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

Density Based Spatial Clustering Application with Noise by Varying Densities

Vikram Neerugatti, Mokkala Kiran Moni, Rama Mohan Reddy A

Abstract: ***Cluster algorithms are used for grouping up of similar points to form a cluster. It has seen mostly in Machine Learning algorithms. The most popular density-based algorithm is DBSCAN. DBSCAN can find the clusters, irrespective of its shapes and sizes of a cluster. DBSCAN algorithm can easily detect the noise in a clustering dataset. In the proposed algorithm we developed a model based on the existing dbscan algorithm. In the developed algorithm we focus mainly on the epsilon parameter value. Whenever the dbscan algorithm fails to form a cluster we increase the epsilon value by half of its original size. We repeat this step until a cluster is formed. Whenever a cluster is newly formed we change existing epsilon parameter value by adding the 10 percent of the previous used epsilon parameter value. We use epsilon for varying the density of a cluster. So, we can use the dbscan algorithm with the varying density values for developing a cluster. We applied this algorithm on the various datasets.***

Keywords: DBSCAN, Clustering, Data mining, Algorithm, Machine Learning.

I. INTRODUCTION

Datamining is used to discovering the patterns in huge data that involves in the methods of statistics, machine learning and database systems [1].

It is inter-disciplinary part of the computer sciences and statistics it extracts information from the data and transforms to a comprehensible structure [1][2][3]. It is the analyzation process of KDD [4]. It does involve in the data management, database aspect and pre-processing of the data and post processing of the structures [1].

The task of the datamining is to extract the data and data records of interesting patterns, by analysing the large amount of data. This analysis uses the various typesofmethods such as spatial indices. These are kind of patterns that are shown as the input and it can also be used as the analysis for the future. Both the data preparation, data collection is used for the interpretation of the results, which belongs to datamining, which is a part of the KDD.

Manuscript published on November 30, 2019.
*Correspondence Author
VikramNeerugatti*, Department of CSE, Sri VenkateswaraColllege of Engineering, Chittoor, India. Email: vikramneerugatti@gmail.com
MokkalaKiramMoni, Department of CSE, S,V University, Tirupati, India. Email: kiranmonireddy@gmail.com
Rama Mohan Reddy A, Department of CSE, S,V University, Tirupati, India.. Email: ramamohansvu@yahoo.com

Data mining provides services that are used for gathering of information, where the data goes through some sort of identification process. At the end of the process, it determines all the characteristics of the data mining. We are collecting the data from the simple to more difficult information, that collects in the form of digitally [5]. Some of the areas are Computer-aided drafting, Scientific data or from the Scientific applications, Medical, GIS images, Sensor data, Cloud data etc... There are various problems that arises in data mining such as Security concerns, Social issues, User interface issues, Mining etc…

Datamining, Machine Learning and Artificial Intelligence, these are all the topics that are inter-related and sometimes datamining is referred to be a part of the Machine learning. Machine Learning will learn algorithms and statistical models so that, it can perform a specific task without the help of any explicit instructions. Machine Learning methods will be used to develop a model to perform the specific task and can make the predictions are the decisions accurately without taking any help from outside [6][7]. It will help in developing the models that gives the best output with the greater performance and accurate results based on the previous data that is being analysed or monitored and then that is given to machine learning program based on this data the machine learning program will develop an optimum model. The model may have several parameters. So, we give the past data to the model so that the model can be trained and in the future the model can give the accurate results. The model can predict the outcome.

There are numerous advantages of the Machine Learning and also has its own disadvantages. Firstly, we see the advantages they are: Identifying the trends and patterns, Human intervention is not required, Continuous growth, it is used in developing the models for the large applications. It is well mannered in maintaining the multi variety, dimensional data. Some of the disadvantages of the Machine Learning are: Acquisition of the data, it requires large number of resources, it requires more time, it requires the interpretation of the results, high errors may occur. Finally, some the Machine Learning applications are: Recognition of images, Speech Recognition, Diagnosis in the medical field, Robotics, Forecasting, Association learning, Classification, Predictions, Regression.

Clustering is one of many types of unsupervised learning techniques. Unsupervised learning is defined as were we drawing the data from the data set and that will have the input data without any labels. Clustering can be defined as the grouping of objects that are similar points that are belonging to the same group and dissimilar point belongs to different groups and otherwise considered to be as outliers or noise. In short it can be defined as object that are similar or dissimilar.

Clustering is very much useful, it groups the various unlabelled data. There are no criteria for a good clustering. For example, we show some interest in the finding of homogeneous groups for finding the "natural clusters" and show the properties in finding the useful groups or in finding dissimilar objects.

Clustering is tasks of grouping set of objects that are similar are put into a same group. The main task to achieve is of the exploratory data mining, statistical type of analysing the data. It is used in differentareas, some of them areDiscovery of the Knowledge [8], Pattern type Recognition [9], Image Processing [10], Bio-informatics [11], Machine Learning [12] [13], Computer Graphics [14], spatial databases [15] provides the challenges to identify the cluster of various shapes. Clustering is not a single algorithm, various types of algorithms of the cluster available. So, that the given task is to be solved.

DBSCAN or Density-Based on Spatial type Clustering of Application that includes Noise[16], is a density-based technique which is used to discover the clusters irrespective of their shapes and handle the noise effectively. It requires two input type of parameters given by user**.**

The DBSCAN is required because the clustering techniques like partition clustering technique, connective based clustering technique, fuzzy based clustering technique and so on will find the clusters that are of the spherically shaped clusters or the convex type clusters. Because of this they are largely affected by noise and the outliers that are present in the data. The data of real time may contain many problems or dissimilarities:Cluster can be of any type of shape, Data may contain noise. Some of the advantages of this method are:Discovering the clusters that are of the different types of shapes, identifying the clusters that are covered by the other types of clusters, DBSCAN is Robust towards the outlier detection.Some of the Disadvantages are:It is not well suited for multiprocessor type of systems, Data sets with varying densities are difficult to perform, Minimum point value and Epsilon value are carefully determined these values are very sensitive, DBSCAN may fail to identify the clusters with the different densities, Density will be affected if we do the sampling on data set.Some of the Applications of DBSCAN are:Scientific applications, GIS images, Satellite sensing, Medical diagnosis, Surveillance type of videos, Sensor data, Marketing analysis, Finances, Business transactions.

II. RELATED WORK

DBSCAN can discover the clusters of different shapes without any noise. Density estimator [17], uses sub groups method and assembly technique to obtain correct clusters and the computational complexity is lesser than other methods. A new clustering method for collaborative filtering [18], using dbscan it uses to improve the prediction accuracy and is proved to be more effective.A VDBSCAN [19], is same as the dbscan algorithm the only major difference between the dbscan and vdbscan is that in the vdbscan it is going to take the epsilon parameter value for the various densities that are shown in the k-distance plot. The DBSCAN-GM [20], it is a combination of GM and dbscan methods. In this method the purpose of Gaussian-Means is to generate the clusters of small size with the estimated centre of the clusters for to determine the parameters of dbscan such as MinPts and Eps and then dbscan algorithm is performed. This method produces better performance and quality of clusters. It handles noises. It provides better accuracy and validity measures.

QIDBSCAN [21], is also a dbscan algorithm rather is an extended version of the dbscan algorithm that is called the quick dbscan algorithm. It is a newly formed from the data type clustering were the data from the large type of data sets so that the data can be mined using this algorithm. We can compare this method with the conventional dbscan algorithm were this method is much faster than the dbscan algorithm this method kind of eliminate all the unnecessary steps such as expanding the seed in dbscan algorithm because the algorithm will increase the speed if we start expanding with selecting and points. The experiment shows the appropriate results that the Quick dbscan algorithm works more efficiently with accurate results.BEDBSCAN [22], method is the one of the various extensions of the dbscan algorithm and in differ to the dbscan algorithms the BEDBSCAN algorithm will expand the size of the cluster based on where it will consider the object that are available in the border as the seeds. This method reduces judgements to objects that belongs to the core and the quires that should be in the same region that of the neighbour, based on this on the dataset it will be reducing the travelling time and this may produce the higher running efficiency and the cost of the input and the output will be reduced and the speed will be greatly enhanced. It shows high effectiveness and feasible to large scale data processing.A regular DBSCAN method has a disadvantage of inability of identifying clusters with different densities to overcome this a new dbscan based method has developed i.e., the EXDBSCAN [23], were the data sets that consists of multi density will be covered by the exdbscan. Here, this algorithm will going to take the parameter for detecting the clusters where the parameter will be given by the user and then each cluster may have densities varied. In this algorithm effectively, outliers will be identified. Here, it shows the clusters were the densities were varied accordingly, and in this algorithm, we are going to use the techniques such as greedy method so that, we can easily find the density of the cluster with the help of the parameter. A new algorithm was developed based on the traditional dbscan algorithm ADBSCAN [24], here it is called adaptive dbscan algorithm were in this algorithm densities of the algorithm for each iteration will be varied on the data set. In this newly changed version of the dbscan method, we will be finding the various epsilon values for the various densities on the data set with respective to distribution of the data objects. Here, in this algorithm it also modified the number of minimum points to be used in the algorithm and it provides the better results for the clusters that are generated after performing the adaptive dbscan algorithm and this method can detect the cluster with different density-levels. The K-DBSCAN [25], is mainly focuses on identifying the clusters with various shapes that are in regions of the variable density and the k dbscan algorithm can focuses on mainly on the points that are of the clusters are identified that are of spatial density or very similar and in this algorithm it can identify the different type of clusters that or overlapped in various spatial regions at different levels of the density and in these case the dbscanalgorithm will be failed for the identification of the overlapped clusters.

In the k dbscan algorithm will identify the different levels of the density based on dividing all the points of the data into the various levels of the density. So, that it identifies the clusters and extract it with the help of a newly changed version of the dbscan algorithm. This k density-based algorithm is showing the more effectiveness.

The DBSCAN-GM identify the clusters and noises, the shape of the clusters will also be very different. Here, the points are also mentioned very doubt full. Because of this the regions of the data may as well be over lapped it means that the point may belong to many different clusters that are already over lapped. Based on these problems a new algorithm is developed or it can be considered as the regular dbscangm algorithm where it originally deals with the points that are soft. This kind of algorithm is DBSCAN-GM based on the soft (SDG) [26], this newly developed is a combination of the two methods that is the dbscangm and the other one it is uses is that the fuzzy set by combining these two the new algorithm is developed and it gives the results of the clusters with a very good quality and it works very effectively.

In dbscan algorithm specifying the parameter to be used in the dbscan algorithm is very hard to determine and the parameter value will determine how many numbers of clusters to formed indirectly this parameter values plays a crucial role in the algorithm. For this a new method was developed on the basis of the density mechanism that is it depends on the objects of their distance in between them includes with theknn based on the centroid [27], and it will not mention the nearest neighbour and also optimization is obtained earlier. These results are very effective to when compared to the other density algorithms.

Density based on the spatial clustering type of the applications that includes the noise data is a data clustering algorithm, it is one of the best performance algorithms however, it shows the reduced performance of clusters of different densities. So, a new method was introduced for clusters with varying densities i.e. Adaptive type Density based on the spatial type clustering of the applications that includes the noise [28], it will identify different type of clusters with varying densities. This method outperforms existing DBSCAN algorithm.

The dbscan algorithm is focused on detecting the unwanted type of data and this algorithm is developed based on the density. The dbscan method is highly sensible to choosing the parameter values that are minimum points to be used and epsilon value. So, a new type of method was developed for configuring the appropriate parameter values to be used in the dbscan algorithm [29], it can detect the appropriate parameters based on considering the various characteristics from the statistics that applies on the dataset that is in the form of a graph. It is carried out in visual display by using MATLAB tools. This method shows the detection results more accurate and scalable but even this method requires the human participation in judgement for parameter setting in the DBSCAN algorithm. But this method also produces trouble in the form of the time consuming during the algorithm execution.

III. PROPOSED SYSTEM

In the proposed model we proposed a model based on the Density based on spatial clustering type of applications that includes noise method. In this model we use two input type parameters one is epsilon parameter value and other one is the minimum point to be given parameter value. Here we perform the normalization on the given dataset. The main reason for doing the normalization on the dataset is to change all the values of the column in the dataset to a common scale. There are various types of normalization are there and the mostly used types are Minmax Scaler and the Standard Scaler. By using these methods, we can translate the entire dataset into the normalized form. One important thing is each and every dataset does not need to be normalized and it is to be done only when there is huge difference between the values.

We are going to calculate the Epsilon parameter value by using the K nearest neighbour graph method. We need to choose how many nearest neighbours you want and calculate their distances and plot the graph by using the K nearest neighbour algorithm. Based on the graph we are going to choose the epsilon value. The place where there is drastic change in the graph or a valley is to be considered as an epsilon value. We should not choose the epsilon value to be very small because it will not consider most of the data and if we choose too large value the clusters will combine. We need to take the Minimum point greater than are equal to the three other wise it will look like the hierarchical or single linkage clustering.

A. Epsilon Calculation:

Input: Select the Dataset consisting of I values
Initialize I as a particular point in a dataset
Loop J = 1 to K
Calculate the distance between I & K
End Loop
Sort all the distances
Plot the graph
Output: Select the Epsilon value as where there is first drastic change or the valley point.

B. DBSCAN:

Input: Dataset, Newly formed Epsilon value, Minimum points, Cluster_Number to 0, Noise value
Initialize I as 1, Cluster has negative values of N length, N as length of the dataset
Loop I to N
If Cluster of I is negative
Call dbscan expansion
If Cluster is formed
Increase Epsilon value to 10 percent of original epsilon value
Else
Increase epsilon value by 50 percent of original epsilon value
End Loop

C. DBSCAN EXPANSION:

Loop I to N
Calculate distance between the point and I
If Distance (point, I)< = epsilon
I will be added to an array group K
End Loop
If K < minimum points value
Assign Noise value to entire K
Else
Loop I in K
Calculate directly reachable points from K group by distance function
End Loop

After calculating the epsilon parameter value, we need to focus on developing the dbscan algorithm. The newly formed algorithm is developed based on the existing dbscan algorithm. By using this procedure, we are going to perform the dbscan algorithm. In the first step we have seen that the dbscan will calculate the epsilon value based on the cluster formation. If the cluster is not going to form than we are going to increase the epsilon value by fifty percent of its original value. If there is a formation of cluster the epsilon value will be increased by ten percent of the existing value. In this way we can say that the density value is varying throughout the algorithm. In other words, the density value will be varying whenever there is a cluster is newly formed. Finally, we than calculate the silhouette score. A silhouette score can be defined as how well the points in the particular cluster are closely matched. Finally, at the end of the algorithm we than plot the cluster graph.

IV. RESULTS AND DISCUSSION

The newly formed algorithm was developed based on the existing dbscan algorithm. In this algorithm we are going to mainly focus on the epsilon parameter value. In this algorithm whenever a new cluster is formed the epsilon value will be going to increase by ten percent of its original value. If there is no cluster is formed the epsilon parameter value will be increased by fifty percent of its original epsilon value. This process will continue for each and every iteration until all the points in the dataset are covered. In a way we can say that we are performing this algorithm by varying densities. By using this algorithm, we mainly focus on silhouette score value. It represents how closely the point is similar to its cluster. We applied this algorithm on Many datasets. We focus mainly on three datasets. At first, we see the Facebook data where we applied this algorithm. Then we see the wi-fi indoor localization data set and the third one is Alcohol dataset. We downloaded these datasets from the UCI machine learning repository. At first, we calculated the epsilon value and plot the graph for these three datasets.

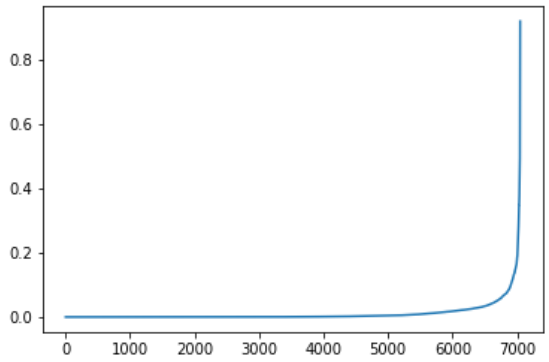

Fig. 1. Facebook Dataset

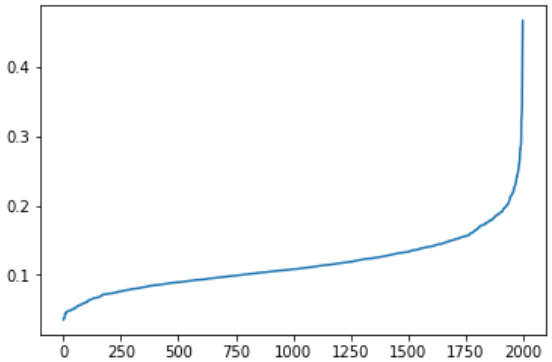

Fig. 2.Wi-fi signal strength Dataset

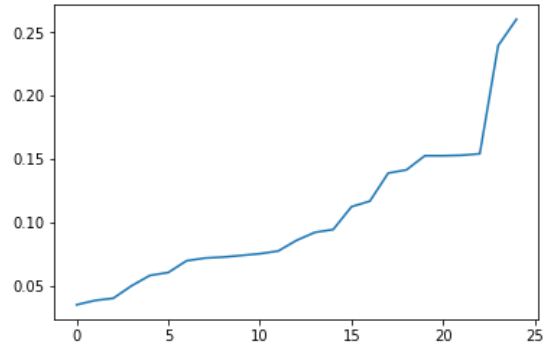

Fig. 3.Alcohol sensor Dataset

We select the epsilon where there is critical range value that come first are the valley point or where there is first drastic change that occur in the graph. After forming the epsilon value, we than perform the developed algorithm on the three datasets, where we earlier generated the K nearest neighbour graph for identifying the epsilon value. We consider these three datasets based on different in their sizes. We are performing the developed algorithm on various sizes of the dataset. The Facebook dataset consists of about 7000 instances and when we come to Wi-fi dataset it has about 2000 instances and finally the Alcohol sensor data it has about 25 instances. We apply the algorithm on these three datasets.

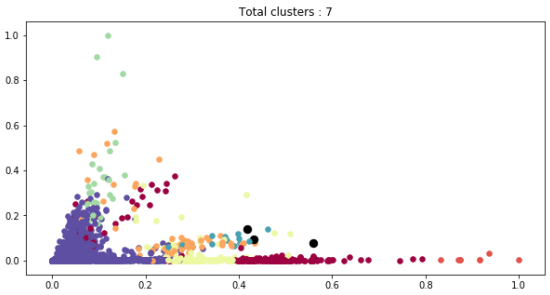

Fig. 4.Facebook dataset after plotting the cluster

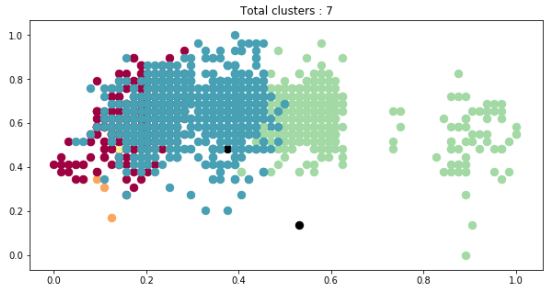

Fig.5. Wi-fi localization dataset after plotting the cluster

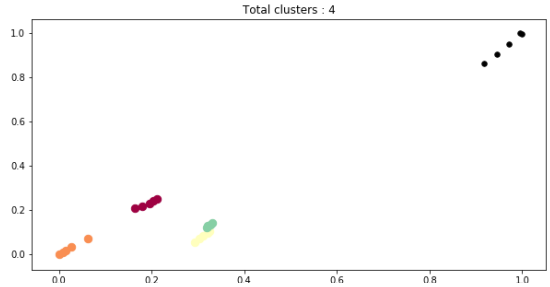

Fig. 6.Alcohol sensor dataset after plotting the cluster

These are the clustered graphs that are formed by applying the developed algorithm on these three datasets. The silhouette score refers to how closely or similar that the point belongs to a particular cluster. We compared the results of the silhouette score of the datasets that is achieved by applying the developed algorithm with the silhouette score of the datasets that is achieved by the existing dbscan algorithm. The developed algorithm silhouette score is most promising and providing good results compared to the existing dbscan algorithm.

V. CONCLUSTION & FUTURE WORK

Cluster algorithms are used for grouping up of similar points to form a cluster. The most popular density-based algorithm is DBSCAN. We developed a model based on the existing Dbscanalgorithm, whenever a cluster is formed the epsilon value will be incremented to ten percent to the existing epsilon value and if no cluster is formed then the epsilon value will be increased to fifty percent of the existing epsilon value. By doing this, we can say the we are performing the algorithm with varied densities. By applying this algorithm on the three different size of datasets we have achieved a better silhouette score from these three datasets than that is achieved by the existing dbscan algorithm. In the future, a new modified version of the dbscan to be developed to reduce the time complexity of the algorithm and need to achieve the parallel programming and use the thread concepts.

REFERENCES

1. "Data Mining Curriculum". ACM SIGKDD. 2006-04-30. Retrieved 2014-01-27.
2. Christopher (2010). "Encyclopædia Britannica: Definition of Data Mining". Retrieved 2010-12-09.
3. Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome (2009). "The Elements of Statistical Learning: Data Mining, Inference, and Prediction". Archived from the original on 2009-11-10.Retrieved 2012-08-07.
4. Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic (1996). "From Data Mining to Knowledge Discovery in Databases"
5. Osmar R. Zaiane, Introudction to data mining 1999.
6. The definition "without being explicitly programmed" is often attributed to Arthur Samuel, who coined the term "machine learning" in 1959, but the phrase is not found verbatim in this publication, and may be a paraphrase that appeared later. Confer "Paraphrasing Arthur Samuel (1959), the question is: How can computers learn to solve problems without being explicitly programmed?" in Koza, John R.; Bennett, Forrest H.; Andre, David; Keane, Martin A. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. Artificial Intelligence in Design '96.Springer, Dordrecht. pp. 151–170.
7. [7] The definition "without being explicitly programmed" is often attributed to Arthur Samuel, who coined the term "machine learning" in 1959, but the phrase is not found verbatim in this publication, and may be a paraphrase that appeared later. Confer "Paraphrasing Arthur Samuel (1959), the question is: How can computers learn to solve problems without being explicitly programmed?" in Koza, John R.; Bennett, Forrest H.; Andre, David; Keane, Martin A. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. Artificial Intelligence in Design '96.Springer, Dordrecht. pp. 151–170.
8. U.M. Fayyad, G. P. Shapiro, P. Smyth, R. Uthurusamy, Advances in Knowledge Discovery and Data Mining, MIT Press, Boston, MA, 1996.
9. S. Theodoridis, K. Koutroumbas, Pattern Recognition, second ed., Academic Press, New York, 2003
10. R.C. Gonzalez, R.E. Woods, Digital Image Processing, third ed., Pearson Prentice-Hall, Upper Saddle River, NJ, 2008.
11. S. Madeira and A. Oliveira Bi, clustering algorithms for biological data analysis: a survey, IEEE/ACM Trans. on Comp. Biology and Bioinformatics, 1(1) (2004) 24–45.
12. Mitchell T M, Machine Learning, McGraw Hill, 2013.
13. Rogers S, Girolami M, A First Course in Machine Learning, CRC Press, 2011.
14. Francis S. Hill (2001). Computer Graphics.Prentice Hall.
15. R.H. Gueting, An introduction to spatial database systems, The VLDB Journal 3(4) (1994)357-399
16. [16] Ester, Martin, Hans-Peter Kriegel, Jörg Sander, and XiaoweiXu. 1996. “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.” In, 226–31.AAAI Press.
17. A.Faroughi, R. Javidan and M. Emami, "A new density estimator based on nearest and farthest neighbor," 2016 8th International Symposium on Telecommunications (IST), Tehran, 2016, pp. 185-190.
18. JiaRongfei, JinMaozhong and Liu Chao, "A new clustering method for collaborative filtering," 2010 International Conference on Networking and Information Technology, Manila, 2010, pp. 488-492.
19. A. K. M. R. Chowdhury, M. E. Mollah and M. A. Rahman, "An efficient method for subjectively choosing parameter ‘k’ automatically in VDBSCAN (Varied Density Based Spatial Clustering of Applications with Noise) algorithm," 2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE), Singapore, 2010, pp. 38-41.
20. A. Smiti and Z. Elouedi, "DBSCAN-GM: An improved clustering method based on Gaussian Means and DBSCAN techniques," 2012 IEEE 16th International Conference on Intelligent Engineering Systems (INES), Lisbon, 2012, pp. 573-578.
21. C. Tsai and T. Huang, "QIDBSCAN: A Quick Density-Based Clustering Technique," 2012 International Symposium on Computer, Consumer and Control, Taichung, 2012, pp. 638-641.
22. D. Chen, Y. Yan and D. Wang, "Density clustering based on border-expanding," 2014 10th International Conference on Natural Computation (ICNC), Xiamen, 2014, pp. 670-674.
23. A. Ghanbarpour and B. Minaei, "EXDBSCAN: An extension of DBSCAN to detect clusters in multi-density datasets," 2014 Iranian Conference on Intelligent Systems (ICIS), Bam, 2014, pp. 1-5.

24. W. Wang, Y. Wu, C. Tang and M. Hor, "Adaptive density-based spatial clustering of applications with noise (DBSCAN) according to data," 2015 International Conference on Machine Learning and Cybernetics (ICMLC), Guangzhou, 2015, pp. 445-451
25. M. Debnath, P. K. Tripathi and R. Elmasri, "K-DBSCAN: Identifying Spatial Clusters with Differing Density Levels," 2015 International Workshop on Data Mining with Industrial Applications (DMIA), San Lorenzo, 2015, pp. 51-60.
26. A. Smiti and Z. Elouedi, "Fuzzy density based clustering method: Soft DBSCAN-GM," 2016 IEEE 8th International Conference on Intelligent Systems (IS), Sofia, 2016, pp. 443-448.
27. X. Wang, Y. Chen and X. L. Wang, "A New K-NN-Centroid-Inspired Density-Based Clustering Algorithm," 2017 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, 2017, pp. 766-771.
28. M. M. R. Khan, M. A. B. Siddique, R. B. Arif and M. R. Oishe, "ADBSCAN: Adaptive Density-Based Spatial Clustering of Applications with Noise for Identifying Clusters with Varying Densities," 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), Dhaka, Bangladesh, 2018, pp. 107-111.
29. J. Song, Y. Guo and B. Wang, "The Parameter Configuration Method of DBSCAN Clustering Algorithm," 2018 5th International Conference on Systems and Informatics (ICSAI), Nanjing, 2018, pp. 1062-1070.

AUTHORS PROFILE

Mr VikramNeerugatti, Assistant Professor,, Department of Computer Science and Engineering, Sri VenkateswaraCollege of Engineering, Chittoor.Mr VikramNeerugatti, is Working on area of Internet of Thing (IoT). He has Submitted his PhD in the area of Internet of Things (IoT) and his research Contributes are to provide detection and prevention mechanisms for RPL attacks like Sinkhole, Black hole, Wormhole, Rank attack, etc in IoT. He completed his Bachelors (B.Tech) & Master's (M.Tech) Degree in the Specialization of Computer Science and Engineering at JNTUA. He has M.S Degree from Brain wells University, London, UK. He has more than 10 years teaching experience from various Institutions like NIT, Goa, Sri Vidyanikhethan, A. Rangampeta, Sri Venkateswara University, Tirupati and Sri Venkateswara College of Engineering, Chittoor. He has more than 5 years of Research Experience form NIT Goa and S. V University, Tirupati. He has attended more than 50 Workshops specifically in the area of IoT. He has published 15+ National and International Conferences and Journals. He got 6 best research paper awards. Recently he got Dr. B.R. Ambedkar Research Fellowship award for his innovative research contributions. His research areas are Internet of Things, Fog Computing, Cloud Computing, etc. He acted as a Resource person for more than 20 workshops specifically in the domain of the IoT& Fog Computing in various institutions like Sri Venkateswara University, Sri PadmavathiMahilaViswavidyalam, Tirupati, Chdalavadavenkatasubbamma engineering college, tirupati, Sir Vidyanekithan Engineering College, etc.

Mr.MokkalaKiranMoni Reddy He Pursuing his M.Tech at SVU, Tirupati.He **r**eceived his B.Tech Degree from the University. His research interests are Software Engineering, Software Architecture, Cloud Computing, Operating System and Data Mining. He is life member of CSI. He has 2 publications and presented 3 papers in National and International conferences.

Dr.A.Rama Mohan Reddy He **r**eceived his B.Tech Degree from JNT University Anantapur in 1986, Masters in Computer Science and Engineering from NIT, Warangal in 1991 and Ph.D. in Computer Science and Engineering from Sri Venkateswara University, Tirupati in 2007. He is currently working as a Professor of Computer Science and Engineering, SV University College of Engineering, Tirupati, India. His research interests are Software Engineering, Software Architecture, Cloud Computing, Operating System and Data Mining. He is life member of ISTE, IETE, ISC and CSI. He has more than 30 years of experience in teaching, 7 scholars completed Phd.'s under his guidance. He has 100+ publications and presented 78+ papers in National and International conferences.