A Survey on K-means clustering algorithm for initialisation of centroid

[1]Swati Nenava [2]Prof. Manoj Chouhan

Abstract:
K-means is one of the simplest unsupervised learning and partitional clustering algorithms. This algorithm classifies a given data set by finding a certain number of clusters (K). The clusters are differentiated by their centers. The best selection is to place them possibly far away. K-means is a most widely used approach in unsupervised machine learning algorithms though it was proposed 50 years ago and is very fast the algorithm is highly sensitive to initial placement of the cluster centers.The survey examined KDD process with care in order to draw facts leading to a enhanced K-means algorithm to give overwhelming results.This study explore various clustering algorithms/techniques, traditional k-means and on the same some recent work.The review has revealed stupendous ideas which can be recommended.

General Terms: **Algorithms**
Additional Keywords: **Unsupervised Learning, K-means ,Data mining,Clustering, Initialisation methods,data set discovery.**

I. INTRODUCTION

Clustering works in a way that objects on the same group are similar to each other as compared to the other group.This group is commonly called a cluster. It is a main task to explore KDD, and a common technique analysis of statistical data used in many fields,such a learning techniques called as machine learning,pattern recognition and information retrieval.The goal of data clustering, also known as cluster analysis, is to find the instinct grouping(s) of a set of patterns, points, or objects.Overall aim is to develop an algorithm which is automatic that will discover the natural groupings in the unlabeled data.

A Google Scholar search found 1,660 entries with the words *data clustering* ,appeared in 2007 alone. This immense literature display enormous importance of clustering in data analysis.Though listing the numerous scientific fields and applications that have utilized clustering techniques as well as the thousands of published algorithms is difficult to handle.An Image segmentation is an significant clustering problem,the most-outstanding technique is Documents clustering.Clustering have various usage from grouping customers into different types for efficient marketing to grouping services delivery engagements for management and planning also to study genome data in biology.With the consequences that the basic problems and method of clustering become well known in a broad scientific community,in statistics,data analysis and in particularly in applications.

One of the major clustering approaches is stand on the sum of squares criterion and on the algorithm known as'k-means'.while tacking the record back this algorithm to its origin developments,we see that it has been proposed by several scientists in different forms and under different assumptions.K-means clustering aim is to divide or partition *n*o. Of data set points into *k* clusters on the basis of variance calculate.Clustering algorithms can be categorize into two groups: hierarchical and partitional [6]. Hierarchical algorithms iteratively find clusters in nested form either in a top-down or bottom-up fashion.In contrary,hierarchical structure does not impose when partitional algorithms detect all the clusters moreover as a partition of the data ,higher complexity in data points can be seen in hierarchical algorithms and thus are not suitable for large data sets,comparatively to partitional algorithms since they have lower Complexity.

II. DATA MINING

According to "Knowledge Discovery in Databases" process data mining is a subfield of computer science which explore database and conduct some meaningful and fruitful information from it,imply as a big data.Data mining is a huge workspace for the analysis of large volumes of data which integrates techniques from several research fields such as machine learning,statistics,pattern recognition, artificial intelligence, and database systems.Large number of data mining algorithms implanted in these fields to accomplish different data analysis tasks. [8] Data mining make a benchmark in education field and lead the researchers to categorize the diversity of research in data mining .Educational data mining methods have had some level of impact on education and related interdisciplinary fields (such as artificial intelligence in education,intelligent

Swati Nenava *IT Department,Shri Vaishnav Institute of Technology and Science,Indore,India.,07314240098*
Manoj Chauhan,IT Department,Shri Vaishnav Institute of Technology and Science,Indore,India,7509655233

ISSN: 2278 – 1323

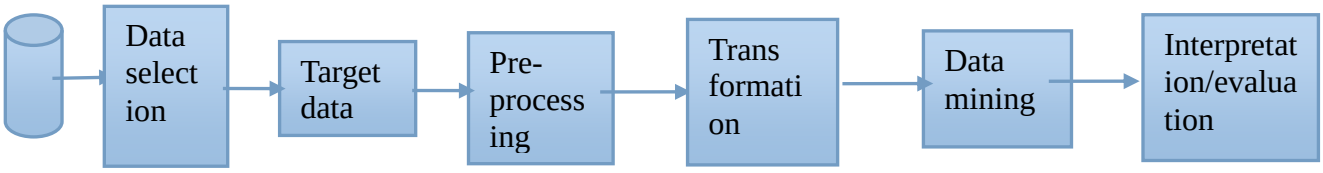

Fig 2(a) Flow chart of data mining process.

grievous computational needs tutoring systems, and user modeling[7].

A. *Application and challenges of data mining*

According to several offended the current state of the art of data mining research is too "ad-hoc." Many techniques are delineate for individual problems, such as classification or clustering,no consolidated theory have been seen.Moreover, a theoretical framework that unite different data mining tasks including clustering, classification, association rules, etc also different data mining approaches (such as statistics, machine learning, database systems, etc.), would help the field and provide a direction for future research.[11]in the late 1980s, data mining attain glorious victory. Several new hurdles have evolve and have been solved by data mining researchers.Moreover,there is still a lack of timely exchange of significant matters in the community,finally the 10 problems can be reviewed:

• Developing a consolidated theory of data mining
• Requirement of High dimensional data scaling and high speed data streams
• Need of Mining the sequence and time series data
• Stand in need of mining complicated knowledge from complicated data
• Data mining in a network configuration and distributed environment
• Various other usage like in biological and environmental problems
• Data mining in security, privacy and data integrity
• Handling non-static, unstable and cost-sensitive data

III. CLUSTERING ALGORITHMS AND TECHNIQUES

The top 10 data mining algorithms identified by the IEEE International Conference on Data Mining that is *k*-Means, SVM, Apriori, EM, PageRank, AdaBoost, *k*NN, Naive Bayes, and CART. These top 10 algorithms are the most influential data mining algorithms and make a difference among various other research communities These 10 algorithms include classification clustering, statistical learning, association analysis, and link mining, which plays the most important topics in data mining research and development.[8]

Classification of clustering algorithms is neither strewed, nor approved. In reality, different classes of algorithms overlap can be seen.Traditionally clustering techniques are categorized in hierarchical and partitioning. Hierarchical clustering is divided into agglomerative and divisive. The basics of hierarchical clustering include classic algorithms SLINK, COBWEB, as well as some other newer algorithms CURE and CHAMELEON [12]in other case hierarchical algorithms moderately dismantle points into clusters,whereas partitioning algorithms grasp clusters directly. Partitioning Relocation Clustering are categorize into probabilistic clustering such as algorithms called as SNOB, AUTOCLASS, MCLUST, secondly k-medoids methods (algorithms PAM, CLARA, CLARANS, and its extension), and lastly k-means methods (different schemes, initialization, optimization, harmonic means, extensions). Such methods aims on building clusters of proper convex shapes by noticing how well points fit into their clusters.

Density-based connectivity used in algorithms such as DBSCAN, OPTICS, DBCLASD, which are called partitioning algorithms,among density-Based Partitioning ,DENCLUE exploits space density functions. These algorithms are less fragile to outliers and can show clusters of irregular shape.They work with low-dimensional numerical data, known as spatial data. From the algorithm GDBSCAN it is found that spatial objects not only include points, but also geometrically extend the objects.

Most of the algorithms work with data indirectly by assembles summaries of data over the attribute space subsets called as Grid-Based Methods. In many other algorithms,it is found that grid-based methodology used as an intermediate step (for example, CLIQUE, MAFIA).

A. *Properties of clustering algorithms*

The characteristic of algorithm in data mining is an substantial factor in examining the performance which includes:

• Algorithm can handle the type of attributes
• Scalability to large datasets
• Power to act with high dimensional data
Power to find clusters of irregular shape and handling outliers
• Time complexity and Data order dependency
• Labeling or assignment (hard /strict vs. soft / fuzzy)
• Confidence on knowledge and user defined parameters
• Interpret-ability of results

The most conventional and the simplest partitional algorithm is

ISSN: 2278 – 1323

K-means.Ease of implementation, simplicity, efficiency, and empirical success are the main cause for its popularity. K-means details are abridge below.

IV. STANDARD K-MEANS

k-means clustering represent each cluster Ci through centroid $\boldsymbol{\mu}_i$ and it tends to quantify resemblance in terms of Euclidean distances between centroids and vector points. If reachability to a common centroid are small, two data points are considered similar on the same if its distance to μ_i is less than that to any other centroid a point **x**j will be assigned to cluster Ci [14]The aim of K-means is to minimize the sum of the squared error(SSE) over all the K clusters .K-means algorithm locate a partition such that the squared error between the empirical mean of a cluster and the points in the cluster is Minimized,here is the objective function:

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2$$

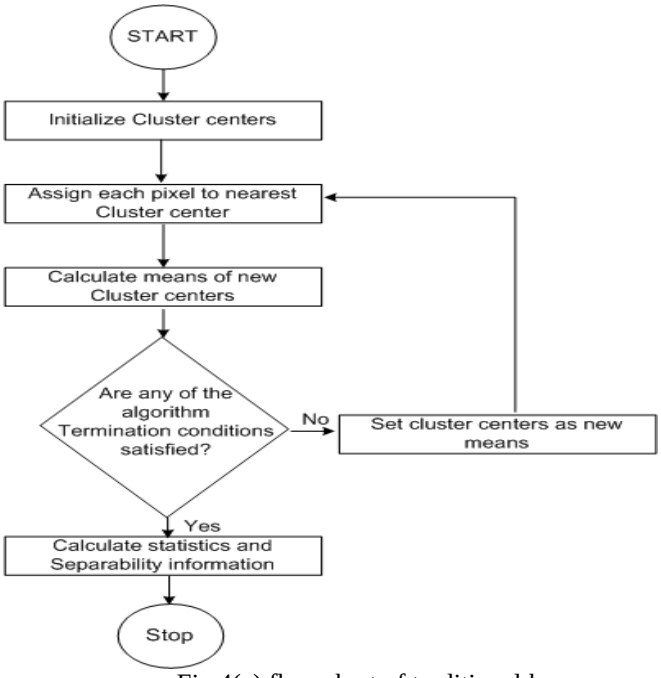

Fig 4(a) flow chart of traditional k-means

Pseudo Code

1.Begin with n clusters, each containing one object and numbering the clusters 1 through n.

2.Estimate the between-cluster distance D(r, s) as the between-object distance of the two objects in r and s respectively, r, s =1, 2, ..., n. Let the square matrix D = (D(r, s)).

3.Next, finding most similar pair of clusters r and s, such that the distance, D(r, s), is minimum in the middle of all the pairwise reachability.

4.Merge r and s to a new cluster t and compute the between-cluster distance D(t, k) for any existing cluster $k \neq r, s$. Once the reachability found, delete the rows and columns corresponding to the old cluster r and s in the D matrix, because r and s do not exist. After that add a new row and column in D corresponding to cluster t.

5.Repeat Step 3 a total of $n - 1$ times until there is only one cluster left.

A. *Problems overview*

K-means algorithm requires three user-specified parameters: number of clusters *K*, cluster initialization, and distance metric. The most critical choice is K. While no mathematical criterion exists, a number of heuristics are available for choosing K. Usually, K-means run separately for different values of K and the partition that appears the most significant to the domain expert is chosen. It is surveyed that different initializations can lead to different last clustering since it is well known that K-means only converges to local minima.[6]

B. *Facts drawn from research about k means*

[14]examined basic ideas behind k-means clustering and pinpoint common hazard in its use and detecting latent structures or evenness within a given sample and display the significance of preprocessing role in data,explaining the fact to decide how a set of data feature structures data points and locate for them.Its aim is on the difficulty of clustering data that can be entrenched in Euclidean vector spaces and show how classical k-mean minimize the objective function;with common hazard in practical applications.A[12] brief review of different clustering techniques in data mining prominence is on clustering in data mining and discuss about probabilistic clustering ,k-medoids and k-means ,showed that these methods aim on how neatly and clearly points are fitted in clusters and build clusters of convex shapes.[6] gives overview of clustering,give a run down of popular clustering methods, not only discusses the major challenges and key issues in designing clustering algorithms,but also figure out some of the emerging and useful research directions, including semi-supervised clustering, ensemble clustering, simultaneous feature selection, data clustering and large scale data clustering.

Facts

a) One way to overcome the local minima is to run the K-means algorithm, for a given K, with several different initial partitions and it is advisable to choose the partition with the smallest value of the squared error.

b) K-means run multiple times say n times with different values of the number of clusters *K*. The similarity between a set of points is defined as the number of times the two points co-exist in the same cluster in N runs of K-means. The last clustering is acquire by clustering the data based on the new pair-wise similarity.

c) The algorithm operates on certain implicit presumption which can superficially lead to unreasonable results.

d) K-means algorithm is a heuristic for tackling with a hard problem which is surprisingly in nature, and the result may

ISSN: 2278 – 1323

entirely depend on the initial clusters.
e) In K-means algorithm initial centroids are formed by random initialization method,and thus this is the reason it cannot be assure of converges at global minimum of its objective function. As it is distinguishing converges to a local minimum, it is crucial to run k-means several times to determine good clusters.
f) It is found that Proper preprocessing raises the probability for the k-means algorithm to recognize structures that are not necessarily Gaussian for very high dimensional data, after dimensionality reduction only k-means clustering should be applied only.
g) If k-means clustering is tried on data sets with no compact convex subsets, it is overly buoyant to anticipate the algorithm to sensible clusters.

V. REVIEW OF RECENT PREVIOUS RELATED WORK

M. Emre Celebi et.al [1]present an overview of initialisation methods with an prominence on their computational efficiency and differentiate eight commonly used linear time complexity initialization methods and vast accumulation of data sets using various performance criteria, analyzing the experimental results of using non-parametric statistical tests and also provide recommendations for practitioners by demonstrating popular initialization methods.They also proved that traditional initialisation methods often perform poorly and in fact there are strong substitutes to these methods.As preprocessing steps is the one of the important step in finding good clusters by identifying structures,attribute normalization is highly preferable preprocessing step which can give overwhelming results.

Fan Cai et.al[2] evaluate different clustering algorithms for analyzing different financial datasets varied from time series to transactions and discuss the pros and cons of each method to raises the understanding of inner structure of financial datasets and other large size database as well as the capability of each clustering method in this context**.**

Ahamed Shafeeq et.al [3**]**presents a enhanced Kmeans algorithm with the intension of not only improving cluster quality but also fix the optimal number of cluster. The K-means algorithm number of clusters need to be fixed. But practically,it is very difficult to affix the number of clusters.They not only work for known number of clusters in advance as well as unknown number of clusters and calculate the number of clusters on the run based on the cluster quality output.This process work on the basis of examining inter-intra clusters only. The input from the user is taken through algorithm and the user tell whether the number of clusters is fixed or not.On finding number of clusters to be fixed then it act same as K-means algorithm. For example, number of clusters is not fixed then it is mandatory to user to give least possible number of clusters as an input. The traditional algorithm procedure repeated by expanding the number of clusters by one in each iteration til it reaches the cluster rank level validity threshold.

Madhu Yedla et.al[4] shows a new method is proposed for finding the better initial centroids and to give an effective way of assigning the data points to the clusters which are suitable with reduced time complexity.

anil k jain[6**]**explain overview of clustering,explained well known clustering methods,and discusses the major dispute also key issues in designing clustering algorithms, and pinpoint about various methodologies,such as ensemble clustering, simultaneous feature selection, data clustering,semi-supervised clustering and large scale data clustering

Fahim A.M. Et.al[10]benefit from previous iteration of *k*-means algorithm, For each data point, they can keep the distance to the nearest cluster. At the next iteration, and calculate the reachability to the previous nearest cluster.Through this the time required to compute distances to $k-1$ cluster centers is saved.

Facts:
[1]With the survey of IMs it is found that IMs have time complexity linear in N reason to be that k-means itself has linear complexity, which is perhaps the most significant reason for its popularity. Therefore, an IM for k-means should not decline this gives advantage of the algorithm. Eight commonly used, order-invariant IMs were involve in the experiments: Forgy's method (F), MacQueen's second method (M), maximin (X), Bradley and Fayyad's method (B) with J = 10, k-means++ (K), greedy k-means++ (G), Var-Part (V), and PCA-Part (P).Here it is noted that in between these methods only V and P are deterministic. Deterministic methods, i.e., V and P, generally outperform the non-deterministic ones,this fact evolve because non-deterministic methods can produce highly variable results during multiple runs. In time-critical applications when considering large dataset where determinism is important ,methods V or P is preferred. These methods have a very good impact in a way that executed only once and resulting in very fast k-means converge. The efficiency difference between the two can be notice only on high dimensional data sets. This is because method V determine the direction of split by determining the coordinate axis with the greatest variance (in O(D) time), whereas method P attain this by calculating the principal eigenvector of the covariance matrix.It is noted that though its higher computational complexity, method P can, in some cases, be more efficient than method V .Facts shows that former method converges significantly faster than the latter. But due to their hierarchical formulation these methods are hard to implement.

In clustering tasks, attribute normalization is a prominent step in preprocessing that intercept attributes with large ranges from dominating the distance reachability calculations and also to avoid numerical instabilities in the computations. Normalization methods are two commonly used the first is (min-max normalization) and second is (z-score normalization).Studies shows that the first method is preferable to the second since the second is likely to detach valuable data points in between-cluster variation [15, 16]. [2] show that density-based clustering does not suit financial dataset. Normalised centroid-based clustering with higher DI or lower DBI gives the best number of clusters to help understanding financial data classification. Original attribute scales do not reflect the behaviour similarity since Euclidean

ISSN: 2278 – 1323

distance is controlled by large scaled attributes, best average tightness does not indicate the best case by departing the worst case.Some constrains are found e.g., K-means clustering tends to find spherical clusters, centroid-based clustering does not handle the noise, etc. [3**]** The algorithm works well for the unknown data set with better results than K-means clustering.On fixing number of cluster to be very small then there is a chance of putting dissimilar objects into same group and for example, the number of fixed cluster is large, then there is more possibility that more similar objects will be placed into different groups.The main hitch of the proposed method is more computational time than the K-means for vast data sets.

[4] shows the k-means algorithm is one of the often used clustering method in data mining, its because of the staging in clustering huge data sets. The result of the final cluster in k-means algorithm substantially rely on the accuracy of the initial centroids,which are selected randomly.More importantly original k-means algorithm converges to local minimum, not the global optimum.[6] concluded the summons in data clustering is to assimilate domain knowledge in the algorithm, finding meaningful representation and finding similarity measure,calculating accuracy of cluster,planning a rational basis for comparing methods,and develop efficient algorithms for clustering large datasets.[10**]**use the idea which comes from the fact that the *k*-means algorithm discovers spherical shaped cluster, whose center is the gravity center of points in that cluster, this center moves as new points are added to or removed from it. This concept makes the center closer to some points and distant apart from the other points, whatever points get closer to the center it will stay in that particular cluster,making no necessity to find its distances to other cluster centers. The points far apart from the center may change the cluster, so only for these points their reachability to other cluster centers are determined, and reallocate to the nearest center

VI. DISCUSSION/COMPARISONOF TECHNOLOGY AND PROGRESS

The above survey underscores one of the important facts about clustering; *there is no best clustering algorithm*. Each clustering algorithm imposes a structure on the data either explicitly or implicitly. When there is a better match between the model and the data, overwhelming partitions are obtained.In the search for good models,it is advisable to include all the available information, no matter whether it is unlabeled data, data with constraints, or labeled data,this can help in using the right dataset solving the contradiction whether to take labeled or unlabeled data.Intelligent initializations are the important term for future research and arbitrary initialization are considered unfavorable; initial centroids should at least be selected among the available data points in k-means clustering.

VII. PROPOSED IDEA ON THE BASIS OF SURVEYING:

Clustering is an interesting domain of separating points geometrically to classify the objects in a given solution space.In k-means which is one of the popular partitioning clustering algorithn faces in handling,empty clusters,outliers detection,and reducing SSE(sum of squared error).The past processing is required when a strategy is needed to choose a substitute centroid,in other case the squared error will be very larger as expected.The resulting cluster centroids (prototypes) may not be as typical and when there efficient initialisaton method does not exist ,this happens because the result of k-mean is entirely depends on the initial centroids,which need to me handle with care.Lots of strategies has come to be outcome from this survey as follow:

In order to diminishing the total SSE by increasing the number of clusters includes, splitting a cluster with the largest SSE to be chosen, but we could also split the cluster with the largest standard deviation for one particular attribute.Secondly,introduce a new cluster centroid: Often the point that is farthest from any cluster center is chosen.

In order to decrease the number of clusters by trying to minimize the increase in total SSE, are the following:

Strategy 1: Disperse a cluster: This is achieve by removing the centroid those cor-responds to the cluster and reassigning the points to other clusters.

Strategy 2: Merge two clusters: Merge the two clusters that result in the smallest increase in total SSE.

Though this ways are not as such effective in decreasing total SSE,but instead is the use of initialization methods,the best initialization found to be PCA-PART and VAR-PART,among them PCA-PART give overwhelming results comparatively.

A.Key properties to be achieved with a good quality of clusters:

1.When the intra-class (that is, intra-cluster) similarity is high and inter-class similarity is low.

2.When clustering method have high expertise in measuring of its ability to discover hidden patterns.

B. QoS parameters as time domain properties which is also required to achieve:

1.Computing reachability between the two instances is O (m) where m is the dimensionality of the vectors.

2.Reallocated clusters: O(kn) distance computations, or O(knm).

3. Quantify centroids: Each instance vector gets added once to same centroid: O (NM).

4.Assume these two steps are each done once for I iterations: O(Iknm).

*C. Pros and Cons of Clustering:*This includes features which give incitement in order to achieve best results from traditional clustering method for enactment different kind of application.

ISSN: 2278 – 1323

Pros:
Clustering approach is simple to perceive and implement,in a way that data patterns are scattered in a space by which patterns becomes closer to recognize.

1. Items automatically assigned to clusters

Cons:
1. Number of clusters need to be known before hand.
2.It usually cease at a local optimum.
3.Sometimes fruitless items used in data-set is also makes inefficient clusters.
4.Too sensitive to outlier.

In order to make traditional clustering approach to be better,this cons need to be removed,which can be dramatically,achieve by improving preprocessing of data,through the survey it is found that by normalizing the data vector and making the better sufficient data set for further process every point need to be included in some applications. Therefore in proposed work,normalized values is found for having good data and no need to traverse whole data for calculating reach-ability between points By doing attribute normalization not only decrease memory usage but also time usage too.Since through normalization anomaly time series is also detected .But in order to get prodigious results its required to apply transformation to normalized values and not overestimate the importance of the normality assumption but rather removing normality of the errors or residuals is favorable ,since the purpose of transformation is about stabilizing the variance,which has to do with the errors (e.g., residuals).We are going to implement a new way concept of three shield protection for handling dataset and giving efficient as well as fruitful input before applying algorithm and thus reducing SSE trivially.

VIII. CONCLUSION

K-means found to be in the top 10 data mining algorithms identified by the IEEE International Conference on Data Mining.Despite its drawbacks, k-means remains the most widely used partitional clustering algorithm in practice. The algorithm is simple, easily understandable and reasonably scalable, and can be easily modified to deal with streaming data.In order to deal large datasets,it is required to organize data in sensible groupings which arises naturally in various scientific fields.Because of this reason it is not surprising to notice the increasing vogue of data clustering. It is significant to remind that cluster analysis is an exploratory tool; the output of clustering algorithms only suggest hypotheses.On the contrary,thousands of enhanced clustering algorithms have been seen and new ones continue came into exist.To find the best algorithm is again the part of research but among the research developments ,most algorithms including the popular K-means, are justifiable algorithms. The *k*-means algorithm and its neural implementation, the Kohonen net, are most successfully used on large data sets. This is because *k*-means algorithm is simple to implement and computationally attractive because of its linear time complexity substantial effort,in seeing all the consequences and possibilities of enhancing k-means method in efficient way,we are looking forward to proposed efficient and powerful procedure in rescuing from errors or residuals by the proposed new way concept of three shield protection mechanism.

IX. REFERENCES

[1] M. Emre Celebi,Hassan A. Kingravi,Patricio A. Vela ,Expert Systems with Applications, 40(1): 200–210, 2013

[2] Fan Cai, Nhien-An Le-Khac, M-Tahar Kechadi, Clustering Approaches for financial data analysis: a survey,2012

[3] Ahamed Shafeeq B M and Hareesha K S,Dynamic Clustering of Data with Modified K-Means Algorithm *IPCSIT vol. 27 (2012) © (2012) IACSIT Press, Singapore*

[4] Madhu Yedla, Srinivasa Rao Pathakota T M Srinivasa,Enhancing K-means Clustering Algorithm with
Improved Initial Center(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 1 (2) , 2010, 121-125

[5] https://scholar.google.co.in/

[6] Anil K. Jain, Data Clustering: 50 Years Beyond K-Means,2010

[7] Ryan S.J.D. Baker, Kalina YACEF,The State of Educational Data Mining in 2009: A
Review and Future visions

[8] XindongWu · Vipin Kumar · J. Ross Quinlan · Joydeep Ghosh · Qiang Yang ·Hiroshi Motoda · Geoffrey J. McLachlan · Angus Ng · Bing Liu · Philip S. Yu ·Zhi-Hua Zhou · Michael Steinbach · David J. Hand · Dan Steinberg, Top 10 algorithms in data mining,Knowl Inf Syst (2008) 14:1–37

[9] T. Su, J. G. Dy, In Search of Deterministic Methods for Initializing K-Means and Gaussian Mixture Clustering,

Intelligent Data Analysis 11 (4) (2007) 319–338.

[10] Fahim A.M, Salem A.M, Torkey F.A, Ramadan M.A, An efficient enhanced *k*-means clustering algorithm,1626 *al. / J Zhejiang Univ SCIENCE A 2006 7(10):1626-1633*

[11] Qiang Yang,Xindong Wu,10 Challenging Problems In Data Mining Research,Vol. 5, No. 4 (2006) 597–604

[12]Pavel Berkhin,A Survey of Clustering Data Mining Techniques,2003

[13]RON KOHAVI,FOSTER PROVOST,Applications of Data Mining to Electronic Commerce,Data Mining and Knowledge Discovery, 5, 5–10, 2001

[14]Christian Bauckhage,Notes on k-Means Clustering (I),2010

[15]H. Hotelling, Simplified Calculation of Principal Components, Psychometrika 1 (1) (1936) 27–35.

[16]G.Milligan, M. C. Cooper, A Study of Standardization of Variables in Cluster Analysis, Journal of Classification
5 (2) (1988) 181–204.

ISSN: 2278 – 1323

Swati Nenava ,Swati Nenava was born in Bhopal madhyapradesh India,she received B.E(Bachelor of Engineering)degree in Information and Technology from Patel Institute of Technology,Bhopal and currently pursuing M.E in Shri Vaishnav Institute of technology and Science,Indore,in Department of Information Technology

Prof. Manoj Chauhan was born in India,he received B.E(Bachelor of Engineering) degree from Madhav Institute of Technology and Science,Gwalior,M.P and M.E from Institute of Engineering &Technology, Indore (M.P.) His researches publications can be seen in International and National Journals as well as membership as Editorial Board Member and Reviewer in International Research Group and Journal, currently he is professor in Shri Vaishnav Institute of Technology and Science, Indore in Department of Information Technology

ISSN: 2278 – 1323