IOP Conference Series: Materials Science and Engineering

An Efficient Fuzzy C-Least Median Clustering Algorithm

An Efficient Fuzzy C-Least Median Clustering Algorithm

Moksud Alam Mallik[1], Nurul Fariza Zulkurnain[2], Mohammed Khaja Nizamuddin[3], and Aboosalih KC[4]

[1]International Islamic University Malaysia, Kuala Lumpur, Malaysia
ACE Engineering College,Jawaharlal Nehru Technological University,Hyderabad,India, malamdot@gmail.com
[2] International Islamic University Malaysia, Kuala Lumpur, Malaysia.nurulfariza@iium.edu.my
3ACE Engineering College,Jawaharlal Nehru Technological University Hyderabad,India mknizams@yahoo.com
4 Homar Technologies, Hyderabad,India, abswalih@gmail.com

Abstract. In today's reality 'World Wide Web' is considered as the archive of extremely enormous measure of data. The substance and complexity of WWW are increasing day by day. Presently the circumstances are such that we are suffocating in data yet starving for knowledge. Because of these circumstances data mining is extremely important to get valuable data from WWW.Clustering data mining is the process of putting together meaning-full or use-full similar object into one group. It is a common technique for statistical data, machine learning and computer science analysis. Clustering is a kind of unsupervised data mining technique which describes general working behavior, pattern extraction and extracts useful information from time series data. In this paper we are discussing our new procedure for clustering called Fuzzy C-least median of squares algorithm which is an improvement to Fuzzy C-means (FCM) algorithm. As it is concerned with the least value among medians, it wipes out means squared error and eliminates the effect of outliers. We compared our clustering result got by applying FCM and FCLM by using Xie-Beni Index, Fukuyama-Sygeno Index and Partition Coefficient. The outcomes demonstrate a clear improvement of our algorithm than existing FCM algorithm.

Keywords: Clustering, Fuzzy Clustering, Fuzzy C-means Clustering, Fuzzy Clustering Least Median,

1. Introduction

Clustering is the process of grouping data elements given the likeness or similarity. Clustering is the process of grouping the data into *clusters*, so that objects within a cluster have high similarity in comparison to one another but are very different from objects in other clusters [1]. So items in the same class will be more similar and items in different classes will be more divergent. We can view cluster as subsets of a data set. One of the classifications of clustering is hard clustering and fuzzy clustering. Hard clustering is similar to mutually exclusive sets in set theory. In hard clustering a data set completely has a place with a cluster or does not have a place with a set i.e. on the off chance if it is part of a particular set it will not be part of any another set[9, 14]. Clustering has its roots in any areas, with data mining, statistics, biology, and machine learning.

Fuzzy clustering promotes fuzzy membership. Here one data set can have amembership to more than one cluster. It implies that one data set can fit in withseveral clusters simultaneously. Each data set will have a level of membership toevery cluster; membership to some cluster will be high while membership to someother cluster will be low. The membership worth will be between zeros to one. Theaggregate estimation of membership of one session to every cluster centers will beone. Data fuzzy clustering should be managing to practically fit reality. For example if a data set is on the boundary between two or moreclusters fuzzy clustering will give it partial membership among clusters [9, 14].

Fuzzy C-means clustering algorithm or FCM is one of the well-known fuzzy clustering algorithms. FCM allows one data point to be a part of more than one clusters or cluster centres. Assume we have a set of data points X = {x_1; x_2; :::; x_n} where n is the total number of data points and V = {v_1; v_2; :::; v_c} is the set of cluster centres where c is the total number of cluster centres. Each data point has membership in every cluster centre. The membership value of ith session to jth cluster centre can be represented as u_{ij} , it can be calculated as[9, 14].

$$\text{Uij}=\frac{\frac{1}{d_{ij\,(\text{xi,vj})}^{2}}\left(\frac{1}{m-1}\right)}{\sum_{k=1}^{c}\frac{1}{d_{ij\,(xi,\text{vk})}^{2}}^{\frac{1}{(m-1)}}}\text{..} (1)$$

The algorithm is based on the following objective function.

$$\text{Jm(UV X)} =\sum_{i=1}^{n} d_{ij}^{2} \sum_{j=1}^{c} u_{ij}^{m}(\text{Xi - Vj}) \text{ ....} (2)$$

Optimal fuzzy clustering of X is characterized as pairs (U, V) that locally minimize J_m. It means cluster centres which give lesser worth for J_m is the set of optimal cluster centres[9, 14].The fuzzifier m decides the level of cluster fuzziness. It is also known as weighting exponent. The value is 1≤ m≤1. If that m = 1 the clustering is same as hard clustering. As m increases past 1 the clustering turns out to be fuzzier. No theoretical or computational confirmation recognizes an ideal m. For most data, the range 1:5≤ m ≤3:0 gives good results [9, 14].

FCM Algorithm

The step by step algorithm is described as following:

Step 1: Arbitrarily select some data points as initial cluster centres. The number of initial cluster centres can be from 2 to n-1, where n is the aggregate number of data points.
Step 2: Calculate the membership value uij for every i and j by using the above equation for uij.
Step 3: Calculate new cluster centres Vj by using the above equation for Vj with the new membership values got from Step 2.
Step 4: Calculate objective function Jm(UV X) by using equation (2).
Step 5: Repeat Steps from 2 to 4 till we get minimum value for the objective function or till the difference between the previous and current value is very less.
Step 6: The cluster centres which are having minimum value for objective function is considered as optimal [9, 14].

$d^2_{ij}(X_{i-}V_j)$ is the square of Euclidean distance between data point X_i and cluster centre V_j. The Euclidean norm is the only choice for which broad experience with geological data is accessible [14] In each iteration new cluster centres will be calculated. The cluster centres will be calculated by the following formula.

$$\text{Vj} =\frac{\sum_{i=1}^{n} u_{ij}^{m}\, Xi}{\sum_{i=1}^{n} u_{ij}^{m}} \text{ ....} (3)$$

This work provides an improvement of the Fuzzy C-Means clustering algorithm known as Fuzzy c-least median of squares algorithm which minimizes mean squared error and also reduces the effect of outliers.

Section 2 gives the Literature Review, Section 3 gives the proposed algorithm, Section 4 elaborates the experimental results and the conclusion and future enhancement of this paper are made in Section 5.

2. Proposed Algorithm

Here we are proposing a new Fuzzy clustering algorithm called "Fuzzy C-least median of squares algorithm" which deals with the least of medians while selecting cluster centres.Our method reduces mean squared error and eliminates the effect of outliers.

2.1 Algorithm

Most of the existing systems are using FCM. It is an improvement over Crisp clustering algorithm. But it causes mean squared error. These issues can be minimized by making utilization of fuzzy set based clustering approaches. This work utilizes another new algorithm named as "Fuzzy c- least median of squares algorithm" which tries to minimize mean squared error and it eliminates the effect of outliers in the result.
Fuzzy c-least median of squares algorithm is similar to Fuzzy c-means algorithm but with some differences. Like Fuzzy c-means it also uses fuzzy set theoretic approach of partial membership [2, 10].The algorithm calculates cluster centers and assigns membership values (u_{ij}) to each data item corresponding to each cluster. The membership value ranges from 0 to 1.

2.1.1 Fuzzy Membership function

Assume X = {x_1; x_2; :::; x_m} is the set of data points or sessions. Each point is a vector of the form i = 1…m , x_i = {x^i_1; x^i_2; :::; x^i_n}. Let V ={v_1; v_2; :::; v_c} is a set of n dimensional vectors corresponds to c cluster centers and each cluster center is a vector of the form 8j = 1:::n , v_j = {v_1^j; v_2^j; :::; v_n^j}. Let u_{ij} represents membership of data point(or session) x_i in cluster j. The m×c membership matrix U = [u_{ij}] shows allocation of sessions to various cluster centres. It satisfies following criteria.

$$\sum_{j=1}^{c} u_{ij} = 1;\ \forall \mathrm{i} = 1 \ldots \mathrm{m}$$

$$0 < \sum_{i=1}^{m} u_{ij} < \mathrm{m},\ \forall \mathrm{j} = 1 \ldots \mathrm{c}$$

The membership value is calculated by using following formula [14].

$$\mathrm{Uij}=\frac{\frac{1}{d^2_{ij\,(\mathrm{xi,vj})}}^{(\frac{1}{m-1})}}{\sum_{k=1}^{c} \frac{1}{d^2_{ij\,(xi,\mathrm{vk})}}^{\frac{1}{(m-1)}}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Initial cluster centres are randomly selected from available sessions. Then the membership value of each cluster is calculated using the equation for u_{ij}. Euclidean distance between various data points and cluster centres can be calculated using the following equation [9, 14].

$$d^2_{ij\,(xi,vj)}=\sum_{k=1}^{n} d^2_{ik\left(x^i_{k,} v^j_k\right)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Where n is the number of dimensions of each data point, $x^i_{k,}$ is the value of k^{th} dimension of x_i and v^j_k is the value of k^{th} dimension of v_j which is the jth cluster center.

2.1.2 Cluster Center calculation

In each iteration steps new cluster centres are calculated using the following formula:
Di = Median{(d_{ij} (s_k - s_i) *u_{ij})}; ∀i ≠ k; k = 1…n
p = Argmin{(Di:n); ∀i = 1…n}
vj = sp

Initially we will take one session say s_1 and find the distance between this session to every other session (say s_2; s_3; s_4; ::::s_n) multiplied by membership function of s_1 to cluster center 1(v_1). Next we will sort these values in ascending order and take the median. The above step will be done for all session's s_1; s_2; s_3; s_4; ::::s_n. Now from these medians got from above steps least value will be taken. The session corresponding to the least value will be taken as the first cluster center in this round .All above steps will be continued for cluster center 2 upto cluster center c(v_1; v_2; v_3; ::::v_c). In this way we will get new sets of cluster centres in one round. New cluster centres will be calculated upto a specific number of rounds till we get optimal cluster centres.

2.1.3 Objective function calculation

The objective function is used to measure the performance of the clustering algorithm. The cluster centres which are having less value for the objective function will give compact clusters or better clustering results. The performance index is calculated using the following objective function:

$$\mathrm{Oi} = \sum_{j=1}^{c}(\mathrm{uij} * \mathrm{dij}(\mathrm{vj} - \mathrm{si}))\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

2.1.4 Algorithm Termination

The steps in sections 2.1.1, 2.1.2, 2.1.3 will be continued till a specified number of steps or till we get minimum value for the objective function. The corresponding value of the cluster centres will be taken as final.

3. Clustering Assessment

Fuzzy clustering validity criterion assesses the nature of fuzzy segments delivered by the clustering algorithm. Clustering can be evaluated in light of the information that was clustered itself, this is called inner evaluation. These methods as a rule appoint the best score to the algorithm that delivers clusters with high similitude within a cluster and low comparability between clusters. The inner assessment measures are most appropriate to get some knowledge about the circumstances where one algorithm performs better than other [9]. Diverse routines can be utilized to survey the quality of clustering algorithms taking into account the internal criterion. Here we are utilizing two fundamental strategies:Xie-Beni Index and Fukuyama-Sugeno Index:-

3.1 Xie-Beni Index

This function relies on the data set, geometric separation measure, the distance between cluster centroids and more basically on the fuzzy allotment created by any fuzzy algorithm utilized.

3.2 Legitimacy criteria for Hard and Fuzzy Clustering

Partition coefficient F measures the sum coinciding between clusters.

$$\mathrm{F} = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n}\left(u_{ij}^{2}\right)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

F is inversely proportional to the overall average overlap between the fuzzy subsets. If there is no membership sharing between any pair of fuzzy clusters then F=1. Solving $\mathrm{Max}_c\{F\}$where (c = 2, 3,…..,n-1) is accepted to deliver substantial clustering of the data set X. A disadvantage is the monotonic diminishing tendency of partition coefficient with increase in the number of cluster centres.[11, 12].

3.3 Compact and Separate Fuzzy validity Criterion

This legitimacy function measures the general average compactness and separation of a fuzzy C-partition. Consider a Fuzzy c-partition of data set X = {x_1; x_2; ::::; x_n} and cluster centroids of every cluster is V = {v_1; v_2; ::::; v_c} and u_{ij}(i = 1; 2; ::::; c; j = 1; 2; :::::; n) as the fuzzy membership of data point j in cluster i. Distance between data point X_j and cluster center V_i can be found by using the equation.

$$d_{ij}^{2}(\mathrm{X_j}, \mathrm{V_i}) = \sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{d}_{\mathrm{ik}}^{2}\left(\mathrm{x}_{\mathrm{k}}^{\mathrm{j}}, \mathrm{v}_{\mathrm{k}}^{\mathrm{i}}\right)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8)$$

Also $n_i = \sum_{x\,j} u_{ij}$ the fuzzy number of vectors in cluster i, then we can say $\sum_{x\,i} n_i$ =n is the total number of data points in X[11].

For every cluster i summation of the squares of the fuzzy deviation of every data point signified by σ_i is known as the variation of cluster i or class i. i.e. $\sigma_i = \sum_x j(d_{ij})^2 = (d_{i1})^2 + (d_{i2})^2 + \ldots + (d_{in})^2$. The summation of the variation of all classes signified by σ is called an aggregate variation of class X concerning Fuzzy C segment,i.e. $\sigma = \sum_{xi} \sigma_{i=} \sum_x i \sum_x j(d_{ij})^2$. A better C partition ought to result in a smaller σ . In a situation when Fuzzy C partition is obtained by utilizing the Fuzzy C-means algorithm with m=2 the estimation of σ will be equal to the Fuzzy C-means objective function[11].

The ratio (denoted by π) of the aggregate variation to the size of the data set is called Compactness of Fuzzy C-partition of the dataset ,$\pi = (\sigma/n)$. Estimation of π will be smaller for more compact clusters. It is a function of dispersion attributes of data set itself and a function of how we separate the data points into clusters. It is free of the number of data points. For a given data set a smaller π shows that we have come to a partition with more compact clusters, thus demonstrating a better partition. The quantity $\pi_i = (\sigma_i/n_i)$ is called compactness of class i or average variation of class i[11].

The separation of Fuzzy C-partition is $s = (d_{min})^2$, where d_{min} is the minimum distance between cluster centroids, i.e. $d_{min} = \min_{ij}\|V_i - V_{j\|}$. A greater s demonstrates that all clusters are all around isolated[11].

Compactness and Separation validity function(or Xie-Beni Index) S is characterized as a proportion of compactness to the separation s, ie S = π /s. Substituting for π,s we get $S = (\sigma/n)=(d_{min})^2$. A smaller S demonstrates a partition in which all clusters are generally compact and separate to one another. So our objective is to discover Fuzzy C-partition with smallest S. S can be composed as

$$S = \frac{\sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^2 \|v_i - x_j\|2}{n.min_{ij} \|v_i - v_j\|2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9)$$

S is not internal to the clustering algorithm because the definition of S does not depend on the algorithm that is used to acquire u_{ij}. For Fuzzy C-Means algorithm with m=2 S can be demonstrated to be

$$S = \frac{j_2}{n} * (\mathrm{dmin})2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

where J_2 is the objective function of FCM. Minimizing S means minimizing J_2. Further the clusters, bigger the value of $(d_{min})^2$ and smaller the value of S. In this way smallest S undoubtedly shows an optimal partition [11].

We already stated that S diminishes monotonically when c gets substantial and near n i.e. S will tend to in the long run diminish when c is very large. So the estimation of S is unimportant when c draws near to n.Fortunately this is not a serious problem since in practice the practical number of clusters c is much smaller than the number of data points n. There are distinctive heuristics to determine stop value for c. One of the technique is that plotting the ideal value of S for c=2 to n-1, then selecting the initial stage of monotonically diminishing inclination as the largest c to be considered [11]. Let C_{MAX} indicate such a then c can be found by[11]

$Min_{(2 \le c \le C_{max}} \{\mathrm{Min}\Omega_c(\mathrm{S})\}$.

3.4 Fukuyama-Sugeno Index

It exploits cohesion and separation of clusters. Here the first term denotes compactness measure and the second term is the level of separation between every cluster and the mean(V) of cluster centroids. As the separation goes higher and the compactness measure comes down, the better is the clustering. For compact and all around isolated clusters we expect little values for FS Index. It is calculated using the following formula [13].

$$\mathrm{FSId} = \sum_{j=1}^{n} \sum_{i=1}^{c} (u_{ij}) m \left\|X_j - V_i\right\|2 - \sum_{i=1}^{c} \|V_i - \tilde{V}\|2, \text{ Where } \tilde{V} = \frac{1}{c} \sum_{i=1}^{c} V_i \ldots\ldots\ldots\ldots\ldots\ldots..(11)$$

4 . Experiment & Results:

4.1 Starting Cluster center choice

Above all we have to randomly choose initial cluster centres for processing. With a specific end goal to choose the number of optimum clusters we computed three validity indexes Xie Beni Index, Fukuyama Sugeno Index and Partition coefficient on initially selected cluster centres.The Initial cluster centres chosen by this method can be utilized as input to both FCM and Fuzzy Least median of squares Algorithm. We have a sum of 951 sessions in our input file.

For Cluster centre choice we haphazardly pick sessions in distinctive reaches. Fuzzy c-Mean clustering is initially applied by choosing the number of clusters as 2.Amid each of the cycles we increase the number of clusters by 1 till the number of clusters reaches the available upper limit. We repeated the above process for upper limit of cluster centres 50,60,100,150,250,350,450,550,650,750,850,950.The Algorithm for Cluster Center determination is as per Table 1. After completing this step, we will get an initial set of cluster centres.Now we can run FCM and FCLM on this set of cluster centres.

Table 1: Algorithm to select initial cluster centres.

```
Step 1: Put all Session numbers in an Array List and shuffle it to different positions in
 Array List.
Step 2: Then select first n sessions From Array List as initial cluster centres.
Step 3: Calculate three validity indexes Xie Beni Index,Fukuyama SugenoIndex and Partition
  coefficient on the FCM Algorithm.
Step 4: Store the yield in file for future comparison.
Step 5: Do Steps from 1 to 4 for n = 2; 3; :::; Limit.
Step 6: Do Step 5: For Limit ranging from 50, 60, 100, 150, 250, 350, 450, 550, 650, 750,
850 and 950.
Step 7: For every Limit i.e. for cluster centres from n = 2 to Limit draw column chart
and analyze the Validity indexes.
Step 8: In every Limit we can choose cluster centres which are having less esteem for
 Xie-Beni Index and FS Index and high esteem for Partition coefficient.
```

There is one issue with Validity indexes. That is it is having monotonic diminishing tendency with expanding the number of clusters. To keep away from outcomes of this impact we can drop values from where monotonic diminishing tendency begins. To accomplish this, we have drawn chart for Fukuyama Sugeno Index,Xie-Beni Index and Partition coefficient for each number of cluster centres of all ranges or "Limits". Figure 1 shows FS Index After evacuating Index over 14,000 and cluster centres more noteworthy than 115 because for time being we are mostly taking Xie beni Index as standard and it is having just cluster centres up to 115. Figure 2 shows Xie Beni Index by deleting value for cluster centres more than 115 and for index values less than 20,000. Figure 3 shows PC Index after deleting indexes less than 0.1 and cluster centres up to 81.We have a table for each Index i.e. each chart. Now from these three tables of Figure 1, 2, 3 (i.e. three index values) we will choose Index value "For 50" and draw the graph. This procedure will be proceeded for each Limit from For 50, For 60, For 100, For 150, For 250,..., For 950. From each Limit we are selecting at least three cluster centres. Taking after are cluster centres for every Limit. It is shown in Table 2. These cluster centres are given as input to both FCM algorithm and FCLM algorithm.

Table 2. Cluster centre Selection Table for FCM.

Limit	Cluster centres selected
From For 50	2,5,34
From For 60	2,4,19
From For 100	2,4,5
From For 150	3,7,15
From For 250	7,9,22
From For 350	2,4,32
From For 450	2,3,18
From For 550	2,4,18
From For 650	5,7,39
From For 750	2,6,17
From For 850	3,10
From For 950	4,6,19

Chart for Index value comparison for the range "For 50" is indicated as a graph in Figure 4. Chart for Index Value Comparison of "For 250" is shown in Figure 5. Like this we have a chart for every ranges or limit of cluster centres.

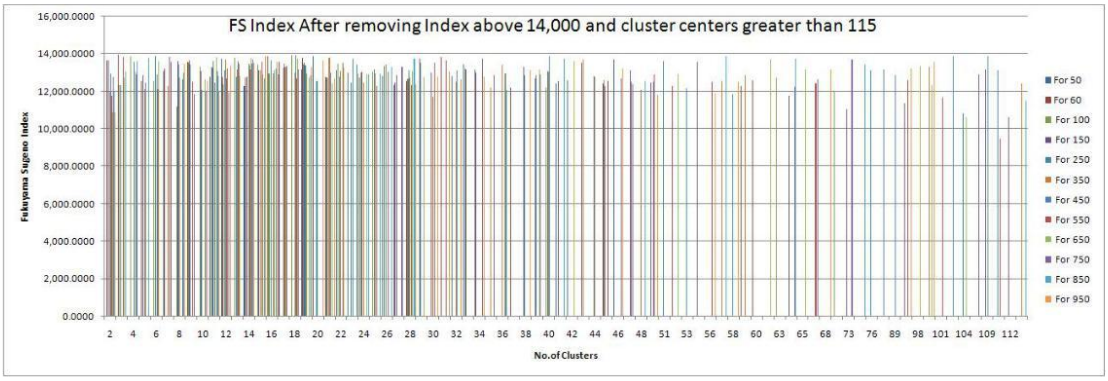

Figure 1. FS Index After removing Index above 14,000 and cluster centres greater than 115

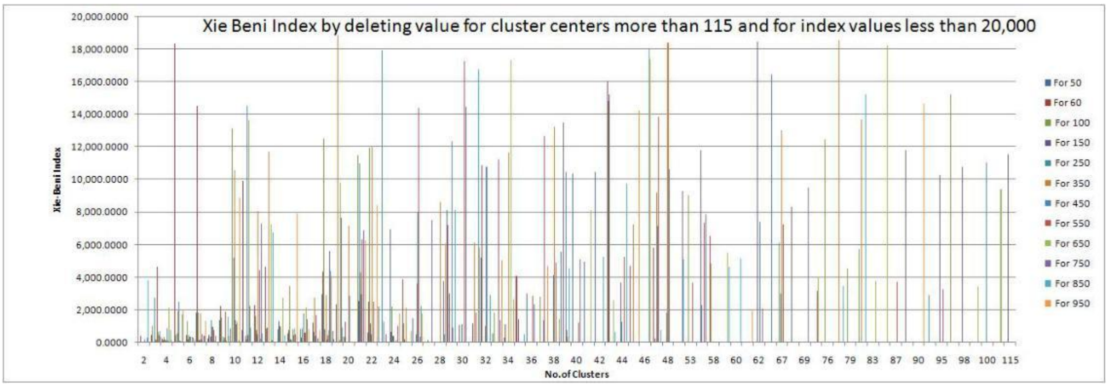

Figure 2. Xie Beni Index by deleting value for cluster centres more than 115 and index values less than 20,000

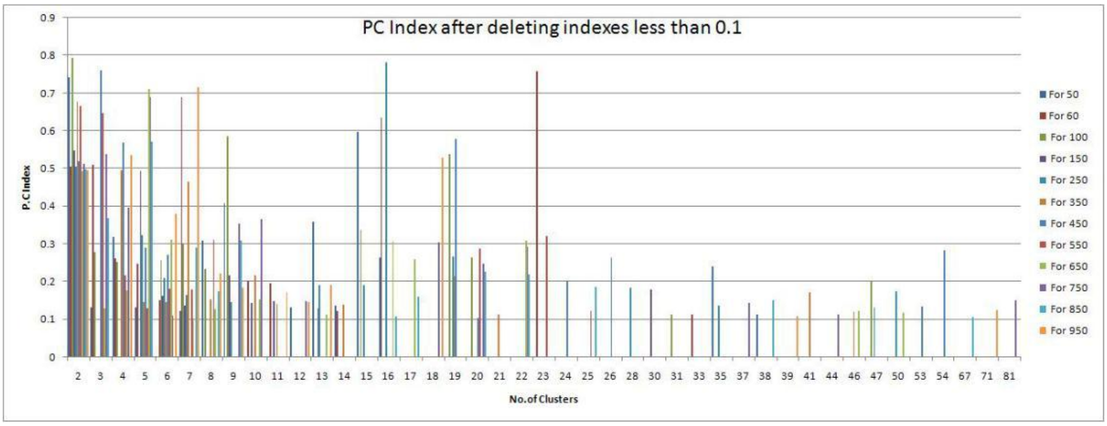

Figure 3. PC Index after deleting indexes less than 0.1

Comparison between FCM and FCLM

For both FCM and FCLM algorithms, we will give with the same input data. That is, we will give same introductory cluster centres and calculate Xie-Beni Index, Partition Coefficient, FS Index, Deviation, Compactness, Separation of cluster centres as a whole. Also we will compute Fuzzy compactness and fuzzy deviation of individual cluster centres. After that we will compare index values acquired by both FCM and FCLM algorithm. The algorithm which is getting less esteem for validity index performs better other than Partition coefficient. For partition coefficient the algorithm which gets higher value performs better (It's most extreme value is "one").

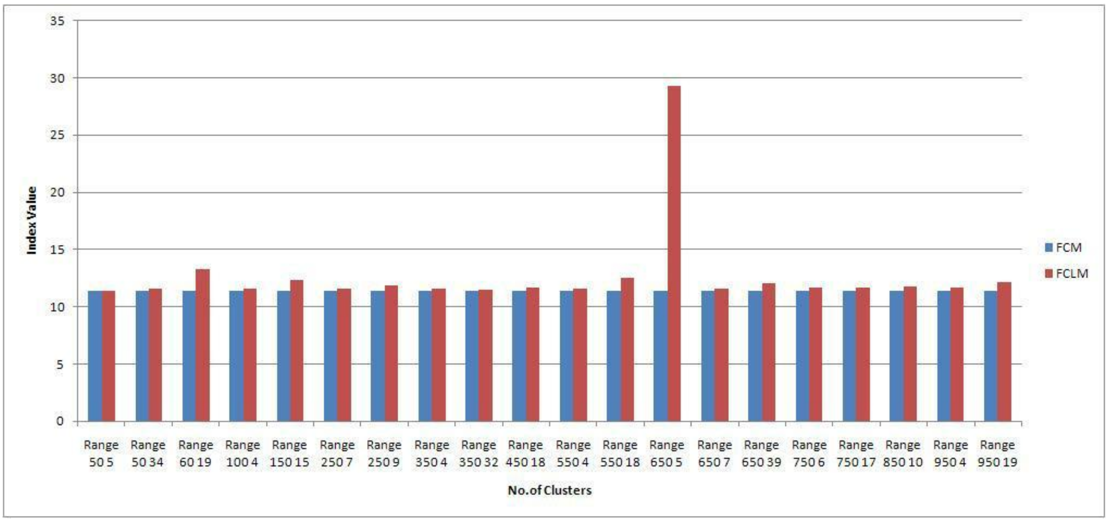

Figure 4. Comparison of Compactness of FCM and FCLM for different input Cluster centres

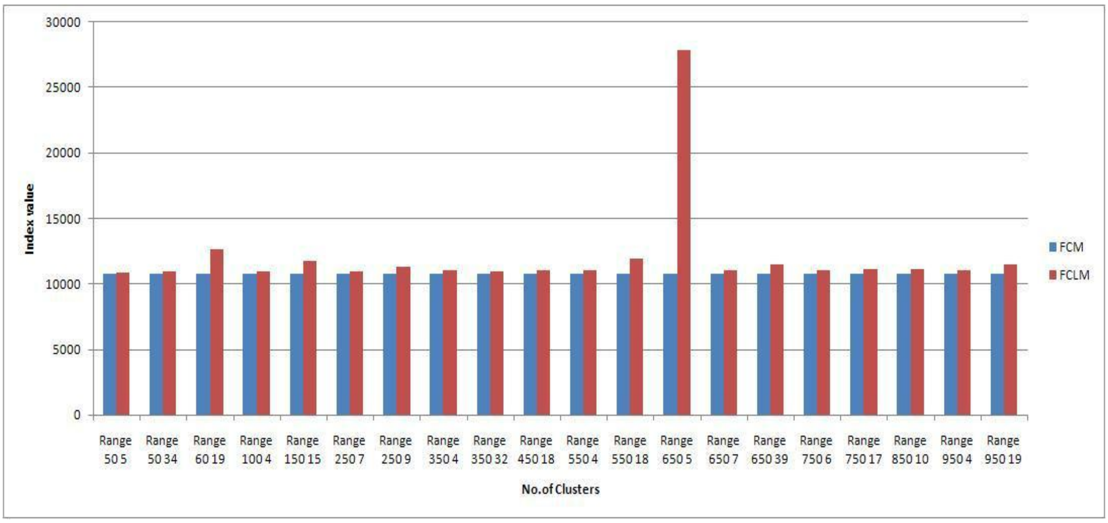

Figure 5. Comparison of Deviation of FCM and FCLM for different input Cluster centres

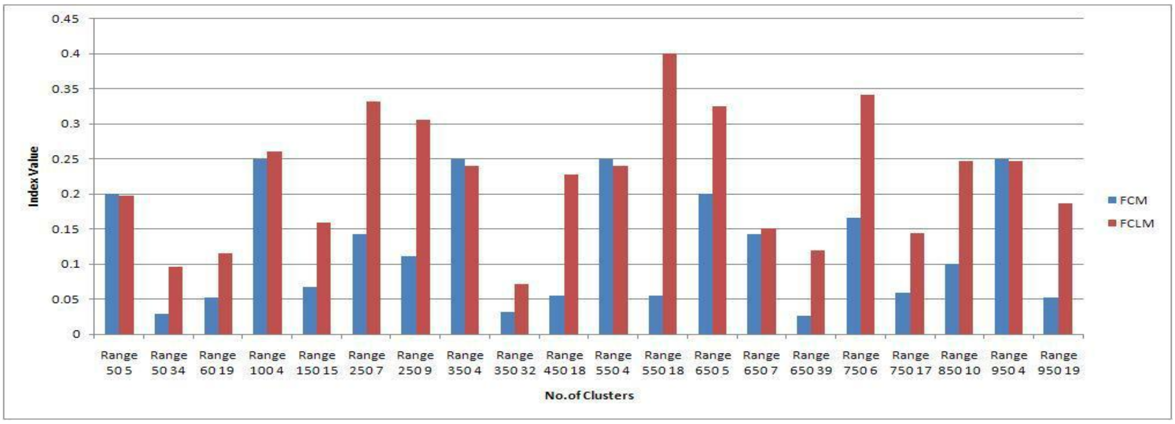

Figure 6. Comparison of Partition Coefficient of FCM and FCLM for different input Cluster centre

In all our comparison charts we are getting improvements for FCLM over FCM other than Deviation and Compactness. Here in every one of the cases we are getting less value for Xie-Beni Index of FCLM over FCM, less value for FS Index of FCLM over FCM, high value for Partition Coefficient of FCLM over FCM, high value for Separation of FCLM over FCM. By examining enhancements in these values we can say that our algorithm enhances performance in clustering.

5. Conclusion and future work

Fuzzy c-least median of squares algorithm improves clustering. It reduces the effect of mean squared error and also eliminates the effect of outliers and hence the noise. So we are getting better clustering solutions than in fuzzy c-means algorithm. With better data sets we can get better solutions in our algorithm.

Both FCM and FCLM performances are influenced by initial cluster centre choice. To escape this negative impact we can apply mountain density function to choose introductory cluster centres. So we will get fitting starting cluster centres. In preprocessing also we need to do a lot of improvements like applying new methods to filter robot requests.

References

[1] Han, J., Kamber, M., and Pei, J.,*Data mining concepts and techniques*, third edition," 2012.

[2] Zahid Ansari, Mohammad Fazle Azeem, A. V. B. and Ahmed, W., A fuzzy clustering based approach for mining usage pro_les from web log data," (IJC- SIS) *International Journal of Computer Science and Information Security*, Vol. 9, No. 6, 2011, vol. 9, pp. 70{79, JUN 2011.

[3] Cooley, R., Mobasher, B., and Srivastava, J.*, Web mining: information and pattern discovery on the world wideweb*," in Tools with Artificial Intelligence, 1997. Proceedings., Ninth IEEE International Conference on, pp. 558{567, Nov 1997.

[4] Castellano, G., Fanelli, A., Mencar, C., and Torsello, M., *Similarity-based fuzzy clustering for user pro_ling*," in Web Intelligence and Intelligent Agent Technology Workshops, 2007 IEEE/WIC/ACM International Conferences on, pp. 75{78, Nov 2007.

[5] Castellano, G., Mesto, F., Minunno, M., and Torsello, M. A., *Web user profiling using fuzzy clustering*.," in WILF (Masulli, F., Mitra, S., and Pasi, G., eds.), vol. 4578 of Lecture Notes in Computer Science, pp. 94{101, Springer, 2007.

[6] Nasraoui, O., Frigui, H., Krishnapuram, R., and Joshi, A., \Extracting web user profiles using relational competitive fuzzy clustering.," *International Journal on Artificial Intelligence Tools*, vol. 9, no. 4, pp. 509{526, 2000.

[7] Castellano, G., Fanelli, A. M., and Torsello, M. A., \Mining usage profiles from access data using fuzzy clustering," in Proceedings of the 6th WSEAS *International Conference on Simulation, Modelling and Optimization*, SMO'06, (Stevens Point, Wisconsin, USA), pp. 157{160, World Scientific and Engineering Academy and Society (WSEAS), 2006.

[8] Ansari, Z., Azeem, M. F., Babu, A. V., and Waseem, A., \A fuzzy approach for feature evaluation and dimensionality reduction to improve the quality of web usage mining results," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 2, no. 6, pp. 67{73, 2012.

[9] Wikipedia, Wikipedia The free encyclopedia, 2014.

[10] Ansari, Z., Babuy, A., Ahmed, W., and Azeemz, M., A fuzzy set theoretic approach to discover user sessions from web navigational data," in *Recent Advances in Intelligent Computational Systems (RAICS*), 2011 IEEE, pp. 879{884, Sept 2011.

[11] Xie, X. L. and Beni, G., A validity measure for fuzzy clustering," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 13, pp. 841{847, Aug. 1991.

[12] Bezdek, J., Numerical taxonomy with fuzzy sets," *Journal of Mathematical Biology*, vol. 1, no. 1, pp. 57{71, 1974.

[13] Y. Fukuyama, M. S., *A new method of choosing the number of clusters for fuzzy c-means method*," in Proceedings of the 5th Fuzzy System Symposium, pp. 247{250 (in Japanese), 1989.

[14] Bezdek, J. C., Ehrlich, R., and Full, W., Fcm: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2, pp. 191{203, 1984.