NFDDC: Novel Neuro-Fuzzy Logic based Methodology for Distributed Data Classification

Shahina Parveen M

Abstract— With the growing culture of Internet applications and their usage lead to challenging task for storing a massive volume of high-velocity data from different fields. This result an evolution of big data with integrated, i.e. Volume, Velocity, and Variety (3V's). The voluminous data extraction is a very complex task which is not possible form classical data mining techniques. Therefore, a big data mining technique is introducing by modifying traditional data mining scheme using a novel of Neuro-Fuzzy Logic based approach, i.e. named as NFDDC. The proposed distributed data classification model performs into three stages first- reduce the data set dimension, second- data clustering, and third-data classification using the neuro-fuzzy method. The performance of the NFDDC system is analysed using two different datasets, i.e. medical data and e-commerce datasets. Additionally, comparative analysis is performed by measuring the accuracy of existing CCSA algorithm with proposed NFDDC algorithm and will get 90% accuracy in data classification.

Keywords—Distributed data Data, Data Mining, Neuro-Fuzzy Logic, Classification.

I. INTRODUCTION

Due to the growing technology in the today's Internet world, the availability of low-cost commodity components with high-speed network technologies is changing the era of computing world (e.g., cloud computing, big data, IoT and many more) [1]. These advanced technologies have the opportunities have led to use distributed computers as unified, single computing system, leading to grid computing [2]. The data grid is emerging integrated architecture which allows connecting of clusters of distributed computing resources located at the different location and facilitates the data and services in a distributed manner. Additionally, the grid computing system can provide multiple applications, such as data exploration, high throughput, collaborative computing and distributed computing [3]. The grid-based distributed computing architecture enables connected computing resources within dynamic platforms and provides parallel processing paradigms for dynamic resource allocation, fast computation with security protocols.

Manuscript published on November 30, 2019.
* Correspondence Author
Dr. Shahina Parveen M*, Information Science Engineering Dept, CMR Institute of Technology, Bengaluru, INDIA. Email id: Shahina.m@cmrit.ac.in

The primary aim of grid computing is to allow organizations and application developers to design and create a distributed computing platform where clusters of computing devices can load and process the data and distribute the workloads by optimizing the time and cost. More clearly, grid computing allows users to achieve more effective and faster results by performing a larger operation within a small period and at a lower cost. From the research study [4], the most likely and significant challenge in the grid computing is data management during processing and storing of large data in the repository which may lead data redundancy problem. The data redundancy will increase by storing of same data at multiple systems. Indeed, data redundancy may lead to several issues by ensuring extensive benefits. [5] Data grid mainly offers data availability by enhancing the operational probability (i.e., grid system contains a data replica if the request is made). Additionally, data grid may also facilitate to distribute the generated or processed data by user request between the grid systems and reduces the data replication possibility with improvising the response time with low energy consumption.

However, extraction of a large set of data from the cloud repository is the challenging task. Therefore, in that context, Data mining becomes a powerful technology which extracts the meaningful information from Big Data [6] [7]. Data mining technology performs various operations like clustering, classification, regression and many more. Among this clasification is the useful approach which classifies the data into different categories based on data features. The classification process performs after the training phase; i.e., training process is especially utilized to create a model for classification which helps for classification in further phase. According to [8], the training process will not get affected by data size. Different authors mainly worked on big data classification problems [9] [10]. The big data classification approach has been introduced by [11], where SVM classifier optimized the performance, and decision tree based algorithmic approach was reconfigured by integrating with k means clustering algorithm.

The study of Distributed data analytics by machine learning approach facilitate high data delivery in a distributed environment [12]. The different clustering techniques for data mining were discussed in the comprehensive study [13]. In this author mainly highlighting distributed data characteristics. An improved approach of data handling and feature extraction from the bulky data sets has been introduced in [14], which provide effective results to the users. The distributed data architecture design and implementation model [15] can optimize the performance through software configuration based on open source.

A neuro-fuzzy distributed data classification model has been introduced by [16], which classify the bulky data based on the feature selection process.

The structure of the proposed research study can be framed as; section II discusses a prior research study in state of the art. Section III defines the problem statement.

Section-IV illustrates about detail system design followed by performance analysis in section V. The last section VI gives the conclusion of the proposed experimental study.

II. LITERATURE REVIEW

This section provides an extensive literature study towards investigating the scale of effectiveness in existing distributed data classification and data mining schemes. Also explores a research gap found from the prior research studies. Many research studies have worked on the problem of distributed data analysis and classification using data mining approach, such as in the study of [17] Patil et al. conducted research on data mining algorithm and knowledge discovery approach by network property. With the help of KEEL software tool, authors analyzed and classified the graphical and structural pattern of data using machine learning approach. Zhu et al. [18], have mainly worked on the analysis of stealing smart meter reading through the wireless network and introduced three different types of data mining algorithms (i.e., Non-intrusive load monitoring algorithm) for distributed data analysis over wireless smart grid. Additionally, authors also analyzed prior wireless data transmission or aggregation methods in smart grid and developed a privacy-preserving module to mitigate the eavesdropping attacks. In large scale grid architecture, the data repository resources are decoupled with other computing resources. Hence that, data transmission time takes tremendous effect on total task processing time and scheduling decision task. Therefore, developing a grid-based resource allocation infrastructure for distributed data computing is a very challenging task. [19] Qureshi et al. addressed a problem of global optimization by considering heterogeneous computing resources with different processing capabilities. Such computing resources are connected with a data repository system by linking networks with different bandwidth rate. Also formulated an analytical model using optimization technique. Consequently, many researchers worked on distributed data mining approach for distributed datasets for example in [20] Angiulli et al. introduced a distributed data mining method, i.e. distance-based outlier detection method for distributed data sets. This approach adopted a parallel data mining scheme in order to generate huge time savings. From the experimental results, authors conclude that proposed outlier detection scheme for large data sets is very efficient w.r.t. processing time increases quite well by adding with extra nodes. More importantly, this approach is well applicable for a distributed environment. The significant challenge of large-scale data mining is to analyze and classify massive data and to design and develop a precise model for each featured data. The distributed data mining is an efficient technique which processes the massive volume of data, but the additional requirement is to minimize the computational processing time without any data loss. Therefore, in that context, [21] Bakry et al. investigated a map-reduce model using the crisp and fuzzy technique. The objective was a deep investigation on distributed data classification scheme and compares both fuzzy and non-fuzzy algorithm using distributed data classification approach. The simulation results of different datasets showed that the proposed map reduce fuzzy method outperform as compare with the crisp method. With the developing of IoT enabling technologies, most of the research study specifically interested in identifying risk problems emerging while discovering and integrating the useful data within the IoT infrastructure [22]. Furthermore, at present, distributed data scenario, different standard data classification algorithms and software modules have been introduced [23] and are utilized for data analysis, classification as well as data aggregation. The social IoT has become a larger social network which supports multiple applications and advances networking services for IoT in a very powerful and productive way. In [24] Lakshmana Prabu et al. introduced a hierarchical model for social IoT systems using map-reduce and supervised classification paradigm. More even, a Gabor filtering scheme is applied to reduce the noisy and redundant as well as unwanted data from the data repository. A map-reduce scheme is applied for mapping and reducing the large-scale databases and improve the system efficiency. One of the fastest growing and more adaptable technologies in the urban area is the smartphone usage. Hossain and Muhammad [25], introduced a communication framework for urban environment distributed data classification using deep learning approach. This method classifies the noisy environmental data from the urban distributed data and reduces the noisy effect during having communication through smartphones. From the experimental analysis, the author demonstrated that propose deep learning algorithm efficiently able to classify the noisy environmental data. Piczak [26] evaluated the potentiality of convolutional neural network performance by classifying the short audio clips generated from the surrounding environment. The network accuracy was evaluated based on three public data samples collected from the surrounding environment and urban recordings. The standard machine learning algorithms and methods allow us to predict and classify the data from the processed data and solve the various issues in the area of data mining in programmatical mode. The integrated technology of distributed data mining and machine learning schemes are considered as significant research methodology which solves the cloud data analysis problem for IoT mobile data. The research work of IoT mobile data monitoring and security framework was developed by Kotenko et al. [27]. In this study, a combined approach of distributed data processing and machine learning concept was introduced for IoT mobile security monitoring. Also, an IoT security framework specified numerous machine learning schemes intended to address classification jobs. Another Distributed data security and privacy module were introduced Xu et al. [28] with the aim to solve the privacy risk obtained from the data mining process. Particularly, the authors identified four kinds of manual data mining applications (i.e., data provider, collector, data miner, and decision maker).

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-4, November 2019

III. PROBLEM STATEMENT

The last two decade of communication media has witnessed the rapid use of the internet in every application which has to lead the drastic increment in the volume, a variety of data, i.e., distributed data. Currently, the biggest concern is arising about extracting this large voluminous distributed data and which is expected to be increased in the future. Also, the extraction of data in profit perspective is needed to be optimized with the data mining approaches. However, due to increasing volume, velocity and variety (3V's) in the distributed data, it's quite a tough task to mine the data more effectively. The data mining includes the operational stages like clustering, classification, association, and regression, etc. The prime intention of the paper is to classify the distributed data. The classification of data involves steps like training and classification. During the training process, a model is built to classify the data in the classification step. The various researches have suggested that the classification is not much concern about the amount of data, but this data is affected in the classification process. Various research tried to classify the large voluminous distributed data but failed to give an optimized performance in classification. Also, the traditional approaches of data mining associated with distributed data (considering 3V's) were failed to classify properly even with other modern techniques. Thus, there is a need of classification mechanism which can handle 3V's of distributed data and perform classification more effectively by adopting a systematic approach.

IV. SYSTEM DESIGN

This problem associated with 3V's of distributed data were handled with developing a novel classification model which considers Neuro-fuzzy logic-based algorithm and reduce 3V's and perform data mining in terms of classification. This paper aims to design the same technique and is named as "Novel neuro-fuzzy logic-based approach for distributed data classification (NFDDC)." The design of this NFDDC model is performed in three different stages like Stage-1 for reducing the data set dimension, Stage-2 for clustering of the reduced data sets and Stage-3 for classification of data by adopting fuzzy logic generated rules through attributes with the neuro-fuzzy logic system. The architectural model of the NFDDC is presented in Figure.1 that indicates all the internal processes involved in it. The details of each stage are given below.

A. Stage-1: Reducing the data sets dimension:-

The large dimensional distributed data generated due to various attributes exist in the dataset. The perception analysis is performed to distinguish the attributes based on their features. In case, the input attribute has the number of classes (Cn) for the number of attributes (An) of the data set. The reduction process of 3V's is derived by using Eq.1 to Eq.7. The mean of the desired class can be derived by using Eq.1 while Eq.2 is derived to compute the mean of the entire distributed dataset with "An"-a number of attributes. in indicates the number of attributes at the class 'i'.Every stage contributes its input to the next stage in a cascading fashion. The architectural design yields better performance with the improvement in measure of parameters for both datasets.

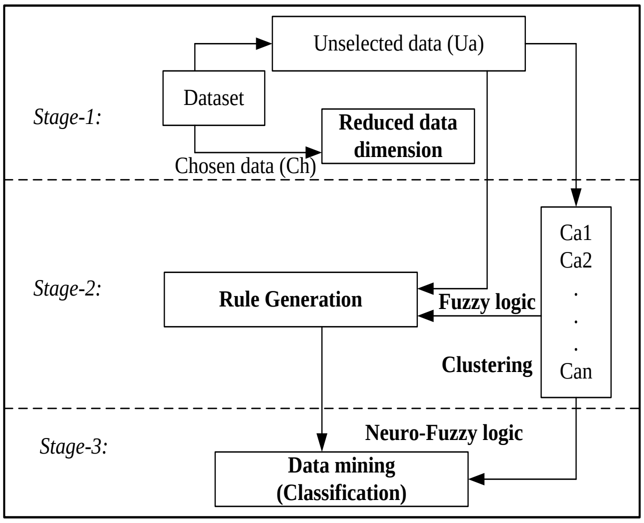

Figure.1. The stage-wise architecture of system design

$$m_i = \frac{1}{i_n} \times \sum An_i \quad \text{...Eq. (1)}$$

$$m = \frac{1}{n} \times \sum_{i=1}^{n} A_i \quad \text{...Eq. (2)}$$

The scattered matrix (S_M) among the class (aS_M) can be calculated as:

$$aS_M = \sum_{i=1}^{C_n} Q_i (m - m_i)(m - m_i)^t \quad \text{...Eq. (3)}$$

Where,

$$Q_i = \frac{i_n}{C_n} \quad \text{...Eq. (4)}$$

The scattered matrix (S_M) within the class (wS_M) class can be calculated as:

$$wS_M = \frac{1}{A_n} \sum_{i=1}^{C_n} S_i \quad \text{...Eq. (5)}$$

The value of S_i can be calculated as:

$$S_i = \sum_{i=1}^{C_n} (An_i - m_i)(An_i - m_i)^t \quad \text{...Eq. (6)}$$

The total scatter matrix (TS_M) can be calculated as:

$$TS_M = aS_M + wS_M \quad \text{...Eq. (7)}$$

The eigenvectors (E) of TS_M can be used to choose the attributes (Ic) as chosen attributes (Ch) and Unselected attributes (Ua) etc. and is given in Eq. (8)

$$Ic = \min(Ev(TS_M)) \quad \text{...Eq. (8)}$$

B. Stage-2: Clustering and Rule generation (Fuzzy logic):-

This stage of the design involves two modules involving clustering and fuzzy logic-based rule generation. The clustering module uses Schwarz based K-mean algorithm on 'Ch' and Rule generation uses rough set theory. The detailed design of this stage is given below.

i Clustering module This module is considered to cluster the data (Ch) with the help of the algorithm given below:

Algorithm: Clustering module

Start:
L-1: initialize → K
L-2: Apply K-mean →Generate clusters (Ca_1, Ca_2 …Ca_n)
L-3: Every test case→ data Tc
L-4: Derive→ Schwarz creation (Sc) for Ca_i

$$Sc = -2 \times l_n \hat{L} + kl_n(Ca_n) \quad \text{…Eq. (9)}$$

L-5: Apply K-mean →Ca_1 clusters
L-6: Derive → Schwarz creation (Sc) for Ca_1 and Ca_2

$$Sc_1 = -2 \times l_n \hat{L} + 2l_n(n) \quad \text{…Eq. (10)}$$

L-7: If Sc > Sc1
n=n+1
$ai=a_{i1}$ and $an=a_{i2}$
End
L-8: i=i-1
End

The clustering algorithm is initialized (L-1) by considering the default smallest value (K=2). Then the K-mean is applied for generation of clusters (Ca) (L-2). Similarly, for each test cases of dataset the Schwarz criterion (SC) is computed for each cluster by using Eq. (9) That includes Ca_n as a number of total elements in Ca, the maximized value of the model function (L), i.e., $\hat{L} = p(x) | \hat{\theta}, C)$ in which $\hat{\theta}$ indicates the parameter value to maximize function (L-3 to L-6). Further, K-mean is subjected to get clusters. The Sc of clusters is computed by using Eq. (10) Where the parameters are double because of clusters. A condition is applied to the new model (L-7).

ii. Rule generation module:

This module works on the unselected datasets of the clustering module, i.e., Ua. The module considers the fuzzy logic for dynamic rule generation, and the implemented algorithm is represented below:

Algorithm: a Rule generation module

Start:
L-1: Scan → Cd
L-2: Indicate→1 for Ch and 0 for Ua
L-3: Count → frequency (Ua)
L4: Generate → common item sets
L-5: Check→ data transaction=0
L-6: Add→support value (V) →weight
L-7: Calculate →weight (min & max)
L-8: Compute → distance (D) using Ed formula
L-9: Generate → value (Dv)
L-10: Initalize→t=1
L-11: Compare→ D & Dv
L-12: if V> Dv
L-13: Perform → data encoding
L-14: Encoding format→ binary
L-15: Perform → Offspring
L-16: Generate→Set of rules
End

The algorithm for rule generation scans all the data (L-1) and are represented with 1 and 0 for Ch and Ua respectively (L-2). The frequency of Ua is counted (L-3) and generated the common item sets (L-4). Further, a checklist is done for data transaction=0 (L-5). A support value (V) is added as weight and computed the minimum and maximum by using weight (L-6 to L-7). Then, the distance is calculated by using Euclidean (Ed) formula and compared the value distance (D) by referring generated distance vector (Dv) using initialized by population set (t) value (L-8 to L-10). Further, it is checked that V>VD and performed the encoding in binary format (L-11 to L-14). The operation offspring is performed to generate the set of rules (L-15 to L-16).

C. Stage-3: Classification stage:-

This stage is used for classification by adapting interfaced neuro-fuzzy logic system and by rules generated. The following Figure.2 indicates the architecture of interfaced neuro-fuzzy logic system.

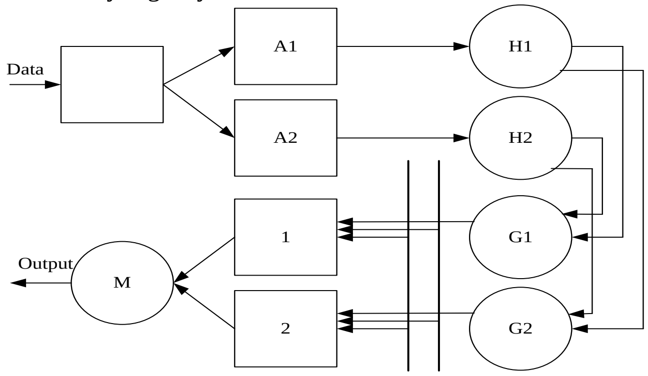

Figure.2. The architecture of interfaced neuro-fuzzy algorithm

In this, the membership function is utilized for fuzzy of the system as shown in Figure.3. The membership function has least (Li), lower (Lo), medium (M), high (H) and highest (Hi) values. The outcomes of this stage are analyzed for each test case of each dataset in results analysis.

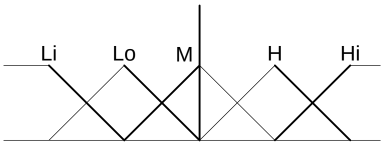

Figure.2. Membership function used in fuzzy logic

V. PERFORMANCE ANALYSIS

The performance of NFDDC algorithm is analyzed by implementing it on eclipse IDE. The analysis is carried out over the dataset consisting of 400 test cases (Tc) having 25 attributes (medical dataset) and 32.561K test cases for 15 attributes (E-commerce dataset). The description of the dataset is represented in Table.1.

Table.1. Description of Dataset

Description / Type of data	An	No. of Tc	Cn	Attribute Type	References
Medical	25	400	2	Nominal/ numeric	UCI Repository
E-commerce	15	32.561 K	2	Nominal/ numeric	UCI Repository

The performance of NFDDC system is analyzed by executing the algorithms overs the datasets through parameters like accuracy, true positive (Tp), false positive (Fp) rate, recall (Re), F-measure and precision (Pr). The obtained parameter analysis for both the data set is given in Table.2.

Table.2. Performance analysis

Param-eter	Accuracy (%)	Tp	Fp	Pr	Re	F-m
Medi-cal	99.75	99.8e-2	0.2e-2	99.8e-2	99.8e-2	99.8e-2
E-commerce	88.025	88e-2	0.22e-2	88e-2	87.3e-2	87e-2

The comparative analysis of the proposed NFDDC system is compared with decision tree (J48) and constant criterion surface algorithm (CCSA). The outcome is given in Figure.4 by considering the accuracy parameter.

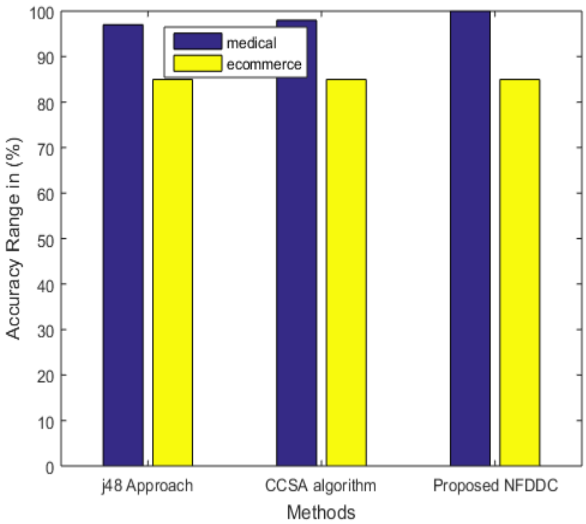

Figure.4. Accuracy analysis

From Figure.4, it is observed that the accuracy of the proposed NFDDC system is compared with decision tree (J48) and constant criterion surface algorithm (CCSA) over medical and E-commerce dataset. From the analysis it is observed that the proposed NFDDC system has improved the accuracy of about 87% and 90% than existing j48 and CCSA algorithm respectively. This improvement can represent the enhanced elements mapping with the correct class.

The comparative analysis in terms of Tp and Fp value is represented in Table.3 for all the three methods where it is observed that proposed NFDDC system has enhanced Tp value and decreased Fp value that gives the better operation of receiver characteristics than existing CCSA and J48 systems.

Table.3. Tp and Fp analysis

Parameter	Analysis	J48	CCSA	Proposed NFDDC
Medical	Tp	0.96	0.98	0.99
	Fp	0.97	0.99	0.99
E-commerce	Tp	0.80	0.82	0.85
	Fp	1.1	1.09	1.05

Similarly, the comparative analysis in terms of recall parameter specifying the output relevancy is given in Figure.5, where it is observed that the proposed NFDDC system has more relevant data than CCSA and J48 algorithm.

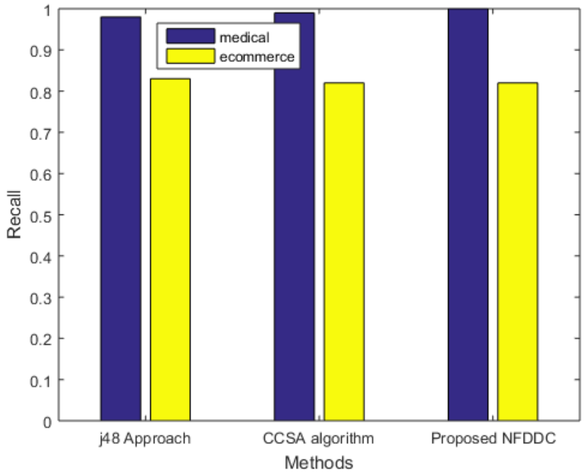

Figure.5. Recall analysis

Following, the comparison of test accuracy is measured in terms of F-measure that gives the weighted harmonic mean indicating the algorithm effectiveness (in table.4). From the table, it is observed that the proposed NFDDC has got better F-measurement than existing J48 and CCSA algorithm which indicates that the proposed system is more effective than others.

Table.4. F-measurement analysis

Parameter	Medical	E-commerce
J48	0.98	0.83
CCSA algorithm	0.99	0.82
Proposed NFDDC	1.1	0.85

VI. CONCLUSION

The proposed study introduced a novel of Neuro-Fuzzy Logic-based Approach for Distributed data Classification (i.e., NFDDC) and compared the results with existing J48 and CCSA algorithms using two different datasets. The aim was to reduce 3V's and perform data mining scheme by analyzing and classifying distributed data. The proposed NFDDC algorithm is performed by classifying distributed large datasets. Also, can observe that the classification accuracy was improved up to 90% which indicates that the proposed algorithm is more effective than others.

REFERENCES

1. Cannavaro, M., Talia, D.: The knowledge grid. Commun. ACM 46(01), 89–93 (2003)
2. Zhuge, H.: China's e-science knowledge grid environment. IEEE Intell. Syst. 19(1), 13–17 (2004).
3. Sánchez, Alberto, Jesús Montes, Werner Dubitzky, Julio J. Valdés, Maria S. Pérez, and Pedro de Miguel. "Data mining meets grid computing: Time to dance?." Data Mining Techniques in Grid Computing Environments (2008): 1.

4. Hamrouni, Tarek, Sarra Slimani, and F. Ben Charrada. "A survey of dynamic replication and replica selection strategies based on data mining techniques in data grids." Engineering Applications of Artificial Intelligence 48 (2016): 140-158.
5. Hamrouni, Tarek, Sarra Slimani, and Faouzi Ben Charrada. "A data mining correlated patterns-based periodic decentralized replication strategy for data grids." Journal of Systems and Software 110 (2015): 10-27.
6. Han, J., Kamber, M., Pei, J., 2011. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers.
7. Zaki, M. J., Wagner Meira, J., 2014. Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press.
8. Li J, Xu Z, Jiang Y, Zhang R. The Overview of Big Data Storage and Management. 2014 13th International Conference on Cognitive Informatics and Cognitive Computing, London, ICCICC'14. 2014. p. 510 –13.
9. Pitre R, Kolekar V. A Survey Paper on Data Mining with Big Data. International Journal of Innovative Research in Advanced Engineering, (IJIRAE). 2014; 1(1):178–80.
10. Fan W, Bifet A. Mining Big data: Current Status and Forecast to the future, SIGKDD Explorations. 2013; 14(2):1–5.
11. Cavallaro G, Ridel M, Richerzhagen M, et al. On understanding big data impacts in remotely sensed image classification using support vector machine methods. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2015; 8(10):4634–46.
12. Thiyagarajan VS. Platfora method for high data delivery in large datasets. Indian Journal of Science and Technology. 2015; 8(33):1–13.
13. Sajana T, Sheela Rani CM, Narayana KV. A Survey on Clustering Techniques for Big Data Mining. Indian Journal of Science and Technology. 2016 Jan; 9(3). Doi no:10.17485/ijst/2016/v9i3/75971
14. Karthick N, Agnes Kalarani X. An Improved Method for Handling and Extracting useful Information from Big Data. Indian Journal of Science and Technology. 2015 Dec; 8(33) Doi no:10.17485/ijst/2015/v8i33/60744
15. Kyoo-sung N, Doo-sik L. Bigdata Platform Design and Implementation Model. Indian Journal of Science and Technology. 2015 Aug, 8(1 8). Do: 10.17485/ijst/2015/v8i18/75864.
16. Gill, Nasib Singh. "An Optimized Algorithm for Big Data Classification using Neuro Fuzzy Approach." Indian Journal of Science and Technology 9, no. 28 (2016).
17. Patil, Shilpa, PS Ashok Kumar, Prasadgouda Patil, and Puneet Palagi. "Data analysis & classification methodology for knowledge discovery in big data." In Inventive Systems and Control (ICISC), 2017 International Conference on, pp. 1-4. IEEE, 2017.
18. Zhu, Liehuang, Meng Li, Zijian Zhang, Xiaojiang Du, and Mohsen Guizani. "Big Data Mining of Users' Energy Consumption Patterns in the Wireless Smart Grid." IEEE Wireless Communications 25, no. 1 (2018): 84-89.
19. Qureshi, Muhammad Bilal, Mohammed Abdulrahman Alqahtani, and Nasro Min-Allah. "Grid resource allocation for real-time data-intensive tasks." IEEE Access 5 (2017): 22724-22734.
20. Angiulli, Fabrizio, Stefano Basta, Stefano Lodi, and Claudio Sartori. "Distributed strategies for mining outliers in large data sets." IEEE transactions on knowledge and data engineering 25, no. 7 (2013): 1520-1532.
21. Hegazy, Osman, Soha Safwat, and Malak El Bakry. "A mapreduce fuzzy techniques of big data classification." In SAI Computing Conference (SAI), 2016, pp. 118-128. IEEE, 2016.
22. S. F. Ochoa, G. Fortino, and G. Di Fatta, ``Cyber-physical systems, Internet of Things and big data,'' Future Generat. Comput. Syst., vol 75, pp. 82 84, Oct. 2017.
23. M. M. Rathore, A. Ahmad, and A. Paul, ``IoT-based smart city development using big data analytical approach,'' in Proc. IEEE Int. Conf. Autom. (ICA-ACCA), Oct. 2016, pp. 1 8.
24. Lakshmanaprabu, S. K., K. Shankar, Ashish Khanna, Deepak Gupta, Joel JPC Rodrigues, Plácido R. Pinheiro, and Victor Hugo C. De Albuquerque. "Effective Features to Classify Big Data Using Social Internet of Things." IEEE Access 6 (2018): 24196-24204.
25. M. S. Hossain and G. Muhammad, "Environment Classification for Urban Big Data Using Deep Learning," in IEEE Communications Magazine, vol. 56, no. 11, pp. 44-50, November 2018.
26. Piczak, Karol J. "Environmental sound classification with convolutional neural networks." In Machine Learning for Signal Processing (MLSP), 2015 IEEE 25th International Workshop on, pp. 1-6. IEEE, 2015.
27. Kotenko, Igor, Igor Saenko, and Alexander Branitskiy. "Framework for Mobile Internet of Things Security Monitoring Based on Big Data Processing and Machine Learning." IEEE Access 6 (2018): 72714-72723.
28. Xu, Lei, Chunxiao Jiang, Jian Wang, Jian Yuan, and Yong Ren. "Information security in big data: privacy and data mining." IEEE Access 2 (2014): 1149-1176.

AUTHOR PROFILE

Dr. Shahina Parveen M is working as Associate Professor, ISE department, CMR Institute of Technology, Bangalore. She has got 14 years of teaching experience. She has obtained Bachelor of Engineering from JNT University in the year 2005. She studied Masters of Technology from ANU, AP. She was awarded Ph.D in the year 2019 in CSE at JNT University, Hyderabad, India. She has published 12 papers in both national and international conferences and Journals, such as Springer and CAE, USA. Published two Patent in September, Oct 2019.