The cloud can execute most operations pertaining to the BPN network learning algorithm without knowing any private information. The cost of each party in our scheme is independent to the number of parties. This work tailors the BGN homomorphic encryption algorithm to support the multi-party scenario, which can be used as an independent solution for other related applications.

As shown in table 1, literature review of various papers has been listed, giving possibility of research gap.

Table 1: Survey Table

Sr no.	*Title*	*Publication/ year*	*Techniques*	*Advantages*	*Research gap*
1	Privacy Preserving Deep Computation Model on Cloud for Big Data Feature Learning	IEEE 2016	BGV encryption scheme	efficiently deal with deep computation model for big data	accuracy performance of our scheme is a litter lower than that of the non privacy-preserving deep computation model
2	Privacy Preserving Back-Propagation Based on BGV on Cloud	IEEE 2015	Privacy Preserving High-order Back-propagation	Secure and efficient.	lower accuracy performance
3	Privacy-preserving access control model for big data cloud	ICSEC 2015	symmetric encryption	more efficient and practical deployment in supporting access control	bigger data size and performance is not evaluated
4	PSLP: Privacy-preserving single-layer perceptron learning for e-Healthcare	ICSEC 2015	Paillier Cryptosystem	achieves the training target	Can not deal with more efficient and privacy-preserving big-data medical model training algorithm
5	Big data machine learning and graph analytics: Current state and future challenges	IEEE 2014	present some current projects	need innovative designs in both hardware and software	---
6	An Information Framework for Creating a Smart City Through Internet of Things	IEEE	RFID, WSN and crowd sourcing	identify the key IoT building blocks of smart cities, as well as provide the approaches and resolutions to meet their respective communications, computing and computation requirements	proper business model of smart city is not created
7	Performance Analysis of Cloud Computing Services for Many-Tasks Scientific Computing	IEEE 2011	trace-based simulation	empirical performance evaluation of four public computing clouds	extend the performance evaluation with other real and synthetic applications, toward creating a performance database for the scientific community
8	Privacy Preserving Back-Propagation Neural Network Learning Made Practical with Cloud Computing	IEEE 2014	BPN network learning algorithm	scheme is scalable, efficient and secure	multiparty collaborative learning without the help of TA can be enabled

3. Proposed System

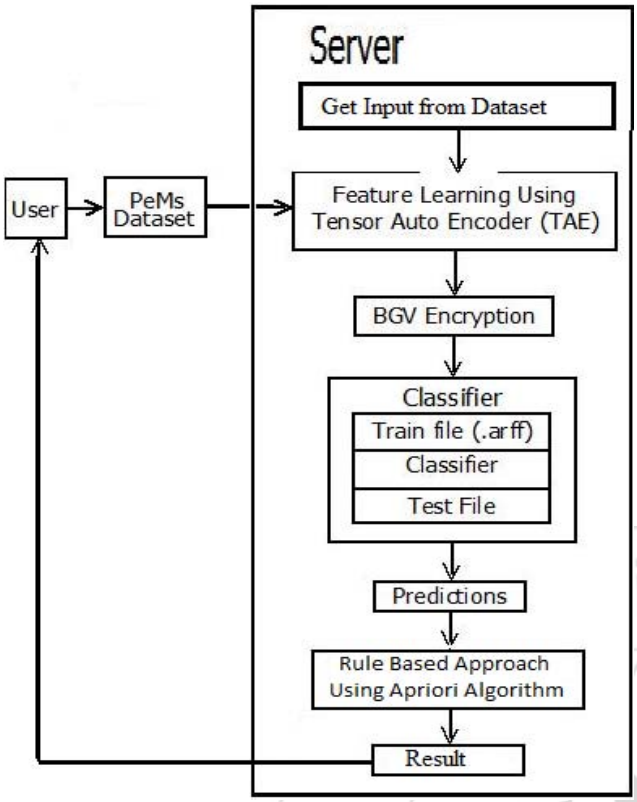

Figure: System Architecture

4. Conclusion

This paper analyses various big data analytics and deep learning systems. Also given the advantages and drawbacks present in the different studies performed by various researchers. To deal with drawbacks in present systems we presented an idea of the new system.

References

[1] Q. Zhang, L. T. Yang and Z. Chen, "Privacy Preserving Deep Computation Model on Cloud for Big Data Feature Learning," in IEEE Transactions on Computers, vol. 65, no. 5, pp. 1351-1362, May 1 2016.
[2] F. Bu, Y. Ma, Z. Chen and H. Xu, "Privacy Preserving Back-Propagation Based on BGV on Cloud," 2015 IEEE 17th International Conference on, New York, NY, 2015, pp. 1791-1795.
[3] S. Fugkeaw and H. Sato, "Privacy-preserving access control model for big data cloud," 2015 International Computer Science and Engineering Conference (ICSEC), Chiang Mai, 2015, pp. 1-6.
[4] G. Wang, R. Lu and C. Huang, "PSLP: Privacy-preserving single-layer perceptron learning for e-Healthcare," 2015 10th International Conference on Information, Communications and Signal Processing (ICICS), Singapore, 2015, pp. 1-5.
[5] H. H. Huang and H. Liu, "Big data machine learning and graph analytics: Current state and future challenges," Big Data (Big Data), 2014 IEEE International Conference on, Washington, DC, 2014, pp. 16-17.
[6] J. Jin, J. Gubbi, S. Marusic and M. Palaniswami, "An Information Framework for Creating a Smart City Through Internet of Things," in IEEE Internet of Things Journal, vol. 1, no. 2, pp. 112-121, April 2014.
[7] A. Iosup, S. Ostermann, M. N. Yigitbasi, R. Prodan, T. Fahringer and D. Epema, "Performance Analysis of Cloud Computing Services for Many-Tasks Scientific Computing," in IEEE Transactions on Parallel and Distributed Systems, vol. 22, no. 6, pp. 931-945, June 2011.
[8] J. Yuan and S. Yu, "Privacy Preserving Back-Propagation Neural Network Learning Made Practical with Cloud Computing," in IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 1, pp. 212-221, Jan. 2014.