AIP Conference Proceedings

The suitable distance function for fuzzy C-Means clustering

Cite as: AIP Conference Proceedings **2578**, 060006 (2022); https://doi.org/10.1063/5.0106185
Published Online: 03 November 2022

Joko Eliyanto and Sugiyarto Surono

View Online

Export Citation

ARTICLES YOU MAY BE INTERESTED IN

Proposing IoT and big data implementation model for business performance
AIP Conference Proceedings **2578**, 050002 (2022); https://doi.org/10.1063/5.0106211

Logarithmic transformation for enhancing keypoint matching of SIFT in augmented reality
AIP Conference Proceedings **2578**, 060010 (2022); https://doi.org/10.1063/5.0106221

Comparison of random forest algorithm, support vector machine and neural network for classification of student satisfaction towards higher education services
AIP Conference Proceedings **2578**, 060003 (2022); https://doi.org/10.1063/5.0106201

AIP Publishing

© 2022 Author(s).

The Suitable Distance Function for Fuzzy C-Means Clustering

Joko Eliyanto and Sugiyarto Surono [a)]

Department of Mathematics, Ahmad Dahlan University, Yogyakarta, Indonesia

[a)] Corresponding author: sugiyarto@math.uad.ac.id

Abstract. Fuzzy C-Means clustering is a form of clustering based on distance which apply the concept of fuzzy logic. The clustering process works simultaneously with the iteration process to minimize the objective function. This objective function is the summation from the multiplication of the distance between the data coordinates to the nearest cluster centroid with the degree of which the data belong to the cluster itself. Based on the objective function equation, the value of the objective function will decrease by increasing the number of iteration process. This research provide how we choose the suitable distance for Fuzzy C-Means clustering. The right distance will meet the optimization problem in the Fuzzy C-Means Clustering method and produce good cluster quality. They are Euclidean, Average, Manhattan, Chebisev, Minkowski, Minkowski-Chebisev, and Canberra distance. We use five UCI Machine Learning dataset and two random datasets. We use the Lagrange multiplier method for the optimization of this method. The result quality of the cluster measure by their accuracy, Davies Bouldin Index, purity, and adjusted rand index. The experiment shows that the Canbera distances are the best distances which provide the optimum result by producing minimum objective function 378.185. The suitable distance for the application of the Fuzzy C-Means Clustering method are Euclidean distance, Average distance, Manhattan distance, Minkowski distance, Minkowski-Chebisev distance, and Canberra distance. These six distances produce a numerical simulation that derives the objective function fairly constant. Meanwhile, the Chebisev distance shows the movement of the value of the objective function that fluctuates, so it is not in accordance with the optimization problem in the Fuzzy C Means Clustering method.

INTRODUCTIONS

The Fuzzy Clustering method is a clustering method based on the membership degree of the data [1]. This method allows the membership degree of the data can be considered into certain classification. In Fuzzy Clustering, the method which mostly used is Fuzzy C-Means [2]. The objective function of the Fuzzy C-Means is the multiplication value of the membership degree of the data in certain cluster with the square value of the distance difference between the data coordinate and cluster centroid [3]. There are several distances that can be used on Fuzzy C-Means clustering. This distance is defined as a function that meet the required criteria [4]. First, the value of distance between two coordinates is non-negative. Second, the value of the distance is zero if and only if there are two points in the same coordinate. Third, the distance between A and B or B and A is equivalent. Fourth, the distance function has to meet the concept of triangle inequality [5].

The research of the effect of the distance in clustering method has been done a lot. A few of the research conclude with the conclusion where the various distance function that being used produces an output which are not much different and there are no sign of the distance that appear dominant [6]–[8]. The effect of the variation of distance function which applied to clustering method, Euclidean distance, Manhattan / City Block distance, Chebisev / Maximum distance, and Minkowski distance, have already identified on K-Means Clustering algorithm [6], [9], [10]. In another research, Euclidean distance, Manhattan/City Block distance, Canberra distance, and Chebisev distance,

The 3rd International Conference on Engineering and Applied Sciences (The 3rd InCEAS) 2021
AIP Conf. Proc. 2578, 060006-1–060006-12; https://doi.org/10.1063/5.0106185
Published by AIP Publishing. 978-0-7354-4233-7/$30.00

are applied and evaluated on fuzzy clustering algorithm. The result of that research concluded that the result of the clustering process is very dependent on the dataset that being applied [11], [12].

Apart from the distance function that already mentioned, there are another variation of the distance function which can be applied on clustering algorithm. There are Standardized Euclidean distance, Mahalanobis distance, Cosine distance, Spearman distance, Canberra distance, Bray Curtiz distance, Average distance, Chord distance, Weighted Euclidean distance, Hausdorf distance, and Minkowski-Chebisev distance [7], [13]–[16]. The different application of a certain distance function can increase the performance quality of the clustering algorithm. For example, the combination of the Minkowski and Chebisev distance is proven to improving the algorithm of the Fuzzy C-Means Clustering [15]. The output result can be improved more by applying the reduction of PCA dimension on the high dimension data [17], [18]. Another way to improve the performance quality of the clustering algorithm is by using Average distance [19], [20].

The main objective from the Fuzzy C-Means clustering is to minimize the value of the objective function [21]. The objective function of Fuzzy C-Means clustering is a function that shows the error value. Through the optimization process, the more the number of iterations, the lower the value of the objective function. This is the basis for optimization of Fuzzy C-Means clustering. Intuitively, this method improving the result on each iteration by updating the membership degree on every dataset point. The membership degree matrix represent the membership degree of a certain data based on its distance to the existing centroid clusters [22]. This research focused on the observation to determine the suitable distance to be applied on Fuzzy C-Means clustering. We use five UCI machine learning data, which are Hill Valley dataset, Iris dataset, Seeds dataset, and Sonar dataset. We also provide two another random dataset with certain specification functioned as an addition variable to analyse. We limit our research to the development of the basic Fuzzy C-Means clustering method with the aforementioned seven distance functions.

METHODS

Fuzzy C-Means Clustering

Fuzzy C-Means (FCM) is a form of clustering method which enable one set of data to be classified into two or more clusters [20]. This method is invented by Dunn in 1973 and improved further by Bezdek in 1981. This method is commonly used to introduce certain pattern of data. This method is based on minimalization the following objective function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\|x_i - c_j\right\|^2, 1 \le m < \infty \tag{1}$$

Where m is any real number valued more than 1, u_{ij} is a membership value of x_i, on cluster j, x_i shows the data in i, c_j is a cluster centroid in j, and $\|*\|$ is a norm which determine the distance between data and cluster centroid. Fuzzy partition is determined by repeated optimization from the objective function which mentioned below with an update of matrix membership u_{ij} and cluster centroid c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}} \tag{2}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \tag{3}$$

The iteration process will stop when $max_{ij}\{|u_{ij}{}^{(k+1)} - u_{ij}{}^{(k)}|\} < \varepsilon$, where ε is a termination criteria between 0 and 1, where k is the number of iteration process. This procedure is convergent to local minimum value of J_m.

The Fuzzy C-Means Clustering algorithm is consist of the following steps [23]:

a. Initialization of $U = matrix[u_{ij}], U^{(0)}$
b. In steps of k : measure the vector of the cluster centroid $C^{(k)} = [c_j]$ with $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m}$$

c. Updating $U^{(k)}, U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$

d. If $\|U^{(k+1)} - U^{(k)}\| < \varepsilon$, then stops the iteration, if not, back to step 2.

Objective function and membership function are closely related in this method. We assume that every data coordinate already has their own value of membership degree for each existing cluster on the initial process of iteration. Then, this value will be updated over and over by using equation (2). During this process, equation (1) will also be updated until reach its minimum value. Figure 1 (a) illustrate the value distribution of the data membership in the initial of iteration process. Afterwards, the output of the iteration will be updated in every iteration (Figure 1 (b)). The iteration process will be terminated when the value of objective function reached their minimum where each data on every cluster has their ideal membership distribution which illustrated on Figure 1 (c). The ideal membership distribution of the data is occurred when every data are gathered in their closest cluster centroids.

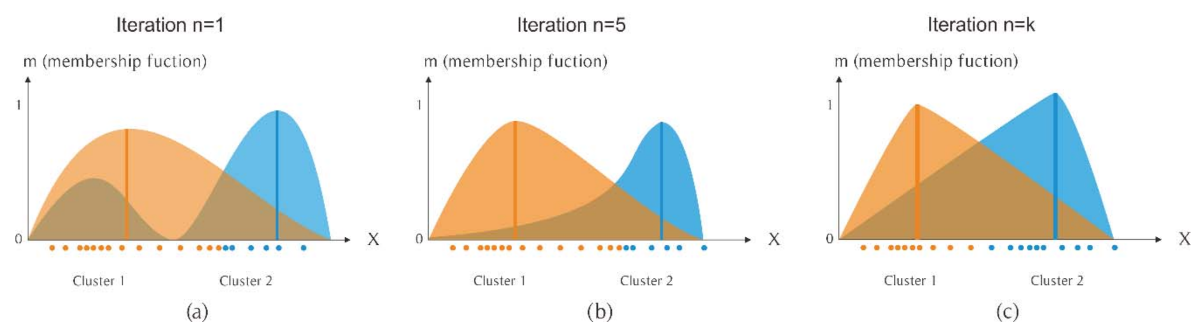

FIGURE 1. Membership Function Update During Iteration Process

Lagrange Multiplier

To maximize or minimize $f(x, y)$ against the constraint then solve the system of equations

$$\nabla f(x, y) = \lambda \nabla g(x, y) \tag{4}$$

And

$$g(x, y) = 0 \tag{5}$$

For x, y and λ. Each point x, y is a critical point for the extreme value problem with the constraint and λ is called the Lagrange multiplier.

Function $f(x,y)$ with constraint $g(x,y)=c$ will reach the minimum or maximum by form it into a Lagrange function which is defined as

$$\mathrm{L}(x,y,\lambda)=f(x,y)+\lambda(g(x,y)-c) \tag{6}$$

To get a solution to the optimization problem, the following necessary and sufficient conditions are required:

Necessary Condition:

$$\frac{\partial L}{\partial x}=0,\frac{\partial L}{\partial y}=0,\frac{\partial L}{\partial \lambda}=0 \tag{7}$$

Sufficient Condition:

- For the maximization problem, then the main determinant of the Hessian matrix is negative definite.
- For the minimization problem, then the main determinant of the Hessian matrix is positive definite.

Distance Functions

Euclidean Distance

Euclidean distance is known as the most common and standard distance function to apply for clustering method. This function for point x and y is define with this following equation [19]:

$$d_{euclidean}(x,y)=\sqrt{\sum_{k=1}^{n}(x_k-y_k)^2} \tag{8}$$

Where x_k and y_k are the value of x and y on dimension n.

Manhattan Distance

Manhattan distance or City Block distance is defined as an addition of each its attributes. Hence, for two points of data x and y on dimension n, the distance between them is define as [19]:

$$d_{manhattan}(x,y)=\sum_{k=1}^{n}|x_k-y_k| \tag{9}$$

Where x_k and y_k are the value of x and y on dimension n.

Chebisev Distance

Chebisev distance is also known as maximum distance. This distance is defined as the maximum value of distance from its attributes that exist. The distance for two points of data x and y on dimension n is define as [19]:

$$d_{chebisev}(x,y)=max_{k=1}^{n}|x_k-y_k| \tag{10}$$

Minkowski Distance

Minkowski distance is a measurement from metric distance which define as [24]:

$$d_{minkowski}(p, x, y) = \sqrt[p]{\sum_{k=1}^{n} |x_k - y_k|^p} \tag{11}$$

With p is a Minkowski parameter. On Euclidean distance $(p = 2)$, Manhattan distance $(p = 1)$, and Chebisev distance $(p \to \infty)$. The metric condition for this function will satisfied as long as $p \geq 1$.

Minkowski-Chebisev Distance

This distance function is invented by Rodrigus (2018) and Novianty (2019) uses the combination of Minkowski-Chebisev distance for Fuzzy C-Means Clustering. The combination of Minkowski and Chebisev distance with weight w1 ¬and w2 are shown on equation (7). When w1 is greater than w2, this distance function will looks similar to Manhattan distance and vice versa. The Minkowski-Chebisev distance is define as [25]:

$$d_{min-cheb}(w_1, w_2, p, x, y) = w_1 d_{minkowski}(p, x, y) + w_2 d_{chebisev}(x, y) \tag{12}$$

Canbera Distance

Canberra distance is an addition of the absolute ratio difference between two points of data [26]. The equation to measure this distance function is [27]:

$$d_{canbera}(x, y) = \sum_{k=1}^{n} \frac{|x_k - y_k|}{|x_k| + |y_k|} \tag{13}$$

This distance is very sensitive to the alteration if the two points of data are close to 0. This distance is applied because of its properties similarities with Manhattan distance.

Average Distance

Average distance is a modification of Euclidean distance. This modification is developed to improve the quality of clustering output [24]. This distance is define with this following equation [19]:

$$d_{average}(x, y) = \sqrt{\frac{1}{n} \sum_{k=1}^{n} (x_k - y_k)^2} \tag{14}$$

Datasets

This research based on five UCI Machine Learning dataset dan two random with three class. TABEL 1 illustrate the detail information of the applied dataset.

TABLE 1. Datasets

Dataset	Row	Column	Class
Random 1	150	4	3
Random 2	210	7	3
Hill-Valley	1484	8	10
Iris	208	60	2
Seeds	606	100	2
Sonar	100	2	2
Yeast	400	2	4

Cluster Evaluation

Cluster Evaluation has to be applied to find out the accuracy level of the clustering algorithm and to perceive the existing classification. This research use 3 testing methods, Accuracy test, Purity test and Davies Bouldin Index, for cluster evaluation [15][18].

Accuracy

Accuracy is calculated by adding up the number of objects that fall into the i-th cluster, where the correct one in the original class is then divided by the number of data objects. Good accuracy results if all clusters match the original class and then divided by the number of data will produce a value of 1.

Purity

Purity is used to calculate the purity of a cluster. A bad cluster has a purity value close to 0. This means that there are no cluster results that match the original class, while a good cluster has a purity value of 1 . So, it can be interpreted that the results of the cluster are in accordance with the original class.

Adjusted Rand Index

Adjusted Rand Index (ARI) is a Rand Index that has been fixed. ARI is used to measure cluster quality by comparing the predicted class with the original class. The Rand Index value is in the interval , while the ARI value is in the interval . The Rand Index value is in the interval [0,1] while the ARI value is in the interval [-1, 1] . The greater the ARI value means the better the quality of a cluster.

Davies Bouldin Index

The Davies–Bouldin index (DBI), introduced by David L. Davies and Donald W. Bouldin in 1979, is a metric for evaluating clustering algorithms. This is an internal evaluation scheme, where the validation of how well the clustering has been done is made using quantities and features inherent to the dataset. This has a drawback that a good value reported by this method does not imply the best information retrieval. Due to the way it is defined, as a function of the ratio of the within cluster scatter, to the between cluster separation, a lower value will mean that the clustering is better.

Research Method

This research is in the form of a numerical simulation. First we define Fuzzy C-Means Clustering with 7 different distances namely Euclidean, Manhattan, Chebisev, Minkowski, Minkowski-Chebisev, Canbera, Average. This method is an optimization problem. Then, we solve this problem using the Lagrange method. From the last step, we obtain the Hessian matrix to see whether the optimization problem that is constructed is a maximization or minimization problem. Then to verify it numerically, we implement Fuzzy C-Means Clustering in python programming language and then record the results. FIGURE 2 is our research flowchart.

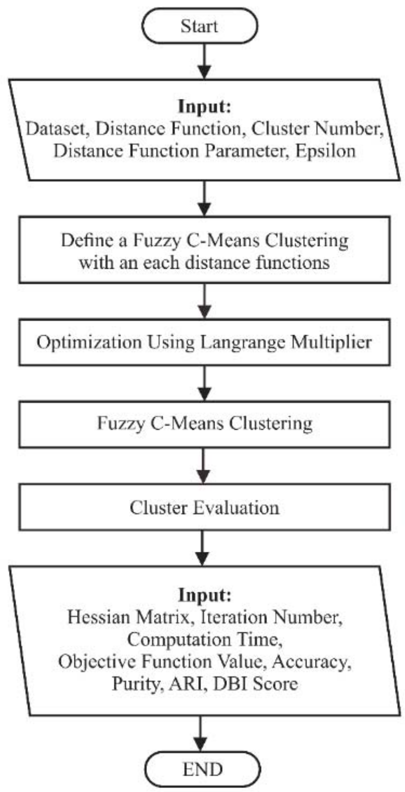

FIGURE 2. Research Flowchart

RESULT & DISCUSSION

First, we define the Fuzzy C-Means Clustering method with seven distance functions. Then we optimize using the Lagrange Mutiplier method to obtain the Hessian matrix results for each distance function presented in TABLE 2. The seven distances produce several possibilities for the Hessian Matrix, namely indefinite, semi-definite positive, or positive definite. This means that there is a high probability that these functions are minimization problems. This result will be verified by further numerical experimentation to see if the pattern of the minimization problem is seen in the iteration process of the method.

TABLE 2. Hessian Matrix of Each Distance Function

Dataset	Matrix Hessian
Euclidean	*indefinite or semi-definite positive or positive definite*
Manhattan	*indefinite or semi-definite positive or positive definite*
Minkowski	*indefinite or semi-definite positive or positive definite*
Chebisev	*positive semi definite*
Minkowski-Chebisev	*indefinite or semi-definite positive or positive definite*
Canbera	*indefinite or semi-definite positive or positive definite*
Average	*indefinite or semi-definite positive or positive definite*

The Fuzzy C-Means Clustering method with 7 kinds of distance functions is implemented in a program written in python. Parameter values which used on this research are :$p = 5, w_1 = 0.5, w_2 = 0.5, m = 2, Max_{Iter} = 100, \varepsilon = 10^{-16}$. The values of objective function is observed in every iteration process. The output comparison of the clustering process on each dataset is done by normalizing the output value of the main objective by using this following equation (15):

$$\hat{x}_i = \frac{x_i}{max(X)} \times 100\% \qquad (15)$$

FIGURE 3 illustrate the result of the cluster evaluation average output in each of the distance which applied on Fuzzy C-Means Clustering algorithm. The result shows that the clustering process output on each distance function is not significantly different. This result is consistent with [6]. Nevertheless, the Euclidean distance and Average distance are achieving a better result than the other distance function. The same result is also achieved on another research [10].

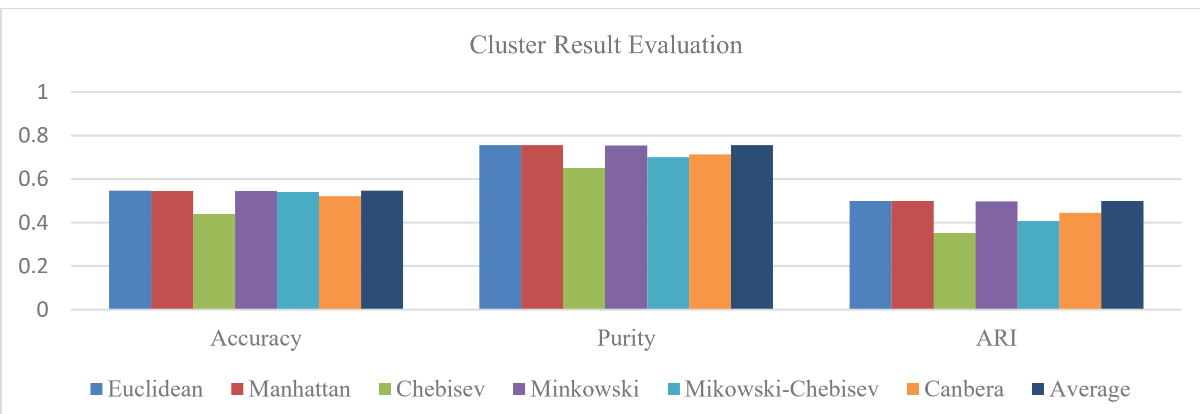

FIGURE 3. Cluster Result Evaluation

Apart from the desired clustering output, the low value of objective function which should be minimum is also the ideal result that we want to achieve. FIGURE 4 shows the average value of objective function value which achieved on each distance function for 100 times iteration. The minimum values of objective function is achieved on 3 distance functions; Euclidean distance, Manhattan distance, and Minkowski-Chebisev distance. This result concludes that the combination of Minkowski distance and Chebisev distance able to improve the output of Fuzzy C-Means Clustering algorithm. This result also mentioned on different research [15] with 0.546 as the minimum value. The value of objective function on Manhattan distance is very different than other distance function which also shown in another research [9].

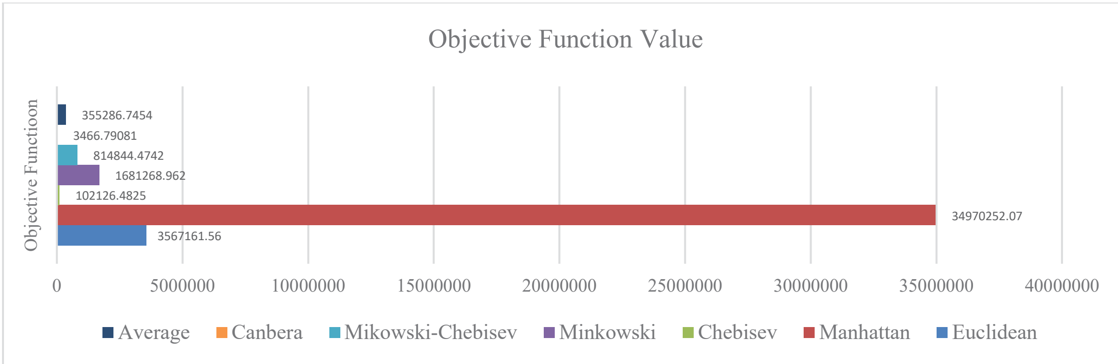

FIGURE 4. Objective Function Value

The results of the clustering evaluation of the Fuzzy C-Means method are presented in TABLE 3. To determine the best distance, scoring is performed which is shown in TABLE 4. The best distance obtained is the Canbera distance with the highest score on all sides.

TABLE 3. Cluster Evaluation

Distance Function	Iteration	Time Computation	Objective Function	Accuracy	DBI	ARI	Purity
Euclidean	40	5,468	694.389.760.824	0,611	1,146	0,307	0,641
Manhattan	72	12,854	67.247.168.149.060	0,624	1,577	0,308	0,657
Minkowski	58	16,399	154.255.495.471	0,606	1,827	0,308	0,659
Chebisev	93	12,284	11.921.246.033	0,368	1,532	0,190	0,577
Minkowski-Chebisev	63	19,923	59.104.443.850	0,481	1,548	0,276	0,642
Canbera	46	7,984	378.185	0,479	1,375	0,339	0,658
Average	58	15,712	6.943.897.619	0,601	1,854	0,309	0,660

TABLE 4. Scoring

Distance Function	Iteration	Time Computation	Objective Function	Accuracy	DBI	ARI	Purity	Total
Euclidean	7	7	2	6	7	3	2	34
Manhattan	2	4	1	7	3	4	4	25
Minkowski	5	2	3	5	2	5	6	28
Chebisev	1	5	5	1	5	1	1	19
Minkowski-Chebisev	3	1	4	3	4	2	3	20
Canbera	6	6	7	2	6	7	5	**39**
Average	4	3	6	4	1	6	7	31

FIGURE 5 shows the comparison of objective function based on iteration number. The other distance function parameter, the combination of Minkowski and Chebisev distance, does not produce the desired output, where the tendency of the objective function values are increasing on Yeast dataset, Sonar dataset, Seeds dataset, and Random 2 dataset. The most undesired result occurs in the application of Chebisev distance where the tendency of the objective function values are fluctuated. However, in different research, the usage of this distance function resulting on suitable performance when another parameters are added [13], [27]. In fact, the performance from the application of Chebisev distance still could be improved even further [9]. The results in this research are very dependent in the starting initialization value of the matrix. That's why the clustering method is also known as non-deterministic method. Nevertheless, the output of this research is considered beneficial for further analysis on how the Fuzzy C-Means Clustering method should be applied.

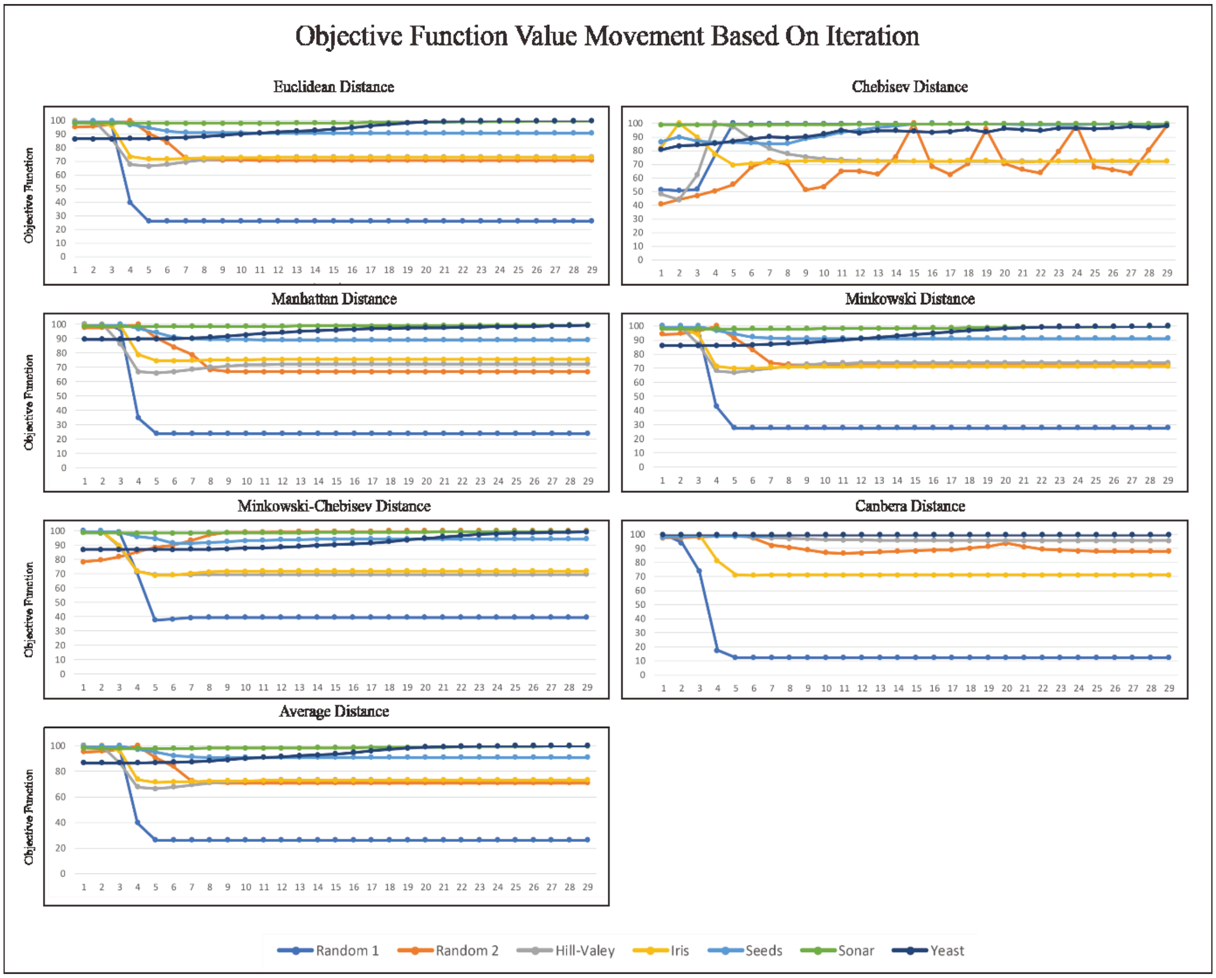

FIGURE 5. Objective Function Value Movement Based on Iteration

Good clustering results are not enough for the Fuzzy C-Means Clustering method. This method is a minimization problem, so the value of the objective function will decrease continuously for each iteration. If this last point is not met, then the modification of the distance function applied is said to be incorrect. For further research, it's essential to determine on how the termination criteria should be based on. The termination criteria in this research is based on the concept where the significant value of clustering result is constant (below the values of ε). This criteria is not exactly accurate, in fact, even when the minimum value of objective function is achieved, the objective function valued below ε are always continuously run the algorithm. On top of that, the possibility where the value of the objective function in the next iteration process got worse could be occurred.

CONCLUSION

This research concluded that the most suitable distance functions to apply for Fuzzy C-Means clustering method are Euclidean distance, Average distance, Manhattan distance, Minkowski distance, Minkowski-Chebisev distance, and Canberra distance. Furthermore, this research shows that not every distance function will satisfy the minimization problem even though the majority of the dataset that being applied are met. Based on the result of this research, it's important to visualize the tendency of the objective function values on Fuzzy C-Means clustering method. On the other hand, focusing on the result of the clustering process is also important. It aims to achieving suitable result through proper method. In the future, the research on the evaluation of the termination criteria and another different application of distance functions could be analyzed further.

REFERENCES

1 M. Zhang, W. Zhang, H. Sicotte, and P. Yang, "A New Validity Measure for a Correlation-Based Fuzzy C-Means Clustering Algorithm," pp. 3865–3868, 2009.

2 N. Grover, "A study of various fuzzy clustering algorithms," *Int. J. Eng. Res.*, 2014.

3 J. Nayak, B. Naik, D. P. Kanungo, and H. S. Behera, "ENGINEERING PHYSICS AND MATHEMATICS A hybrid elicit teaching learning based optimization with fuzzy c-means (ETLBO-FCM) algorithm for data clustering," *AIN SHAMS Eng. J.*, 2016, doi: 10.1016/j.asej.2016.01.010.

4 J. Hernadi, *Analisis Real Elementer dengan Ilustrasi Grafis & Elementer*. Erlangga, 2015.

5 G. R. Bartle and R. Sherbet, *Introduction to Real Analysis (4thed.)*. United States of America: Hamilton Printing Company, 2010.

6 P. Grabusts, "The choice of metrics for clustering algorithms," *Proc. 8th Int. Sci. Pract. Conf.*, 2011.

7 V. Kumar, J. K. Chhabrea, and D. Kumar, "Performance evaluation of distance metrics in the clustering algorithms," *INFOCOMP J. Comput. Sci.*, 2014.

8 Y. S. Thakare and S. B. Bagal, "Performance evaluation of K-means clustering algorithm with various distance metrics," *Int. J. Comput. Appl.*, 2015.

9 O. A. M. Jafar and R. Sivakumar, "Hybrid Fuzzy Data Clustering Algorithm Using Different Distance Metrics : A Comparative Study," no. 6, pp. 241–248, 2014.

10 A. Singh, A. Rana, and U. Pradesh, "K-means with Three different Distance Metrics," vol. 67, no. 10, pp. 13–17, 2013.

11 B. Charulatha, P. Rodrigues, and T. Chitralekha, "A comparative study of different distance metrics that can be used in Fuzzy Clustering Algorithms," *IJETTCS. Natl. Conf. Archit. Softw. Syst. Green Comput.*, 2013.

12 V. P. Mahatme and K. K. Bhoyar, "Impact Of Distance Metrics On The Performance Of K-Means And Fuzzy C-Means Clustering-An Approach To Assess Student's Performance In E-Learning Environment," *Int. J. Adv. Res. Comput. Sci.*, 2018.

13 S. Boddana and H. Talla, "Performance Examination of Hard Clustering Algorithm with Distance Metrics," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 2S3, pp. 172–178, 2019, doi: 10.35940/ijitee.b1045.1292s319.

14 N. Gueorguieva, I. Valova, and G. Georgiev, "ScienceDirect ScienceDirect M & MFCM : Fuzzy C-means Clustering with Mahalanobis and Minkowski Distance Metrics," *Procedia Comput. Sci.*, vol. 114, pp. 224–233, 2017, doi: 10.1016/j.procs.2017.09.064.

15 P. Noviyanti, "Fuzzy c-Means Combination of Minkowski and Chebyshev Based for Categorical Data Clustering," *Univ. Ahmad Dahlan*, 2018.

16 A. S. Shirkhorshidi, S. Aghabozorgi, and T. Y. Wah, "A Comparison Study on Similarity and Dissimilarity Measures in Clustering Continuous Data," pp. 1–20, 2015, doi: 10.1371/journal.pone.0144059.

17 J. Eliyanto, Sugiyarto, Suparman, I. Djakaria, A. H. Mustafa, and Ruhama, "Dimension Reduction Using Core and Reduct to Improve Fuzzy C-Means Clustering Performance," *Technol. Reports Kansai Univ.*, 2020.

18 S. Surono and R. D. A. Putri, "Optimization of fuzzy c-means clustering algorithm with combination of minkowski and chebyshev distance using principal component analysis," *Int. J. Fuzzy Syst.*, 2020, doi: 10.1007/s40815-020-00997-5.

19 G. Gan, C. Ma, and J. Wu, "Data clustering: theory, algorithms, and applications," *Soc. Ind. Appl. Math.*, 2020.

20 J. Han and M. Kamber, "Data Mining : Concepts and Techniques Third Edition," *Morgan Kaufmann. Amsterdam*, 2011.

21 J. C. Dunn, “A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters,” *J. Cybern.*, 1973.

22 J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum Press, 1981.

23 A. K. T. Andu and A. S. Thanamani, *Multidimensional Clustering Methods of Data Mining for Industrial Applications*. 2013.

24 T. Brunello, D. Bianchi, and E. Enrico, *Introduction To Computational Neurobiology and Clustering*. Singapore: World Scientific Publishing, 2007.

25 O. Rodrigues, “Combining minkowski and cheyshev: new distance proposal and survey of distance metrics using k-nearest neighbours classifier,” *Pattern Recognit. Lett.*, vol. 110, pp. 66–71, 2018, doi: 10.1016/j.patrec.2018.03.021.

26 B. G. N. Lance and W. T. Williams, “Computer programs for hierarchical polythetic classification (" similarity analyses "),” 1964.

27 C. Science, “Performance Evaluation of Distance Metrics in the Clustering Algorithms,” no. 1, pp. 38–51, 2014.