International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

IJISAE

ISSN:2147-6799 www.ijisae.org Original Research Paper

Cluster Based Grid Computing with Privacy Preserving Optimization Using Deep Learning Technique

Arvind Kumar Pandey[1], Dr Pankaj M Agarkar[2], Allen Paul L.Esteban[3], Dr V Selvakumar[4] Ankur Gupta[5], Shashikant V. Athawale [6]

Submitted: 15/08/2022 **Accepted:** 20/11/2022

Abstract: Grid computing empowers to involve Grid for enormous scope register and information escalated applications, in science, designing and business. Such applications incorporate, sub-atomic demonstrating for drug configuration, cerebrum movement examination, high energy physical science, protein displaying, beam following and weather conditions determining, etc. The thought behind grouping is to credit the items to bunch so that articles in a single bunch are more homogeneous to different groups.This research propose novel technique in cluster based grid computing with privacy preserving optimization based on deep learning architecture. Here the clustering is carried out using Hadoop based clustering and the privacy based optimization using deep neural network technique. Here the experimental analysis has been carried out in terms of accuracy, precision, data transmission rate, F-1 score. the proposed technique attained accuracy of 95%, precision of 76.5, data transmission rate of 86%, F-1 score of 79%.

Keywords: *Grid computing, clusters, privacy preserving, optimization, deep learning*

1. Introduction

The quick growth of logical applications has prompted the advancement of new age of disseminated frameworks, for example, Grid Computing [1]. Grid frameworks arranges assets that are not expose to concentrated control that implies they are appropriated over networks whose assets are made due, utilized, possessed by a few associations, and are dynamic in nature i.e., Resources and clients can change regularly. This framework consistently gives the admittance to an enormous number of administrations and heterogeneous assets like workstations, organizations, stockpiles, and figuring power that have a place with a few associations and regulatory spaces. Besides, during the last ten years, another age of grid frameworks arose. This new age of grid is known as World Wide Grid (WWG) [2]. Like the notable www, www targets laying out a logical and a computational worldwide grid that anyone, all over the planet, can access and utilize its administrations relying upon his necessities. The upsides of grid framework, for example, dynamicity, the heterogeneity, the dispersion credits of grid assets and the complexity of asset revelation has turned into a test for stretching out grid administration to huge scope frameworks on which the grid framework depends to track down suitable assets for a given work. Consequently in this paper we focus on finding the grid assets for the required positions. Whether it is customary information mining or information examination in a major information climate, bunching, as an essential course of naturally sorting obscure information, can be utilized in the information preprocessing stage as well as in information mining handling. Be that as it may, in the large information climate, bunch examination faces many difficulties. A portion of these difficulties are inborn to the bunching calculation, while others are brought about by the complex information climate [3].

[1] *Professor, Department of Computer Science, Arka Jain University, Jamshedpur, Jharkhand, India. arvind.p@arkajainuniversity.ac.in*
2 SPPU Pune MS, Computer Engineering, Dr ADYPSOE Lohgaon Pune, Pune, pmagarkar@gmail.com
3 Faculty, Graduate School Department, Nueva Ecija University of Science and Technology, Philippines.estebanallenpaul@gmail.com
4 Assistant Professor, Maths and Statistics, Bhavan's Vivekananda College of Science,Humanities and Commerce, Hyderabad-94, Telangana vskselva79@gmail.com,selvakumar.stats@bhavansvc.ac.in
5 Assistant Professor, Department of CSE, RIMT University, Mandi, Gobindgarh, Punjab, India, ankurgupta@rimt.ac.in
6 Department of CE, AISSMS COE, Pune, India, Savitribai Phule Pune University, svathawale@gmail.com

2. Literature Review

Lately, in light of the huge information stage, there has been a ton of exploration work to carry out the conventional information mining calculation in lined up on the disseminated stage and streamline the calculation as per real requirements. To conquer the above issues, work [4] utilized a Map Reduce processing structure joined with the K-choice arranging calculation for equal inspecting to further develop examining proficiency and embraced an example based pre-processing methodology to get the underlying focus highlight get a higher exactness rate. Creator [5] planned map and decrease capabilities to understand the parallelization of the k-implies calculation. Work [6] proposed k-implies nearby optimality. Work [7] proposed a k-implies starting grouping community choice calculation in light of ideal parceling. calculation first partitions the information tests and afterward decides the underlying bunch communities as per the attributes of the example conveyance. Creator in [8] made a thickness based technique. It can really dispose of vagrant places and parallelize them. In light of the issue of an excessive number of cycles of the k-means grouping calculation, [9] in view of the k-implies calculation's qualities of an excessive number of emphasess and too low execution effectiveness proposed a conveyed processing system in view of Spark and applied it in the equal calculation of k-implies text bunching.)en, as indicated by RDD, k-implies prerequisites for complex tasks that should be iterated are addressed. Work [10] applied the MPI equal processing system to the wavelet bunching calculation and proposed the MPI-wave group calculation.

3. Proposed Method

This section discussnovel technique in cluster based grid computing with privacy preserving optimization based on deep learning architecture. Here the clustering is carried out using Hadoop based clustering and the privacy based optimization using deep neural network technique.

Hadoop is an open source execution of Map Reduce equal handling system. Hadoop conceals the subtleties of equal handling, including dispersing information to handling hubs, restarting subtasks after a disappointment, and gathering the consequences of calculation. Here we utilize disseminated document framework in Hadoop (HDFS). Hadoop works on master-slave architecture, HDFS has Name Node and Data Node that works in the similar pattern.

1. Name Node (Master)
2. Data Node(Slave)

Name Node: Name Node return availability information of Data Node when receive the Client request (name node_return_availability). The communication between Name Node and Data Node via heartbeat. After a period time, if Name Node doesn't receive the heartbeat from the Data Node, this Data Node will be considered as dead. Once Data Node has finished the receiving process, it will send done message to Name Node and Client at the same time. The Client gets the done message, it will send done message to Name Node, One data transmission ends.

Data Node: This fills in as a Slave Data Nodes are essentially used for putting away the information in a Hadoop group, the quantity of Data Nodes can be from 1 to 500 or considerably a larger number of than that, the more number of Data Node your Hadoop bunch has More Data can be put away. so it is prompted that the Data Node ought to have High putting away ability to store an enormous number of document blocks. Information hub performs activities like creation, erasure, and so on as indicated by the guidance given by the Name Node.

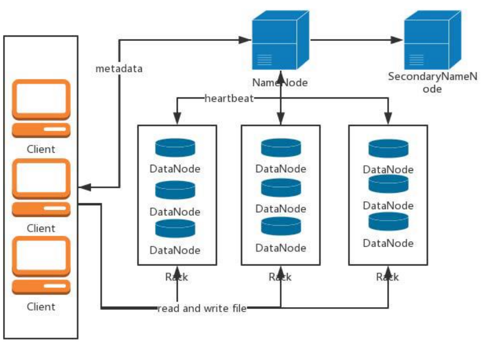

Figure -1 HDFS clustering Architecture

A Hadoop Cluster may contain:

• 40 hubs/rack, 1000-4000 hubs in group
• 1 Gbps data transfer capacity inside rack, 8 Gbps out of rack
• Hub specs (Yahoo terasort): 8 x 2 GHz centers, 8 GB RAM, 4 plates (= 4 TB?)
• Documents split into 128MB blocks
• Blocks reproduced across a few datanodes (normally 3)
• Single namenode stores metadata (document names, block areas, and so on)
• Improved for huge documents, consecutive peruses

We propose ADAM, a procedure for capable stochastic smoothing out that simply requires first-demand tendencies with little memory need. Adam moreover keeps an emphatically decaying typical of past inclines mt, similar to compel. However power ought to be noticeable as a ball running down an inclination, Adam behaves like a significant ball with scouring, which thus incline towards level minima in the goof surface.

$$m_{t,i}` = \frac{m_{t,i}}{(1-\beta_1^t)}$$

$$v_{t,i}` = \frac{v_{t,i}}{(1-\beta_2^t)}$$

where β t1 is β1 power t, β t 2 is β2 power t, and mˆ t,i and vˆt,i are the one-sided rectified first and second minutes, individually. In this way, the boundary update in Adam is consolidated as,

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\alpha_t \times m_{t,i}}{\sqrt{v_{t,i}+\epsilon}}$$

In this paper, an information is addressed utilizing a bunch of provincial highlights. These elements are removed thickly, and their portrayal is given by highlight vectors in the result of the kth brain layer. A Deep brain classifier D1(.) is fitted to all ordinary local highlights created by the FCN (Fully convoluted networks). where h is the size of the component vectors produced by the auto-encoder, which rises to the size of the secret layers. In this step, just the dubious locales are handled. Accordingly, a few focuses (I, j) in grid (wk, hk) are overlooked and not broke down in the grid (w h ' ' ,) k . Like D1, we make a DNN classifier D2 on the ordinary preparation provincial highlights which are all addressed by our peculiarity finder. Those districts which are not adequately fitted to G2 are viewed as unusual. Eqs. (5) and (6) sum up irregularity location by utilizing two fitted Gaussian classifiers. In the first place, that's what we have, $G_1\left(f_k^t(i,j,1{:}m_4)\right) =$

$$\begin{cases} \text{Normal} & \text{if } d\left(G_1, f_k^t(i,j,1{:}m_k)\right) \le \beta \\ \text{Suspicious} & \text{if } \beta < d\left(G_1, f_k'(i,j,1{:}m_k)\right) \\ \text{Abnormal} & \text{if } d\left(G_1, f_k^k(i,j,1{:}m_k)\right) \ge \alpha \end{cases}$$

Then, for a suspicious region represented by $T_k^t(i,j,1{:}h_k)$ we have that:

$$G_2\left(T_k^t(i,j,1{:}h_k)\right) = \begin{cases} \text{Abnormal} & \text{if } d(G_2, T_k^t(i,j,1{:}h)) \ge \phi \\ \text{Normal} & \text{otherwise} \end{cases}$$

4. Experimental Analysis

We utilized Python and Natural Language Tool Kit to prepare and arrange the troupe classifier with bunching based IDS. In absolute we utilized informational collection of size 19340 out of which 18340 were utilized for preparing and 1000 for testing. Tweets are gathered naturally utilizing Twitter API and they are physically clarified as certain or negative.

Table 1: Comparative analysis of Proposed and existing technique

Parameters	MLP	CNN	LBPH	Ha_Clus_DNN
Accuracy	86	87	88	95
Precision	71.6	73.2	75.1	76.5
data transmission rate	73	75	81	86
F-1 score	65	71	75	79

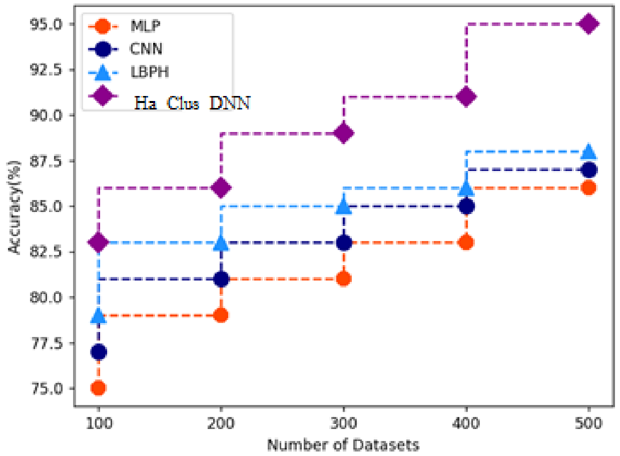

Figure-2 Comparison of Accuracy

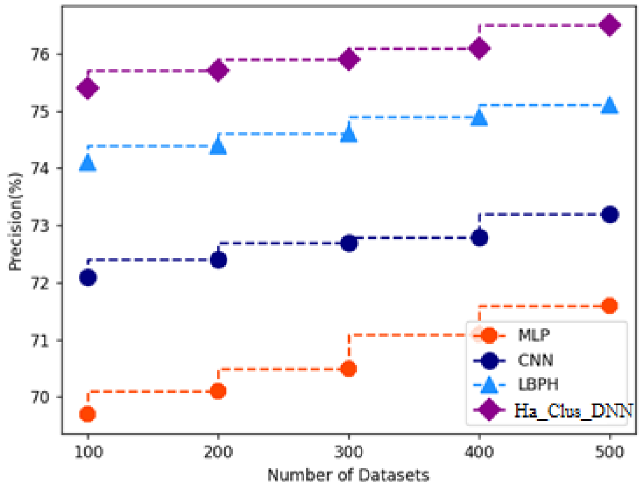

Figure-3 Comparison of Precision

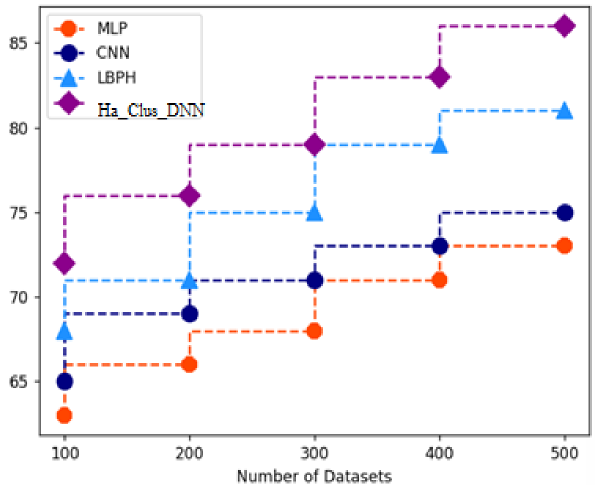

Figure-4 Comparison of data transmission rate

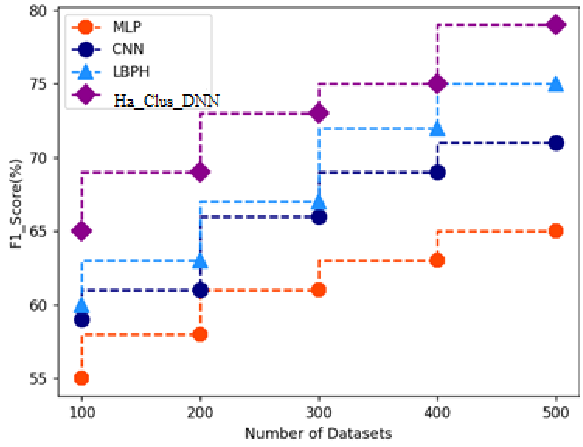

Figure-5 Comparison of F-1 Score

The above table-1 and figure 2-5 shows the comparative analysis of face image classification and feature extraction using proposed **Ha_Clus_DNN**. Here the comparison has been carried out in terms of accuracy, precision, data transmission rate, F-1 score. The existing technique compared are MLP, CNN and LBH among which the proposed technique obtained optimal results in clustering with DNN based optimization. Here the proposed technique attained accuracy of 95%, precision of 76.5, data transmission rate of 86%, F-1 score of 79%.

5. Conclusion

This research propose novel technique in clustering with privacy preserving optimization based on deep learning architecture. The clustering is carried out using Hadoop technique and the privacy based optimization using deep neural network technique. The experimental analysis has been carried out in terms of of accuracy, precision, data transmission rate, F-1 score, where the proposed technique attained accuracy of 95%, precision of 76.5, data transmission rate of 86%, F-1 score of 79%. Subsequently, it deftly decides the quantity of required groups and is unfeeling toward anomalies, while being just component 19x more slow than the present quickest confidential improvement procedure which must be utilized for explicit in data collections.

References

[1] Sai Kumar, S., Reddy, A. R., Krishna, B. S., Nageswara Rao, J., & Kiran, A. (2022). Privacy Preserving with Modified Grey Wolf Optimization Over Big Data Using Optimal K Anonymization Approach. Journal of Interconnection Networks, 2141039.

[2] Bozdemir, B., Canard, S., Ermis, O., Möllering, H., Önen, M., & Schneider, T. (2021, May). Privacy-preserving density-based clustering. In Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security (pp. 658-671).

[3] Dhasarathan, C., Kumar, M., Srivastava, A. K., Al-Turjman, F., Shankar, A., & Kumar, M. (2021). A bio-inspired privacy-preserving framework for healthcare systems. The Journal of Supercomputing, 77(10), 11099-11134.

[4] Kavitha, V. (2021). Privacy preserving using multi-hop dynamic clustering routing protocol and elliptic curve cryptosystem for WSN in IoT environment. Peer-to-Peer Networking and Applications, 14(2), 821-836.

[5] Tumuluru, P., Burra, L. R., BhavaniDasari, D., Saibaba, C. M., Revathi, B., &Venkateswarlu, B. (2021, March). A Novel Privacy Preserving Biometric Authentication Scheme Using Polynomial Time Key Algorithm In Cloud Computing. In 2021 International

Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1330-1335). IEEE.
[6] Wu, X., Zhang, Y., Shi, M., Li, P., Li, R., &Xiong, N. N. (2022). An adaptive federated learning scheme with differential privacy preserving. Future Generation Computer Systems, 127, 362-372.
[7] Singh, S., Rathore, S., Alfarraj, O., Tolba, A., & Yoon, B. (2022). A framework for privacy-preservation of IoT healthcare data using Federated Learning and blockchain technology. Future Generation Computer Systems, 129, 380-388.
[8] Xie, P., Jia, Y., Chen, H., Wu, J., & Cai, Z. (2021). Mixed-stage energy management for decentralized microgrid cluster based on enhanced tube model predictive control. IEEE Transactions on Smart Grid, 12(5), 3780-3792.
[9] Yu, Z., Hu, J., Min, G., Wang, Z., Miao, W., & Li, S. (2021). Privacy-preserving federated deep learning for cooperative hierarchical caching in fog computing. IEEE Internet of Things Journal.
[10] Tian, Y., Zhang, Z., Xiong, J., Chen, L., Ma, J., & Peng, C. (2021). Achieving graph clustering privacy preservation based on structure entropy in social IoT. IEEE Internet of Things Journal, 9(4), 2761-2777.