International Journal of Pure and Applied Mathematics
Volume 116 No. 9 2017, 319-325
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

COMPARATIVE OBSERVATION AND PERFORMANCE ANALYSIS OF MULTIPLE ALGORITHMS ON IRIS DATA

[1]C.Nalini, [2]C. Raghavendra, [3]KRajendra Prasad
[1]Professor, CSE Department, BIST,BIHER, Bharath University, Chennai.
[2,3]CSE Department, Institute of Aeronautical Engineering, Hyderabad.
[1]nalini.cse@bharathuniv.ac.in

Abstract: Information mining innovation has risen as methods for distinguishing examples and patterns from huge amounts of information. Information mining is a computational insight teaches that contributes devices for information examination, disclosure of new learning, and self-sufficient basic leadership. Bunching is an essential information depiction technique in information mining which gathering's most comparable information. There are different calculations used to tackle this issue in this paper, we utilize FCM (Fuzzy C - mean) bunching calculation and MFPCM (Modified Fuzzy Possibility C - mean) grouping calculation. In this paper we analyze the execution examination of Fuzzy C mean FCM) bunching calculation and contrast it and Modified Fuzzy possibility C mean calculation. In this, we looked at FCM and MFPCM calculation on various informational collections. We measure multifaceted nature of FCM and MFPCM at various informational indexes. FCM grouping is a bunching method which is isolated from Modified Fuzzy Possibililstic C imply that utilizes Possibililstic dividing.

Keywords: Data clustering Algorithm, Portioning, Data Mining, Fuzzy C Mean, Modified Fuzzy Possibililstic C mean.

1. Introduction

Information investigation is considered as a critical science in this present reality. Information mining innovation has risen as a method for recognizing examples and patterns from substantial amounts of information. Information mining is a computational insight train that contributes devices for information investigation, disclosure of new learning, and self-governing basic leadership. The assignment of preparing vast volume of information has quickened the enthusiasm for this field. As said in Mosley (2005) information mining is the investigation of observational datasets to discover unsuspected connections and to compress the information in novel ways that are both justifiable and valuable to the information proprietor. Information mining finds depiction through grouping perception[1-2], affiliation, successive investigation. Bunching is an essential information depiction strategy in information mining which gathering's most comparative information. Information grouping is a typical procedure for information examination[3-4], which is utilized as a part of many fields, including machine learning, information mining, design acknowledgment, picture investigation and bioinformatics. Bunch investigation is a procedure for characterizing information; it is a technique for discovering groups of an informational collection with most closeness in a similar bunch and most difference between various bunches. The traditional grouping strategies put each purpose of the informational collection to precisely one bunch. Since 1965, Zadeh proposed fluffy sets so as to come nearer of the physical world. Zadeh presented the possibility of incomplete participations portrayed by enrollment capacities. Bunching calculation segments an unlabelled arrangement of information into bunches as indicated by the similitude. Contrasted and the information order, the information grouping is an unsupervised learning process, it needn't bother with a marked informational collection as preparing information[5-6], yet the execution of the information bunching calculation is frequently significantly poorer. Despite the fact that the information grouping has better execution, it needs a named informational index as preparing information and marked information for the arrangement is regularly extremely troublesome and costly to get. So there are numerous calculations are proposed to enhance the grouping execution. Bunching is the arrangement of comparable items into various gatherings, or all the more definitely, the parceling of an informational index into subsets (groups)[7-8], with the goal that the information in every subset share some normal quality. Bunching strategy is utilized for joining watched objects into groups (gatherings), which fulfill two fundamental criteria: Each gathering or bunch ought to be homogeneous articles that have a place with a similar gathering are like each other.

1.1 Fuzzy C- Mean Algorithm

Fluffy C Mean (FCM) is an information bunching [13-14] procedure in which an informational collection is gathered into n groups with each information point in the dataset having a place with each bunch will have a high level of having a place or participation with that bunch and another information point that lies far from the focal point of a bunch will have a low level of having a place or enrollment with that bunch[9-10].

The means of FCM calculation given underneath. Fix c and c is (2<=c<n) and select an incentive for parameter m'. Introduce the parcel lattice U(0). Each progression in this calculation will be named as r, where r =0, 1, 2…
Calculate the c focus vector{vij} for each progression

Calculate the separation framework D[c,n].

Update the segment framework for the rth step, U® as take after: on the off chance that ||U(k+1)- U(k)||<δ at that point STOP: generally come back to step 2 by iteratively refreshing the bunch focuses and the participation grades for information point. FCM iteratively moves the group focuses "to one side" area inside a dataset.

1.2 Modified Fuzzy Possibililstic C - Mean Algorithm

The FPCM calculation endeavors to parcel a limited gathering of components X={x1, x2, x3… … xn} into an accumulation of c fluffy groups concerning some given model. Given a limited arrangement of information, the calculation restores a rundown of c cluster centers V, with the end goal that V=vi, i=1,2,3… … ,c And a segment framework U to such an extent that U=u_{ij}, i=1,2,3,… … c, j=1,2,… … n. Where, u_{ij} is a numerical incentive in [0, 1] that advises how much the components xj has a place with the i^{th} group. Characterizes a group of fluffy sets {Ai, i=1,2,3… … ..c} as a fluffy c parcel on a universe of information focuses X

- Fluffy set takes into account level of participation.
- A solitary point can have halfway participation in more than one class.
- There can be no void classes and no class that contains no information focuses.

The steps of MFPCM algorithm given below:

1. The objective function of the MFPCM can be formulated as follows:

$$J_{MFPCM}=\sum_{i=1}^{c}\sum_{j=1}^{n}(\mu_{ij}^{m}w_{ji}^{m}d^{2m}(x_j,v)+t_{ij}^{n}$$

2. Calculate U = {μ ij} represents a fuzzy partition matrix, is defined as:

$$u_{ij}=\left[\sum_{k=1}^{c}\left(\frac{d(x_j,v_i)}{d(x_j,v_k)}\right)^{2m/(m-1)}\right]^{-1}$$

3. Calculate T = {t ij} represents a typical partition matrix, is defined as:

$$t_{ij}=\left[\sum_{k=1}^{n}\left(\frac{d(x_j,v_i)}{d(x_j,v_k)}\right)^{2\eta/(\eta-1)}\right]^{-1}$$

4. Calculate V = {v ij} represents c centers of the clusters, is defined as:

$$v_i=\frac{\sum_{j=1}^{n}(\mu_{ij}^{m}w_{ji}^{m}+t_{ij}^{n}w_{ji}^{n})\times x_j}{\sum_{j=1}^{n}(\mu_{ij}^{m}w_{ji}^{m}+t_{ij}^{n}w_{ji}^{n})}$$

1.2 Daugman's Algorithm

This is by a wide margin the most referred to technique in the iris acknowledgment writing. It is authorized to Iridium Technologies1 who transformed it into the premise of 99.5% of the business iris acknowledgment frameworks. It was proposed in 1993 and was the main technique adequately actualized in a working biometric framework. The creator accept both understudy and iris with roundabout frame and the Integro-differential administrator[15-16]. We executed two pre-process operations with the reason for picture differentiate upgrade, trusting that they could add to the change of the outcomes:

1.3 Histogram Equalization

This operation enhances the complexity between each eye's areas, which possibly will encourage the division errand.

1.4 Binarization

The picture binarization - in view of an edge - is an exceptionally regular operation that amplifies the detachability between the iris locales and the staying ones. This procedure has[17-18], notwithstanding, one noteworthy drawback: it is profoundly reliant of the picked edge, and as picture qualities change, the outcomes may truly fall apart. In addition, the binarization bargains one of the Daugman‟s strategy Daugman's calculation depends on applying an integro-differential administrator to discover the iris and student shape.

$$\max(r, x_0, y_0) \,|\, G_\sigma(r) * \frac{\partial}{\partial r_{r, x_0, y_c}} \oint \frac{I(x,y)}{2m} ds$$

1.4 Equation 1: Daugman's integro Differential Equation

Where X0, Y0, ro: the centre and radius of coarse circle (for each of pupil and iris). Gσ(r): Gaussian function. Δr: the radius range for searching for. I(X, Y): the original iris image. *G*σ(*r*) is a smoothing function[19-20], the smoothed image is then scanned for a circle that has a maximum gradient change, which indicates an edge. The above algorithm is done twice, first to get the iris contour then to get the pupil contour. It worth mentioning here the problem is that the illumination inside the pupil is a perfect circle with very high intensity level (nearly pure white). Therefore, we have a problem of sticking to the illumination as the max gradient circle. So a minimum pupil radius should be set. Another issue here is in determining the pupil boundary the maximum change should occur at the edge between the very dark pupil and the iris, which is relatively darker than the bright spots of the illumination. Hence, while scanning the image one should take care that a very bright spot value could deceive the operator and can result in a maximum gradient. This simply means failure to localize the pupil.

1.5 Optimized Daugman's Algorithm

As an answer for this issue, change to theintegro-differential administrator is proposed to disregard all circles if any pixel on this circle has an esteem higher than a specific edge[21]. This edge is resolved to be 200 for the greyscale picture. This guarantees just the splendid spots values normally higher than 245 – will be wiped out.

Another arrangement we considered is to treat the brightening by Truncating pixels higher than a specific limit – splendid spots – to dark. Yet, this technique flopped in many pictures, this is on the grounds that when the spot hits the student the brightening spreads on the understudy so as we treat the enlightenment spots it will abandon a most extreme change edges that can't be resolved and the administrator will think of it as the student limit. The arrangement of the Algorithm methodology is cleared in the flowchart indicated below.The false acknowledgment rate for the iris acknowledgment framework is 1 out of 1.2 million, measurably superior to the normal unique mark acknowledgment framework. The genuine advantage is in the false dismissal rate, a measure of confirmed clients who are rejected. Unique mark scanners have a three percent false dismissal rate, while iris filtering frameworks gloat rate at the 0% level.

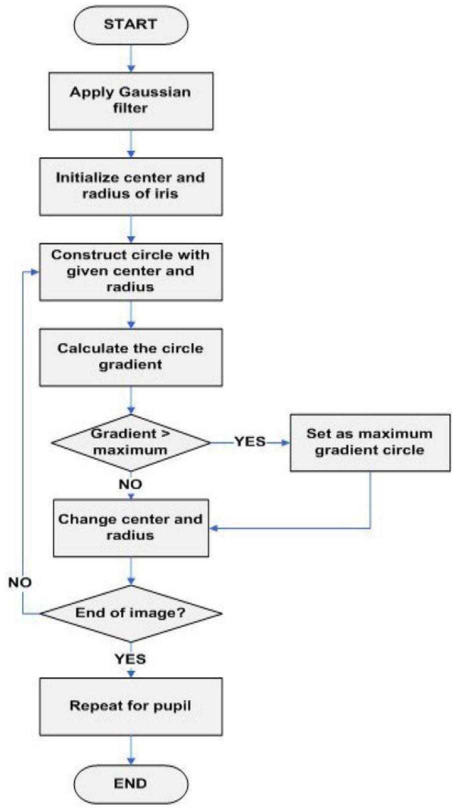

Figure 1. Flow Diagram showing Daugman's algorithm.

2. Results

2.1 Time many-sided quality of FCM and MFPCM by changing no. of Clusters on Iris Data:

The usage of FCM and MFPCM is done on iris Data in MATLAB. The information t contains 3 classes of 150 examples every, where each class alludes to a sort of iris plant. One class is straightly distinct from the other two, the last are NOT directly distinguishable from each other. The informational index contain four property which are given underneath

The time multifaceted nature of FCM [11] is O(ndc2i) and time unpredictability of MFPCCM is O(ncdi). Presently keeping no. of information focuses steady, lets accept n=100, d=3, i=20 and fluctuating no. of bunches, we get the accompanying table and diagram. Where n= number of information point, c= number of bunch, d= measurement, i= number of cycle

Table 1. Time Complexity when Number of cluster varying

Sl. No.	Number of cluster	FCM Time Complexity	MFPCM Time Complexity
1	1	2000	2000
2	2	10000	5000
3	3	26000	8500
4	4	45000	11000

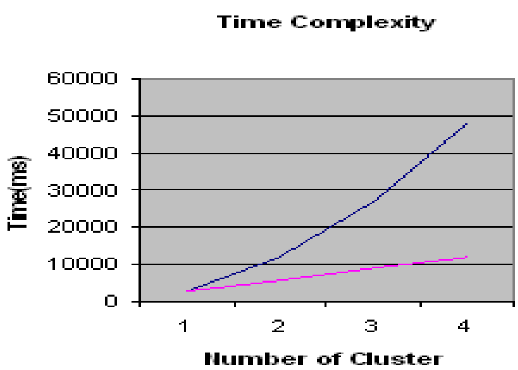

Figure 2. Time complexity of FCM and MFPCM by varying no. of Clusters

Now keeping no. of cluster constant, lets assume n=140, d=3, c=3 and varying no. of Iteration, we obtain the following table and graph.

Table 2. Time Complexity when Number of Iterations varying

Sl. No.	Number of Cluster	FCM Space Complexity	MFPCM Space Complexity
1	10	400	3
2	15	600	6
3	20	800	9
4	25	1000	12

Sl. No.	Number of Iteration	FCM Time Complexity	MFPCM Time Complexity
1	5	5000	4000
2	10	11000	5000
3	15	15000	7000
4	20	25000	11000

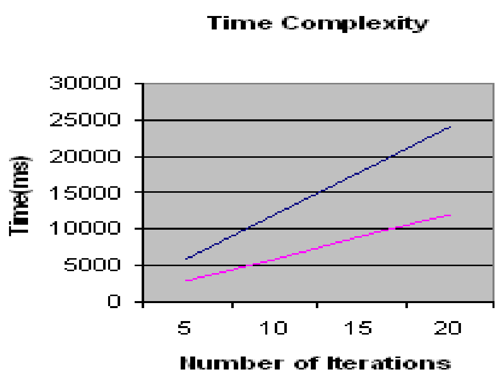

Figure 3. Time complexity of FCM and MFPCM by varying no. of Iterations

2.2 Comparison of space complexity of FCM and MFPCM:

The space complexity of FCM is O(nd+nc) and MFPCM is O(cd). Now keeping no. of data points constant, lets assume n=140, d=3 and varying no. of clusters we obtain the following graph.

Table 3: Space Complexity when Number of Clusters varying

Sl. No.	Number of Cluster	FCM Space Complexity	MFPCM Space Complexity
1	10	400	3
2	15	600	6
3	20	800	9
4	25	1000	12

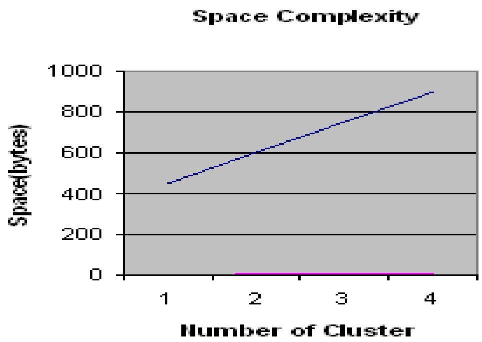

Figure 4. Space Complexities Of FCM And MFPCM By Varying Number Of Clusters

2.3 Complexity Analysis of FCM Algorithm :

The time complexity of the fuzzy c mean algorithm is O(ndc2i), where empirically I grows very slowly with n,c and d.
The memory complexity of FCM is O(nd + nc), where nf is the size of data set and nc the size of U matrix.
For data sets, which cannot be loaded into memory, FCM will have disk accesses every iteration. Thus the disk input output complexity will be O(ndi) It is likely that for those data sets the U matrix cannot be kept in memory too. Thus, it will increase the disk input/output complexity further.

2.4 Complexity Analysis of MFPCM Algorithm:

The asymptotic efficiency of the algorithm has following notations:

- i number of k means passes over entire dataset.
- n number of data points.
- c number of clusters
- d number of dimensions

The time complexity of the hard c mean algorithm is O (ncdi), where empirically I grows very slowly with n, c and d. The memory complexity of MFPCM is cd I/O complexity of MFPCM is ndi

Table 4.4. Comparative Analysis of Complexities of FCM and MFPCM

Algorithm	Time complexity	Space complexity	I/O complexity
FCM	O(ndc2i),	O(nd + nc)	O(ndi)
MFPCM	O(ncdi),	cd	ndi

3. Conclusion

In dividing based grouping calculations, the quantity of definite bunch (k) should be characterized in advance. Additionally, calculations have issues like vulnerable to neighborhood optima, delicate to anomalies, memory space and obscure number of emphasis steps required to group. The time multifaceted nature of the MFPCM is O(ncdi) The memory many-sided quality of MFPCM is compact disc and the information yield many-sided quality will be O(ndi). Fluffy grouping, which constitute the most established part of delicate registering, are appropriate for taking care of the issues identified with understandability of examples, inadequate/uproarious information, blended media data and human cooperation, and can give inexact arrangements speedier. They have been for the most part utilized as a part of finding affiliation rules and utilitarian conditions and picture recovery. The time many-sided quality of the Fuzzy C Mean calculation is O(ndc2i). The memory intricacy of FCM is O(nd + nc),and the circle input yield many-sided quality will be O(ndi).

References

[1] Udehemanth.D, D.Selvathi and J.Anitha,"Effective Fuzzy Clustering Algorithm for Abnormal MR Brain Image Segmentation",Page umber 609-614, International/Advance Computing Conference (IACC 2009),IEEE,2009.
[2] SorinIstrail, "An Overview of Clustering Methods", With Applications to Bioinformatics.

[3] Wei Wang, Chunheng Wang, Xia Cui, Ai Wang, "A Clustering Algorithm Combine the FCM algorithm with Supervised Learning Normal Mixture Model", IEEE 2008.
[4]Deepak Agrawal "Web Data Clustering using FCM and Proximity Hints from Text as well as Hyperlink-structure", IEEE2008.
[5] M. Brej and M. Sonka, "Object localization and border detection criteria design in edge-based image segmentation automated learning from examples", IEEE Transactions on Medical imaging, vol. 19, pp. 973-985, 2000.
[6]Udayakumar R., Kaliyamurthie K.P., Khanaa, Thooyamani K.P., Data mining a boon: Predictive system for university topper women in academia, World Applied Sciences Journal, v-29, i-14, pp-86-90, 2014.
[7]Kaliyamurthie K.P., Parameswari D., Udayakumar R., QOS aware privacy preserving location monitoring in wireless sensor network, Indian Journal of Science and Technology, v-6, i-SUPPL5, pp-4648-4652, 2013.
[8]BrinthaRajakumari S., Nalini C., An efficient cost model for data storage with horizontal layout in the cloud, Indian Journal of Science and Technology, v-7, i-, pp-45-46, 2014.
[9]BrinthaRajakumari S., Nalini C., An efficient data mining dataset preparation using aggregation in relational database, Indian Journal of Science and Technology, v-7, i-, pp-44-46, 2014.
[10]Khanna V., Mohanta K., Saravanan T., Recovery of link quality degradation in wireless mesh networks, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4837-4843, 2013.
[11]Khanaa V., Thooyamani K.P., Udayakumar R., A secure and efficient authentication system for distributed wireless sensor network, World Applied Sciences Journal, v-29, i-14, pp-304-308, 2014.
[12]Udayakumar R., Khanaa V., Saravanan T., Saritha G., Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction, Middle - East Journal of Scientific Research, v-16, i-12, pp-1781-1785, 2013.
[13]Khanaa V., Mohanta K., Saravanan. T., Performance analysis of FTTH using GEPON in direct and external modulation, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4848-4852, 2013.
[14]Kaliyamurthie K.P., Udayakumar R., Parameswari D., Mugunthan S.N., Highly secured online voting system over network, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4831-4836, 2013.
[15]Thooyamani K.P., Khanaa V., Udayakumar R., Efficiently measuring denial of service attacks using appropriate metrics, Middle - East Journal of Scientific Research, v-20, i-12, pp-2464-2470, 2014.
[16]R.Kalaiprasath, R.Elankavi, Dr.R.Udayakumar, Cloud Information Accountability (Cia) Framework Ensuring Accountability Of Data In Cloud And Security In End To End Process In Cloud Terminology, International Journal Of Civil Engineering And Technology (Ijciet)Volume 8, Issue 4, Pp. 376–385, April 2017.
[17]R.Elankavi, R.Kalaiprasath, Dr.R.Udayakumar, A fast clustering algorithm for high-dimensional data, International Journal Of Civil Engineering And Technology (Ijciet), Volume 8, Issue 5, Pp. 1220–1227, May 2017.
[18]R. Kalaiprasath, R. Elankavi and Dr. R. Udayakumar. Cloud. Security and Compliance - A Semantic Approach in End to End Security, International Journal Of Mechanical Engineering And Technology (Ijmet), Volume 8, Issue 5, pp-987-994, May 2017.
[19]Thooyamani K.P., Khanaa V., Udayakumar R., Virtual instrumentation based process of agriculture by automation, Middle - East Journal of Scientific Research, v-20, i-12, pp-2604-2612, 2014.
[20]Udayakumar R., Thooyamani K.P., Khanaa, Random projection based data perturbation using geometric transformation, World Applied Sciences Journal, v-29, i-14, pp-19-24, 2014.
[21]Udayakumar R., Thooyamani K.P., Khanaa, Deploying site-to-site VPN connectivity: MPLS VsIPSec, World Applied Sciences Journal, v-29, i-14, pp-6-10, 2014.