Clustered Comparative Analysis of
Security Sensor Discrimination Data

Prof. Neeraj Bhargava[1], Aakanksha Jain[2],
Abhishek Kumar[3], Pramod Singh Rathore[4], Anuja Bansal[5]
[1]School of system sc. &engg. Professor MDS University Ajmer, India
[2]M.Tech Scholar at MDS University Ajmer, India
[3]Assistant Professor, Computer Science Department ACERC, Ajmer
[4]Assistant Professor, Computer Science Department ACERC, Ajmer
[5]M.Tech Scholar at MDS University Ajmer, India
[1]profneerajbhargava@gmail.com, [2]Jain1994aakanksha@gmail.com,
[3]Abhishekkmr812@gmail.com, [4]pramodrathore88@gmail.com, [5]anujabansal07@gmail.com

Abstract—**Security alarm is used to protect from burglary (theft), property damage and from intruders. These security alarms consists sensors and alerting device to indicate the intrusion. Clustering is data mining technique which is used to analyzing the data. In this paper we discus about different clustering algorithm like DBSCAN, Farthest first. These algorithms are used to evaluate the different number of clusters with the sensor discrimination data base. In any organization Sensor security has many types of security alarm. It may be glass breaking alarm, smoke heat and carbon monoxide alarm, and it may be false alarm. Our aim is to compare the different algorithms with the sensors data to find density clusters i.e. which type of data will provide dense cluster of useful alarm condition. This evaluation will also detect the outliers within data such as empty alarms.**

Index Terms—**Acoustic; Nuisance; Detector; Farthest First Algorithm.**

I. Introduction

SECURITY sensor alarm is a device that gives us indication of abnormal situation within the system. It can sense the environment situation and provide the signal by alarming for protection. In Glass braking alarms, shock sensors is installed in glass pane. When glass is break a shock frequency generated through the glass, surrounding walls and through the window frame to send the frequency to device that generates the alarm of glass breaking.

Smoke heat and carbon monoxide alarm are not turn off for a second anytime. In this two metal plates are used to generate electric current. When a smoke is entering in chambers, it detects the small amount of smoke and sound the alarm for indicates the situation. In security sensor some time may be false alarm are causes due to extra precaution i.e. if small particle of smoke is detected by chamber alarm device then it sounds the false alarm.

DBSCAN is density based clustering [5][18] algorithm with noise. It discover cluster in different size and shapes. [16]DBSCAN is able to identify class in spatial database. It is more powerful clustering [1] algorithm then other algorithm to solving anomaly detection problem. DBSCAN is use in all fields because it determines outlier accurately. In DBSCAN clusters are dense regions separated by low dense regions.

Fa farthestFirst Algorithm is a Clustering approach which selects center groups by choosing data point on basis of their distance from each other. Algorithm works in similar way as K-means do as follows:

(i) All of first it selects the centroid.

(ii) And then it begins the cluster assignment [12].

Di differences of FarthestFirst with K-mean are:

(a) FartestFirst works in single pass only to make a set cluster of selected data points, need not to update regularly after each iteration.

(b) Centroid calculation is done with actual data not with geometric aspects.

Significant advantage of choosing FartestFirst over K-mean is it provides good performance in centroid selection because distance calculated between centroid is optimal distance. Here term centroid can be defined as a location of cluster centers in attribute set space. These centroids can move freely and are not supposed to bind in set of value of instances

II. Literature survey

(1). *Passive infrared detector*: PIR detector is commonly used which is us in household and small business environments. When it determines the heat energy, it started the work [6]. It does not generate its own energy and does not detect motion. It is designed according to wall or ceiling mounted.

(2). *Infra sound detector*: Infrasound detector works when it detects the sound waves at the frequency below 20 hertz. Infrasound is untreatable for human ear, its sound waves traveled along hundred kilometers. This infrasound alarm give indication of earthquake, gravity waves etc. infrasound sensor, an order frequency filter, an analog to digital converter, and for analyze the record signal MCU is consist in the infrasound detector to give us result.

(3). *Motion sensor*: Motion sensor are use to detect [7] the movement using various technologies. In motion sensor infrared, ultrasonic, vibration and contact technology is used to detect the movement. These detection technologies are used to reduce the false alarm.

(4). *Ultrasonic detector*: Ultrasonic alarm or detector emits the waves by transmitter at frequency between 15 kHz to 75 kHz. These waves is inaudible to human hear due to its high frequency, by using sound waves it determine the distance for an object. To measure the distance this device transmit the signals or sound waves, then sound waves are reflected by solid surface and detected by device.

(5). *Vibration sensor*: These devices are designed to detect an attack on structure itself. These devices are placed on barriers. When vibration is occurring, it breaks the current flow by which alarm is started [10]. Its sensor is reliable and there are low chances of generating false alarm.

III. RELATED WORK

DBSCAN Clustering: Density based [3] clustering algorithm measure the density of a region with spatial datasets. DBSCAN depends on two arguments that are epsilon and Minpts. Minpts argument is used to reduce the single-link effect and epsilon [2] distinguishes the radius of the search circle. DBSCAN [8] has the ability to find the complete cluster surrounding with another cluster but not connected with that cluster [13][14]. It uses only two parameter to ordering of the points in the database to recognize the arbitrary pattern, different size clusters and filters noise.

FarthestFirst Clustering: Farthest-first is use to speed up the clustering. But for fast performance it need less move and adjustment. It works like k-mean algorithm in some ways. It choosing data points on basis of their distance from each other to select a center point. In farthest-first algorithm each successive point is as far as possible from the set of previously-selected points [9].

IV. EXPERMIMANTAL EVALUATION

The dataset contains the possible values collected from differnet kind of sensor from any commercial organisation. Substanceroup is the class attributes which have values: (a) Group A which deals with all smoke, heat and Carbon monoOxide detection realed sensors data.

(b)Acoustic is category which deals with all alarms associated with glass breaking events.

(c) Nuisance category includes all False and absent alarm.

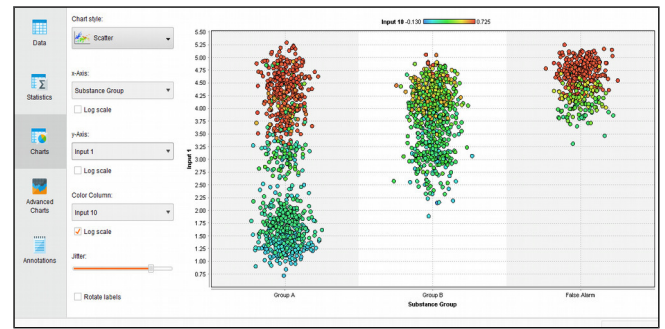Evaluation of attributes with class variable is as shown below:

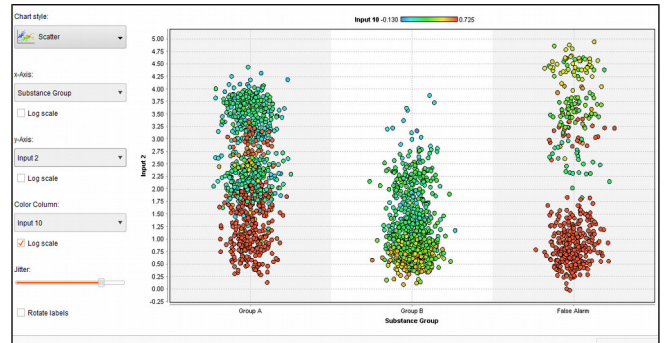Figure 1: Evaluation of category 1 data with class variable

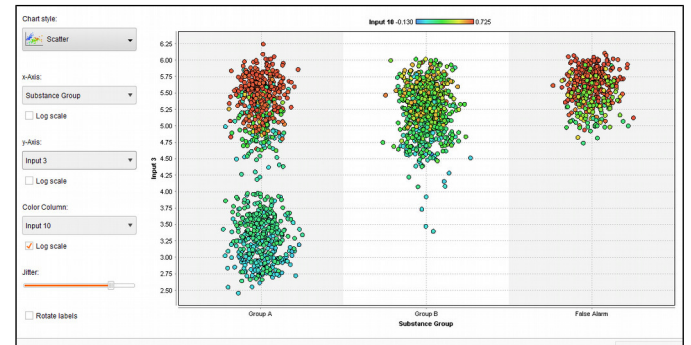Figure 2: Evaluation of category 2 data with class variable

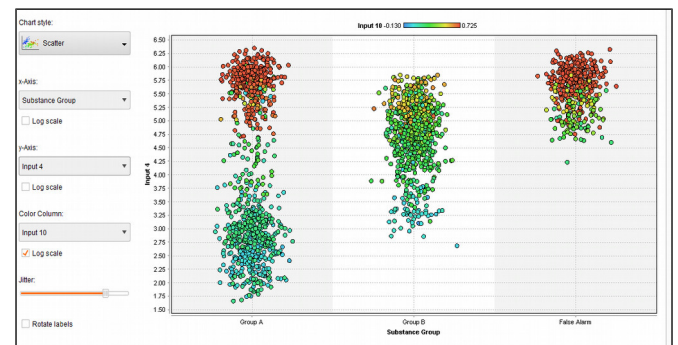Figure 3: Evaluation of category 3 data with class variable

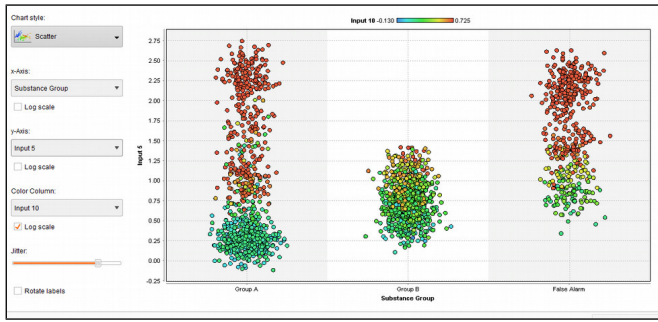Figure 4: Evaluation of category 4 data with class variable

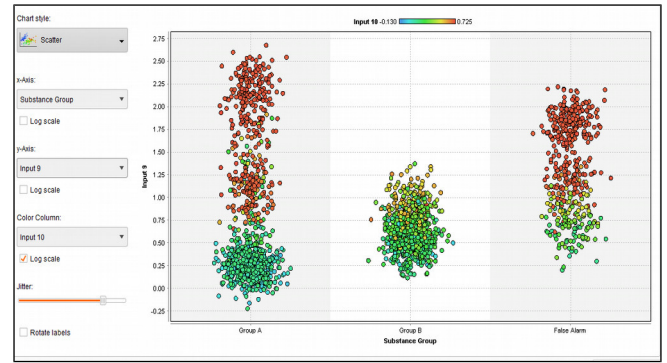Figure 5: Evaluation of category 5 data with class variable

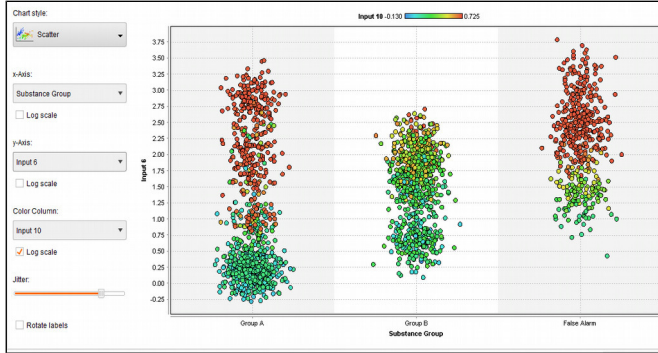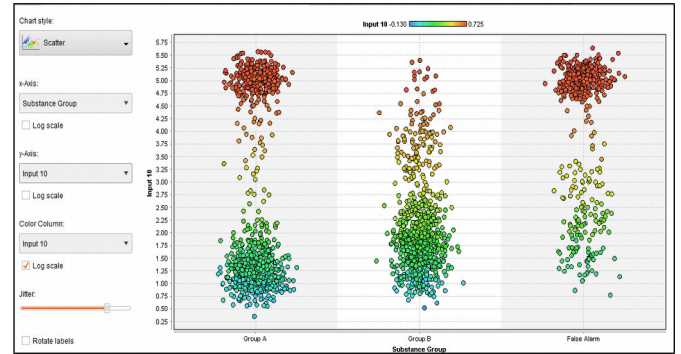Figure 6: Evaluation of category 6 data with class variable

Figure 7: Evaluation of category 5 data with class variable

Figure 8: Evaluation of category 6 data with class variable

Figure 9: Evaluation of category 9 data with class variable

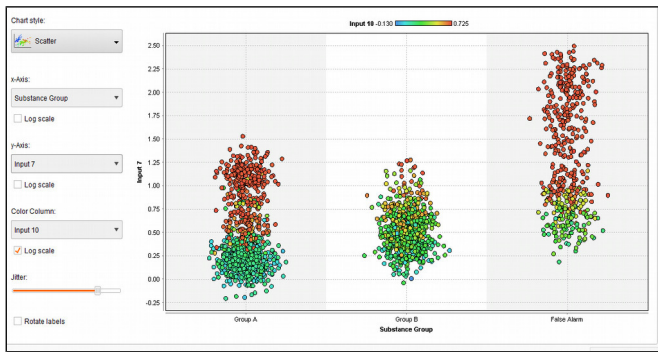Figure 10: Evaluation of category 10 data with class variable

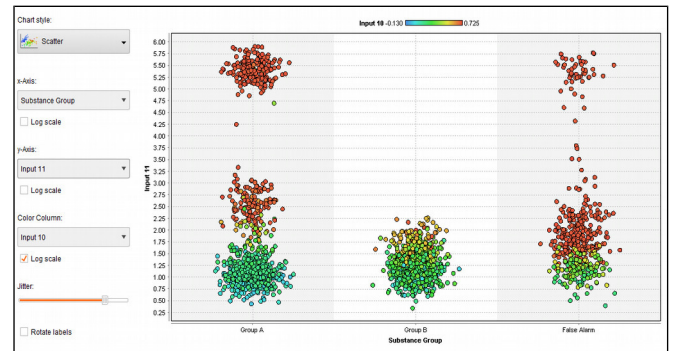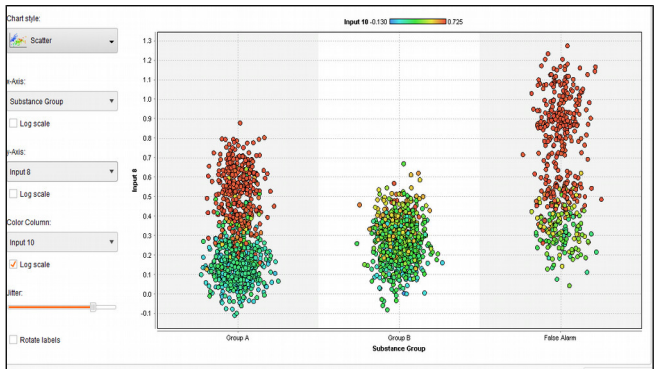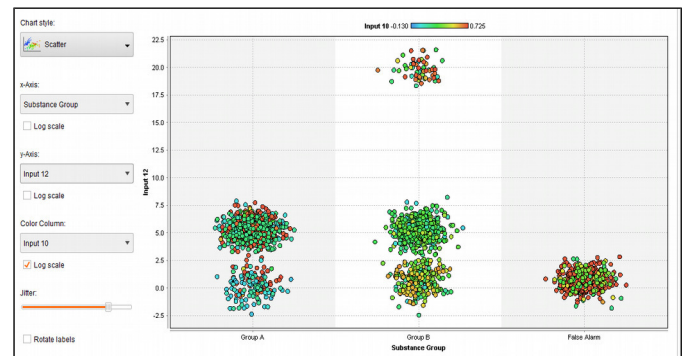Figure 11: Evaluation of category 11 data with class variable

Figure 12: Evaluation of category 12 data with class variable

(1). *Experimental evaluation using Dbscan Algorithm*: Dbscan algorithm makes clusters by iteratively [4] checking neighbor elements of each data points within dataset [11]. In case nearby elements are more than minPts, a new cluster formed with O as core object. In second phase objects which are density-reachable with core object get merged into it [14]. The process terminates when no new point can be added to any cluster. Lower density points are concluded as noise thus gets omitted.
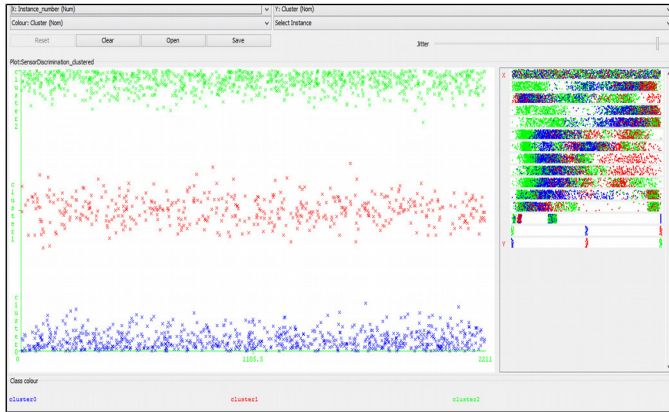

Figure 13: Experimental Evaluation using DBSCAN algorithm (Cluster assignment of each group)

(2) Experimental evaluation using FarthestFirst Algorithm: Farthest-First algorithm may be computed by greedy algorithm to maintain the distance between each point and previously selected points.

It works in following steps:-

(a) Firstly it initializes the empty sequence by sequencing of selected points. And also initialize the distances between each point and selected points to infinity.

(b) While all points have not been selected, the following steps are repeated:

- Find a point p by scanning the list of not-yet-selected points.
- Then remove p from the not-yet-selected point and add it to end of sequence of selected points.
- Now replace the distance stored for *q* by the minimum of its old value for each remaining not-yet-selected points.
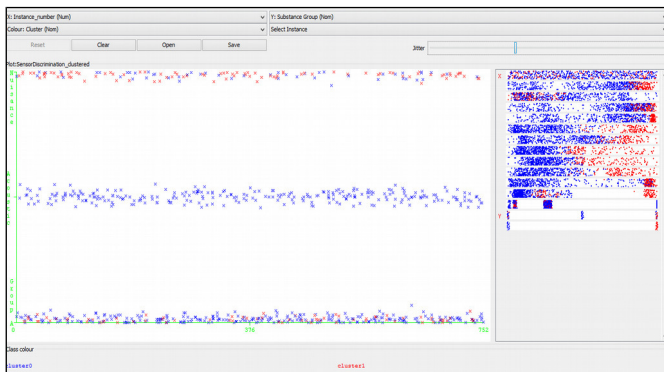

Figure 14: Experimental Evaluation using FarthestFirst algorithm (Cluster assignment in 2 groups)

V. Result Analysis

As experimental evaluation shows how clusters were made for each input category with class attribute, how dense or noisy data are there with respect to category selection in term of class attribute. Now result is analyzed by comparing generated clusters with Dbscan and farthestfirst. Result of Dbscan consist of 3 clusters C0,C1 and C2,each have 799,456 and 947 data points 36% ,21% and 43% of total dataset respectively ,whereas result of FarthestFirst consist of 2 clusters Cluster 0 and Cluster1 have 2955 and 257 data points 88% and 12% respectively of total provided dataset. Another factor is build time which is 3.13 seconds for Dbscan and FarthestFirst take no time in building model.
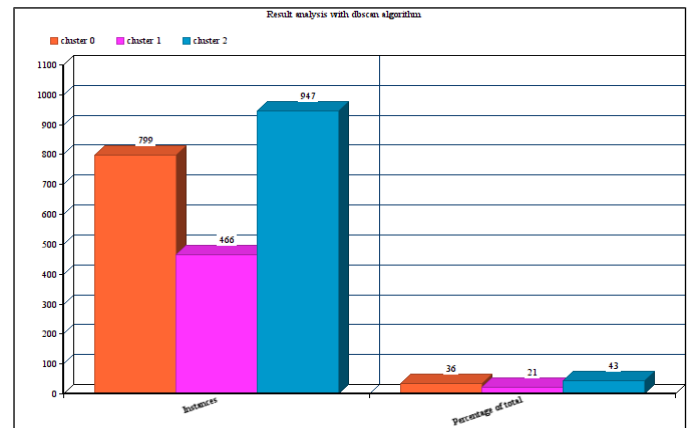

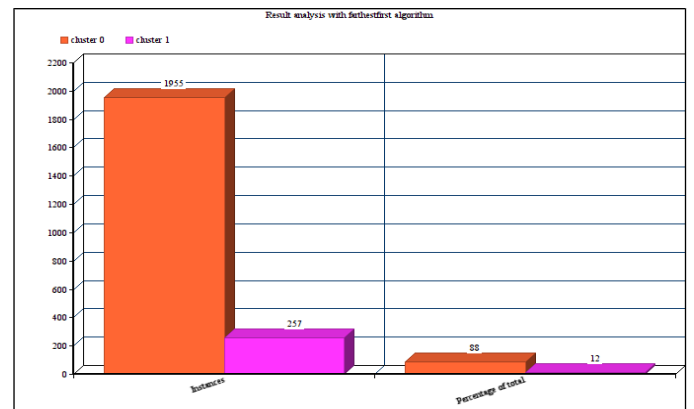. Figure 15: Result analysis with Dbscan algorithm

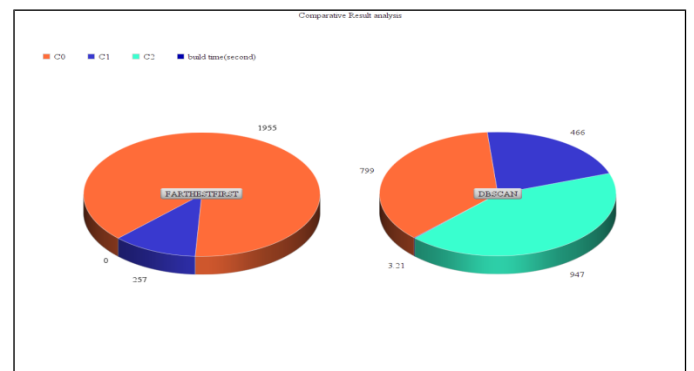Figure 16: Result analysis with FarthestFirst algorithm

Figure 17: Comparative analysis of Dbscan with FarthestFirst

VI. Conclusion and Future Work

Comparative analysis of results obtained after implementing Dbscan and FarthestFirst algorithm it is clear from the generated graph that results of Dbscan algorithm as compare to FarthestFirst are better in term of distinguishing the class variable clustering, three clusters generated will very clear assignments of data points into the corresponding cluster. Each set of input is also appraised with class variable, which clears the cluster assignment each data point of input group. We can easily analyze our desire set of the cluster and deport outliers. Both algorithms above are being analyzed using the trained set parameter based on its class attribute. For further future use this set of the clustering algorithm can be implemented in the large-sized organization to explore data detail obtained from sensors of that organization i.e we can categories type of alarm and on that basis harshness of alarm event can be decided and appropriate action can be quickly taken. Each group of input will contain each department's sensor data, which can individually get evaluated with a class variable i.e severity of sensor's alarm

References

[1] Ajin V W ,Lekshmy D Kumar "Big Data and Clustering Algorithms "International Conference on Research Advances in Integrated Navigation Systems (RAINS - 2016), April 06-07, 2016.

[2] Yuchao Zhang, Hongfu Liu, Bo Deng "Evolutionary Clustering with DBSCAN" 2013 Ninth International Conference on Natural Computation (ICNC) 978-1-4673-4714-3/13/$31.00 ©2013 IEEE.

[3] Bi-Ru Dai, I.-Chang Lin. Efficient Map/Reduce-Based DBSCAN Algorithm with Optimized Data Partition [C]. IEEE Cloud, 2012: 59-66.

[4] Tariq Ali, Sohail Asghar, Naseer Ahmed Sajid," Critical Analysis of DBSCAN Variations" 978-1-4244-8003-6/10$26.00©2010IEEE.

[5] S. R. Pande, S. S. Sambare and V. M Thakre, Data Clustering using data mining techniques, International Journal of advanced research in computer and communication engineering. 2012;142(4):87-91.

[6] M. Akhil Jabbers, Dr. Prirti Chandra and Dr. B. L Deekshatulu, Heart disease prediction system using associative classification and genetic algorithm, International conference on emerging trends in electrical, Electronics and communication technologies. 2012; 25(1): 78-90

[7] P. Venkateswara Rao and A.Ramamohan Reddy, Shrimp disease detection using Back Propagation Neural networks, International Journal of Pharma and BioSciences. 2016; 7:3.

[8] A. O. Quintana, "Clustering: Density-based Clustering", Universidad San Pablo, p. 4, July 2014 [Online]. Available: http://biolab.uspceu. com/clustering/Density-based.pdf.

[9] Liang Sun and Shinichi Yoshida, A Novel Support Vector and K-Means based Hybrid Clustering Algorithm, Proceedings of the IEEE International Conference on Information and Automation. 2010; 64(4):20-23.

[10] Yunseop (James) Kim, Member, IEEE, Robert G. Evans, and William M. Iversen "Remote Sensing and Control of an Irrigation System Using a Distributed Wireless Sensor Network" IEEE Transactions on Instrumentation and Measurement, Vol. 57, No. 7, July 2008

[11] Dr.A.V.Senthil Kumar "Heart Disease Prediction Using Data Mining preprocessing and Hierarchical Clustering" International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), Vol. 4, No.6 Pages: 07 - 18 (2015)

[12] A. O. Quintana, "Clustering: Density-based Clustering", Universidad San Pablo, p. 4, July 2014 [Online]. Available: http://biolab.uspceu. com/clustering/Density-based.pdf.

[13] P. Shrivastava and H. Gupta, "A Review of Density-Based Clustering in Spatial Data", International Journal of Advanced Computer Research, vol. 2, no. 3, 2012.

[14] A. Smith, V. Sucharita, P. Sowjanya and B. Geetha Krishna "A Predictive Model For Heart Disease Using Clustering Techniques" International Journal of Pharma and Bio Sciences, ISSN 0975-6299, Int J Pharm Bio Sci 2017 July; 8(3): (B) 529 -534

[15] Rajkumar, A. and G. S. Reena, " DiagnosisOf Heart Disease Using DataminingAlgorithm." Global Journal of computer science and Technology, 2010. Vol. 10(Issue 10).

[16] Varun Kumar, Nisha Rather,"Knowledge Discovery from Database using an Integration of clustering and Classification", IJACSA, vol 2 No. 3, PP. 29-33, March 2011.

[17] Ritu Chauhan, Harleen Kaur, M.Afshar Alam, "Data Clustering Method for Discovering Clusters in Spatial Cancer Databases", International Journal of Computer Applications (0975 – 8887) Volume 10– No.6, November 2010.