ISSN: 04532198
Volume 62, Issue 06, July, 2020

Dimension Reduction Using Core and Reduct to Improve Fuzzy C-Means Clustering Performance

Joko Eliyanto[1], Sugiyarto[2], Suparman[3*], Ismail Djakaria[4], Mustafa A.H. Ruhama[5]

Department Mathematics Education, Postgraduate Program, Universitas Ahmad Dahlan, Yogyakarta, Indonesia[1,2,3]
Universitas Negeri Gorontalo, Indonesia[4]
Universitas Khairun, Ternate, Indonesia[5]

Corresponding email: 3*

Abstract— Large-volume data is very difficult to find hidden patterns in the data. The complexity and computational time for analyzing large volumes of data to obtain important information are very dependent on the number of data and variables in a dataset. Big data intersects with incomplete data. This study aims to develop a method of data clustering that is sensitive to missing values in big data that is fast and efficient. This research develops data clustering using fuzzy c-means clustering methods. This method can accommodate the incompleteness of data by calculating the datum expertise in the dataset. Dimension reduction is applied to reduce dimensions in a data set while maintaining important information in the dataset. Core and Reduct which is one of the concepts in the rough set theory was chosen to reduce and leave only the core of a dataset. Core and Reduct are applied to look for core data patterns and select important variables in the data. The results showed that the application of Core and Reduct in the Fuzzy C-Means clustering could shorten the computational time and reduce the value of objective functions until the remaining 43.49%. At the same time, the quality of the clusters produced can be better with relatively unchanged purity and far better accuracy. The combined advantage of this method is that it has a better performance compared to the standard fuzzy c-means clustering.

Keywords— Core and Reduct, Dimension Reduction, Fuzzy C-Means, Rough Set.

1. Introduction

The usage of reasonable force source extended fundamentally not long after the principle enormous oil crisis in the late seventies. Around that point, fiscal issues were the hugest parts, from now on the eagerness for such methods lessened when oil costs fell. The present resurgence of eagerness for the use of an economical force source is driven by the need to diminish the high natural impact of fossil-based imperativeness structures. Useful patterns hidden in large volumes of data are very difficult to determine [1]. Statistical methods can be used to analyze data to find useful patterns [2]. Data available in various types can be represented as variables [3]. From a mathematical point of view, variables are also seen as dimensions [4]. The more data records and data variables, the more complex and longer the computational processes that must be carried out [5]. One pattern in the data is the data group. Grouping data is also called clustering. A dataset member can be a member of a data group in two ways [6]. Firstly, the data is entered into a group with certain criteria (e.g. shortest distance). The second way, the data is entered into a group by calculating the degree of membership of data to existing clusters [7]. The second way is called fuzzy clustering [8]. With the fuzzy method, the possibility of membership of data to a group can be considered. In fuzzy clustering, a popular method used is fuzzy c-means [9]. In this method, data converted into fuzzy form is the distance between the object and the cluster center given, the objective function of the fuzzy c-means is the product of the degree of data membership in a cluster with the square of the distance of the data point to the center of the cluster [10]. Clusters whose members are very similar and differences between high clusters will be obtained by minimizing the objective function. A clustering method will be very complex when the data has a lot of variables and records [11]. Dimension reduction is a way to reduce dimensions in a data set while maintaining important information in the dataset [12]. Rough set theory is a new mathematical tool for dealing with uncertainty and unclear decision systems. This theory has been successfully applied in all fields. This theory is used to identify the reduction set of all attributes of the decision system. The reduction set is used as a preprocessing technique for the classification of decision systems to issue potential patterns or association rules or knowledge through data mining techniques [13]. Core and reduct is a technique in a rough set that aims to reduce and leave only the core of a dataset [14].

Various studies have been conducted to apply and develop both fuzzy c-means clustering, dimensional reduction, and core and reduct. M. Premasundari and C. Yamini use fuzzy clustering to group crime with certain levels of crime in the United States [15]. Fuzzy c-means clustering has the main disadvantage of being able to get stuck at some local optimum. To overcome this, A. Baykasoğlu, and İ. Gölcük, F. B. Özsoydan uses the weighted superposition attraction algorithm. This method shows a significant increase in results [16]. Mohamad Faiz Dzulkalnine and Roselina Sallehuddin combine fuzzy c-means clustering, principal component analysis, and support vector machines to improve the efficiency and accuracy of fuzzy c-means clustering [17]. Other research conducted by Akbar Esmaeilzadeh Kurosh Shahriar is proposing a combination of fuzzy clustering methods based on deferential evolutionary to conduct rock mass cluster studies. This new method provides the best value for all criteria and has good stability in various criteria considered in the study [18]. Research on dimension reduction using rough sets was conducted by Shenglong Yu and Hong Zhao. They combine rough sets and laplacian score based costs sensitive to make a variable selection. The experimental results show that the proposed algorithm can improve performance to get a subset of features with low computational costs [19]. M. Sammany and T. Medhat use rough sets to reduce input to artificial neural networks. After the reduction process, an artificial neural network with a multi-layer perceptron is applied. This study increased classification accuracy [20]. G. Yan, G. Ma, and L. Zhu propose a rough set application as a front end preprocessing data for vector classifier support. Experimental results on benchmark datasets show that the proposed method can reduce computational complexity without reducing classification accuracy compared to vector classifier support without preprocessing data [21]. T. R. JeraldBeno and M. Karnan proposed a mechanism for selecting new features based on forwarding and backward fuzzy reduction. This study also presents a new entropy-based modification of the rough set-based approach. The results of this study are applied to find the reduction in rough sets to a minimum. Evaluation is carried out experimentally [21]. Differently, Mustafa Mat Deris, Norhalina Senan, Zailani Abdullah, Rabiei Mamat, and Bana Handaga developed the method of dimensional reduction using conditional entropy for incomplete information systems [22]. Research on the use of dimension reduction using the theory of cores and reducts in rough sets is still very little. This research focuses on fuzzy c-means clustering for grouping data. To improve the performance of this method, dimension reduction using cores and reducts is carried out. This technique will provide data that is the core of the initial data set. The new dataset produced has a lower dimension than the original dataset. So fuzzy c-means clustering will work on data with lower dimensions. The main problem to be solved is how to improve the performance of fuzzy c-means clustering for grouping data. The researcher wants to investigate how effective the combination of these two methods is for numerical data grouping. This study aims to develop a fuzzy c-means clustering method that provides better results while at the same time having a low load and computational time. As a method development process, researchers apply this method to the sample data that is under the desired conditions, namely in the form of sample data that has been previously classified. Data that has a group label can be viewed as an information table so that the core and reduct method can be applied.

2. Methods

This study uses a computer program simulation. Both fuzzy c-means clustering and core and reduct are arranged into an algorithm. Based on this algorithm, an m-file is compiled using MATLAB. Then, the simulation is done by using several benchmark datasets. Regarding the dataset, the dataset is organized into data in the form of information tables. Before the dataset is clustered using fuzzy c-means clustering, the data set will be reduced using core and reduct. So a new dataset with lower dimensions is generated. The reduction data are then clustered using fuzzy c-means. The resulting data clusters are then calculated for their accuracy and purity. For comparison, the value of the objective function and the computational time computation of this method are displayed. To identify the quality of this new method, the dataset is clustered using both fuzzy c-means standard and fuzzy c-means core and reduct. Objective function values, computational time, accuracy, and purity of results for the two methods are compared.

2.1 Fuzzy C-Means Clustering

Given the data $a = \{a_i\}_{i=1}^{n}$, $n > C$, n is a number of record data and C is the number of clusters at different data points in dimensions d. The fuzzy c-means objective function is to minimize (1) as follows:

$$F(U,V) = \sum_{i=1}^{n}\sum_{j}^{k} u_{ij}^{z} d_{ij} \quad (1)$$

with constraints

$$\sum_{j=1}^{k} u_{ij} = 1,\ 0 \le u_{ij} \le 1,\ 1 \le i \le n, 1 \le j \le k \quad (2)$$

where:
n: number of data
k: number of clusters
z: fuzziness index
u_{ij} : the degree of membership in the i-th data and j-th cluster
v_{ij} : center of the i-th cluster and j-category data
d_{ij} : distance function to measure the similarity between datums in a dataset

The steps for clustering with this technique are as follows:

1. Retrieve data containing random variables $X = \{x_1, x_2, x_3, \cdots x_m\}$ and $Y = \{y_1, y_2, y_3, \cdots y_n\}$ states objects and attributes. Data in the form of matrix size $n \times m$, where n is a lot of data and m is the number of data attributes.
2. Determine the number of cluster = c, maximum iteration = MaxIter, the smallest error expected = ε with the initial iteration t=1 and $F_{FCM}{}^{(0)} = 0$.
3. Determine the initial matrix $u_{ij}{}^{(0)}$ with size $c \times n$ consist of random numbers $0 \le u_{ij} \le 1$, such that the total number of new partition matrices in class is 1 according to the constraint (1).

$$U^{(0)} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ u_{c1} & u_{c2} & \cdots & u_{cn} \end{bmatrix} \quad (3)$$

where $u_{11} + u_{21} + \cdots + u_{c1} = 1$.

4. Calculate cluster center v_{ij}.

$$v_{ij} = \frac{\sum_{i=1}^{c}\left(u_{ij}^{m} \times X_{ij}\right)}{\sum_{i=1}^{c} u_{ij}^{m}} \tag{4}$$

5. Calculate objective function $F_{FCM}(U,V)$ using (1).
6. Calculating changes in the partition matrix.
7. Check the conditions for stopping, if $\left|F_{FCM}(U,V)^{t} - F_{FCM}(U,V)^{t-1}\right| < \varepsilon$ $t = \mathrm{MaxIter}$ or then the counting process is stopped. If it does not meet, the calculation continues until one of the criteria is met.

2.2 Core and Reduct

Core and reduct consist of removing excess partitions (equivalence relationships) of excessive base categories in the database such that a set of basic categories in the database is maintained [23]. This procedure makes it possible to eliminate all unnecessary data from the database and only retains portions of the entire dataset that are truly useful. In many data analysis applications, information and knowledge are stored and represented in information tables. Information tables provide an easy way to describe a set of objects up in the universe by a set of attributes up to [24]. Information tables or also called decision information systems T can be expressed as quadruple T = (U, A, C, D), where U is the set of all objects (universe) [25]. A is a variable set. If set A is then divided into conditional attribute set C and decisional attribute set D, i.e. $C \cup D = A$ and $C \cap D = \varnothing$. The information table is a table where each row represents an entity and each column represents a specific attribute variable. Entities in the U set can be classified into several different decision classes and several different conditional attributes. Table 1 is a display of the decision table. The set of universes $U = \{x_1, x_2, \cdots\}$ is a set of objects or entities, $C = \{c_1, \cdots, c_m\}$ is a set of conditional attributes, and $D = \{d_1, \cdots, d_k\}$ is a decisional attribute set. The terms f_{ij} express the jth conditional attribute of the i-th entity and express the j-th conditional attribute of the i-th entity.

Table 1. Information Table

U	**C**				**D**		
	c_1	c_2	$\cdots$	c_m	d_1	$\cdots$	d_k
x_1	f_{11}	f_{12}	$\cdots$	f_{1m}	g_{11}	$\cdots$	g_{1k}
x_2	f_{21}	f_{22}	$\cdots$	f_{2m}	g_{21}	$\cdots$	g_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	f_{n1}	f_{n2}	$\cdots$	f_{nm}	g_{n1}	$\cdots$	g_{nk}

The information table above represents all available information and knowledge [26]. In the information table, for each subset $R \subseteq A$, indiscernibility relationships are defined with

$$IND(R) = \{(x,y) \in U \times U : r \in R : r(x) = r(y)\} \tag{5}$$

In the information table T, the set of attributes $R \subseteq A$ is called a reduct, if R satisfies the following two conditions:

1) $IND(R) = IND(A);$ (6)

2) $\forall a \in R, IND\left(R-\{a\}\right) \neq IND\left(A\right).$ (7)

Condition one states that for each object pair that cannot be distinguished by a subset R, it also cannot be distinguished by A and vice versa. The second condition states that there are object pairs that cannot be distinguished by R - {a} but can be distinguished by A. This means that R is the minimum set of attributes that can maintain the indiscernibility relationship IND (A). Usually, there is more than one reduction in an information table. The set of all reductions from the information table T is denoted as RED (T).

Then, the cores of the attribute set $R \subseteq A$ are as follows:

$$CORE\left(R\right)=\bigcap RED\left(R\right) \quad (8)$$

The following algorithm for dimension reduction uses the concept of core and reduct:

Algorithm 1. *Core & Reduct*

INPUT:

Information table $T=\left(C,D\right)$ where $C=\left\{C_1,C_2,\cdots,C_n\right\}$ is set attributes contribution; $D=\left\{D_1\right\}$ is decision set attribute. Suppose A is the total amount of attributes. Then RED [k] is reduct, COR [m] is the core, and min_reduct [j] is minimum reduction.

k=0; m=0;

PROCESS BEGIN:

For (i = 1; i <= N; i ++)

1. Set the C_i attribute to ‘a 'and delete the C_i column from T and call the remaining table Ti. Save the Ti condition attribute in array b [i].
2. If for the same set of condition attribute values, the value of the decision attribute differs from the T_i table is not consistent or T_i table is consistent. If T_i is consistent then

```
{
        k=k+1
        RED[k]=b[i];
}
else
{
        m=m+1
                COR[m]=a;
        }
end if;
      end for;
```

3. If Core (m) = 1 then

   for (j = 1; j <= N; j ++)

Check the consistency of the COR core attributes [m] with each other of the T_i condition conditions left one by one to find min. reduction.

min_reduct [j] = {COR [m], Cj}

end for;

end if;

4. If Core (m)> 1 then the set of all cores form a min reduct

PROCES END

OUTPUT:

Reduct set RED [k] , core COR [m] and min_reduct [j].

2.3 Cluster Evaluation

To determine the accuracy of the clustering algorithms and class labels available, a cluster evaluation was carried out. In this study, cluster evaluation was carried out two tests, namely: purity and accuracy.

1. Purity

Purity is used to calculate the purity of a cluster. Purity calculation for each cluster obtained is done by taking the most objects entered in the C-cluster where $1 < i < C$ and C' are the original h-class with $1 < h < C'$. As for the overall purity of the C cluster, it is done by adding up each purity in the C cluster and dividing it by the number of objects defined as follows:

$$\text{purity}(P,C) = \frac{1}{n}\sum_i^c \max_{i\le h\le c'} \left|P_i \cap C_h\right| \tag{9}$$

where $P = \{P_1, P_2, P_3, \cdots P_c\}$ is the cluster set and $C = \{C_1, C_2, C_3, \cdots, C_c\}$ is the original class set. Poor clustering 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 value of purity 1. This means that the cluster results are in accordance with the original class.

2. Accuracy

Accuracy is calculated by adding up the number of objects included in the i-cluster, where $1 < i < C$ the exact class is then divided by the number of data objects. Accuracy is defined as follows:

$$r = \frac{\sum_{i=1}^{c} a_1}{n} \tag{10}$$

Information:

a_i : a number of objects in the i-cluster that correspond to the original class.

n: number of n objects.

Good accuracy results if all clusters match the original class and then divided by the amount of data will produce a value of 1.

2.4 Dataset

The dataset used in this study is a dataset taken from the UCI Machine Learning website [28]. Table 2 presents a brief description of the dataset used in this study.

Table 2. Dataset

Dataset	Object	Variables	Class
Iris	150	4	3
Seeds	210	7	3
Yeast	1484	8	10
Sonar	208	60	2
Hill Valley	606	100	2

All of the above datasets are numeric types.

2.5 Proposed Fuzzy C-Means Clustering Method

To improve the performance of fuzzy c-means clustering, the method of reducing core and reduct dimensions in the dataset is applied first. Lower dimensions will reduce the computational burden of the fuzzy c-means algorithm. Then, the purity and accuracy test is performed to see the quality of the clustering results. Figure 1 presents the scheme for the proposed new method. Whereas Algorithm 2 presents a fuzzy c-means clustering algorithm with core and reduct.

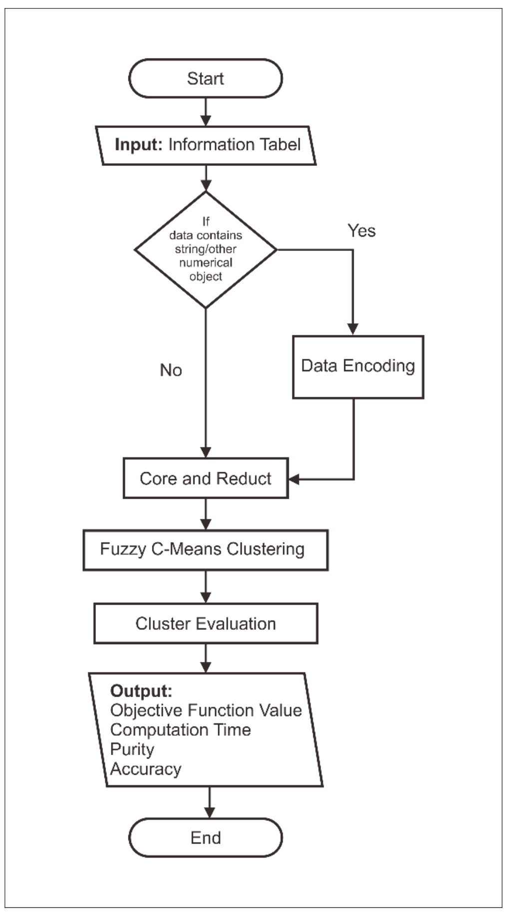

Figure 1. Fuzzy C-Means Core & Reduct Clustering Scheme

Algorithm 2. ***Fuzzy C-Means Core & Reduct Clustering***

INPUT:
Data input is in the form of variables $X = \{x_1, x_2, \cdots, x_n\}$ $Y = \{y_1, y_2, \cdots, y_m\}$ and expresses objects and attributes. Data is a $n \times m$ matrix, where n is a lot of data and m is the number of data attributes.

PROCESS BEGIN
1. If the dataset is not numeric data, then encoding data is done, if not then proceed to the next process.
2. Apply the core and reduct method so that the number of variables $Y = \{y_1, y_2, \cdots, y_m\}$ will be a number of new variables $Y = \{y_1, y_2, \cdots, y_p\}$ $p \le m$,
3. Applying the fuzzy c-means clustering method so that data clusters are obtained.
4. Cluster evaluation.

PROCESS END

OUTPUT:
The value of the objective function, computational time, purity, and accuracy.

3. Result and Discussion

Core and reduct can reduce the dimensions of data with high dimensions. Table 3 presents the results of the core and reduct reduction in each dataset.

Table 3. Core and reduct result

Dataset	Before Reduction	After Reduction
Iris	4	4
Seeds	7	2
Yeast	8	6
Sonar	60	2
Hill Valley	100	2

Based on the above results, the core and reduct work much better on high-dimensional data. In low-dimensional data, it tends to be difficult to determine the core of the data set. The main purpose of fuzzy c-means clustering is to achieve the smallest objective function values. The smaller the value of the objective function, the fuzzy c-means clustering has better results. That means the data groups are separated more clearly. Figure 2 presents a comparison between the objective function values of fuzzy c-means and fuzzy c-means core and reduct. For the five datasets that were simulated, the core and reduct succeeded in decreasing the value of the objective function until the remaining 43.49% was in the sonar dataset.

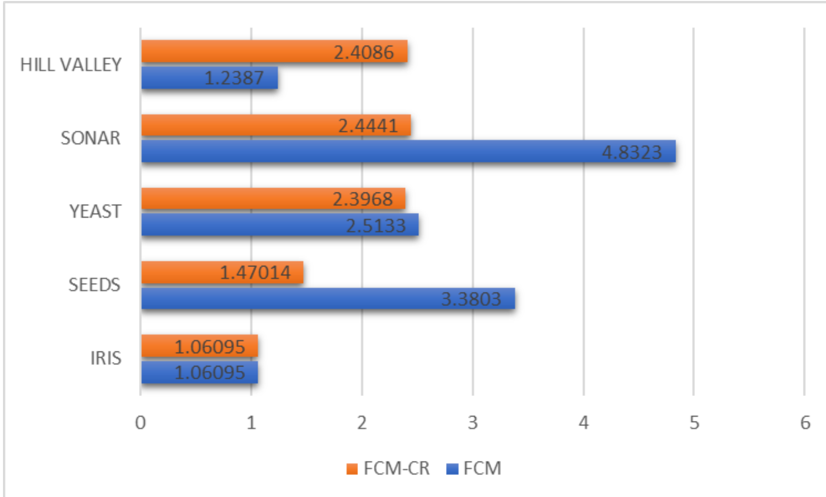

Figure 2. Objective Function Value

When a computational process works at a lower dimension, the computational load is reduced. This will certainly affect the computing time. Figure 3 presents computational time comparisons for these two methods. Research has shown that the core and reduct relative relatively increase the computational speed of fuzzy c-means clustering. So the computing time was successfully reduced. The results obtained do not feel significant because the core and reduct only reduce the number of attributes, but the number of data records remains.

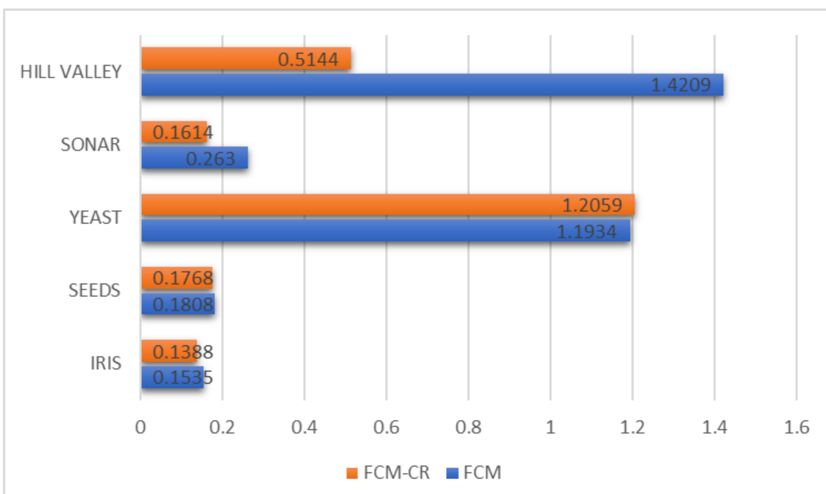

Figure 3. Computation Time

A good cluster result has purity approaching one, whereas a bad cluster is an opposite. Figure 4 shows that fuzzy c-means produces clusters with varying purity. However, the application of core and reduct did not have a significant impact. Purity value can still be maintained.

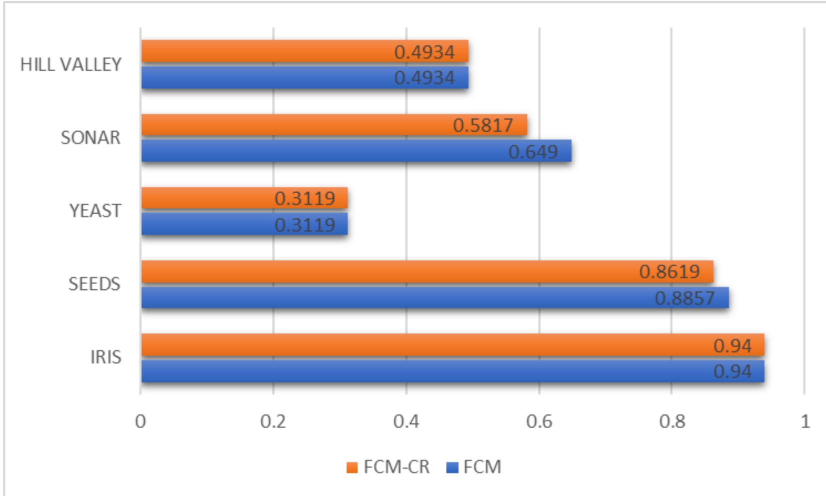

Figure 4. Purity

The most interesting result of this newly developed method is that core and reduct can improve accuracy very significantly. Figure 5 shows that in the seeds dataset, the accuracy value can jump from 0.0142 to 0.8169. Improved accuracy can also be seen in other datasets except for the hill-valley dataset which produces fixed accuracy. However, in other aspects, the fuzzy c-means clustering core and reduct still has better results.

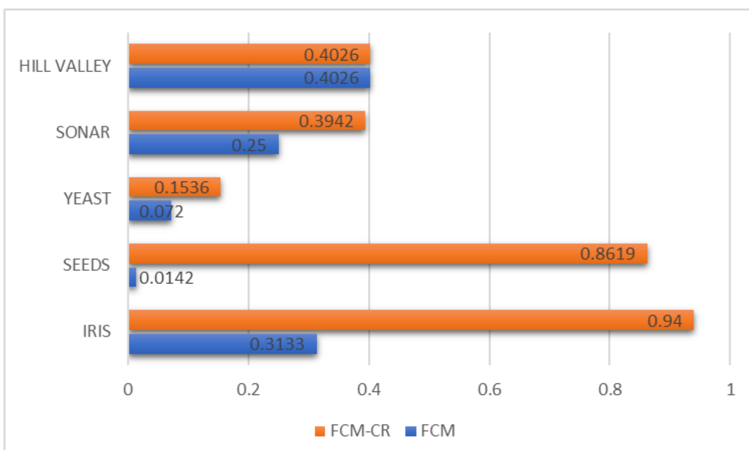

Figure 5. Accuracy

Some of the above results are in line with the results of previous studies [16-21]. R. Zhao, L. Gu, and X. Zhu also conducted research in the same field as this research. They produce a combination of fuzzy c-means clustering with rough set feature selection, namely reduct that can increase accuracy by an average of 1% [28]. The core process that was added to this study guarantees that the result of dimension reduction is only the core of the dataset. The results of this study achieve the best results at a certain random center value to start the fuzzy c-means algorithm. Optimization methods for finding and selecting initial values for the best cluster centers can be developed to further improve the performance of fuzzy c-means clustering. If this clustering can still be run on a low dataset dimension, better results with a lighter load can be obtained.

4. Conclusion

Core and reduct work better on larger dimension datasets. Core and reduct are proven to improve the performance of fuzzy c-means clustering. This dimension reduction presents the essence of the dataset as a result of its reduction. Although the dimensions are much lower, to be used as further analysis (in this case clustering) can be relied upon. The results of this fuzzy c-means clustering are very dependent on the selection of the initial dataset points and the initial center point of the cluster. A new optimization method can be developed to overcome this.

5. References

[1] R. Vashist, "An Algorithm for Finding the Reduct and Core of the Consistent Dataset", International Conference on Computational Intelligence and Communication Networks (CICN), 2015, pp. 741-745.

[2] U. Sivarajah, M.M. Kamal, Z. Irani, and V. Weerakkody, "Critical analysis of Big Data challenges and analytical methods", Journal of Business Research, vol. 70, 2017, pp. 263-286.

[3] S. P. Kaur, "Variables in research", Indian Journal of Research and Reports in Medical Sciences, vol. 3, no. 4, 2013, pp. 36-38.

[4] B. Marr, "Big data in practice: how 45 successful companies used big data analytics to deliver extraordinary results", John Wiley and Sons, Chichester, 2006.

[5] C.C. Drovandi, C. Holmes, J. M. McGree, K. Mengersen, S. Richardson, and E. G. Ryan, "Principles of experimental design for big data analysis", Statistical science: a review journal of the Institute of Mathematical Statistics, vol. 32, no. 3, 2017, pp. 385.

[6] J. Han, M. Kamber, and J. Pei, "Data mining concepts and techniques third edition", Morgan Kaufmann, Amsterdam, 2011.

[7] A. Gosain, and S. Dahiya, "Performance analysis of various fuzzy clustering algorithms: a review", Procedia Computer Science, vol. 79, 2016, pp. 100-111.

[8] D. J. Bora and A.K. Gupta, "A comparative study between fuzzy clustering algorithm and hard clustering algorithm", 2014, arXiv preprint arXiv:1404.6059.

[9] N. Grover, "A study of various fuzzy clustering algorithms", International Journal of Engineering Research, vol. 3, no. 3, 2014, pp. 177-181.

[10] J. Nayak, B. Naik, D.P. Kanungo, and H.S. Behera, "A hybrid elicit teaching learning based optimization with fuzzy c-means (ETLBO-FCM) algorithm for data clustering", Ain Shams Engineering Journal, vol. 9, no. 3, 2018, pp. 379-393.

[11] Z. Shi and L.S.C. Pun-Cheng, "Spatiotemporal data clustering: a survey of methods", ISPRS international journal of geo-information, vol. 8, no.3, 2019, pp. 112.

[12] A. Sarveniazi, "An actual survey of dimensionality reduction", American Journal of Computational Mathematics, vol. 4, 2014, pp. 55-72.

[13] K. Thangavel and A. Pethalakshmi, “Dimensionality reduction based on rough set theory: A review”, Applied Soft Computing, vo. 9, no. 1, 2009, pp. 1-12.

[14] B. R. A. Moreira, R. S. Viana, L. A. M, Lisboa, P. R. M. Lopes, P. A. M. Figuieiredo, S. B. Ramos, C. S. B. Bonini, V. D. R. Trindade, M. G. O. Andreade, and A. May, "Classifying hybrids of energy cane for production of bioethanol and cogeneration of biomass-based electricity by principal component analysis-linked fuzzy c-means clustering algorithm", Journal of Agricultural Science, vol. 11, no. 14, 2019, pp. 246-253.

[15] M. Premasundari and C. Yamini, “A violent crime analysis using fuzzy c-means clustering approach”, ICTACT Journal on Soft Computing, vol. 9, 2019, pp. 1939-1944.

[16] A. Baykasoğlu, İ. Gölcük, and F. B. Özsoydan, “Improving fuzzy c-means clustering via quantum-enhanced weighted superposition attraction algorithm”, Hacettepe Journal of Mathematics and Statistics, vol. 48, 2018, pp. 859-882.

[17] M. F. Dzulkalnine, R. Sallehuddin, “Missing data imputation with fuzzy feature selection for diabetes dataset”, SN Applied Sciences, 2019, https://doi.org/10.1007/s42452-019-0383-x.

[18] A. Esmaeilzadeh and K. Shahriar, “Optimized Fuzzy Cmeans – Fuzzy Covariance – Fuzzy Maximum Likelihood Estimation Clustering Method Based on Deferential Evolutionary Optimization Algorithm for Identification of Rock Mass Discontinuities Sets”, Periodica Polytechnica Civil Engineering, vol. 63, no. 2, 2019, pp. 674-686.

[19] S. Yu, H. Zhao, “Rough sets and Laplacian score based cost-sensitive feature selection”, PLoS ONE, 2018, https://doi.org/10.1371/journal.pone.0197564.

[20] M. Sammany and T. Medhat, “Dimensionality reduction using rough set approach for two neural networks-based applications”, 2007, https://link.springer.com/chapter/10.1007/978-3-540-73451-2_67.

[21] G. Yan, G. Ma, and L. Zhu, “Data Dimension Reduction Using Rough Sets for Support Vector Classifier”, 2006, https://link.springer.com/chapter/10.1007/11795131_67.

[22] T.R. Jerald Beno, and M. Karnan, “Dimensionality reduction: rough set based feature reduction”. International Journal of Scientific and Research Publications, vol. 2 no. 9, 2012, pp.1-6.

[23] Z. Pawlak, “Rough Sets, Theoretical Aspects of Reasoning about Data”, Academic Publishers, Warswa, 1991.

[24] Z. Pawlak, “Rough Sets”, International Journal of Information and Computer Science, 1982, http://dx.doi.org/10.1007/BF01001956.

[25] D. Xue and Y. Q. Chen, “Scientific computing with MATLAB”, CRC Press, 2018.

[26] Y.Y. Yao and Y. Zhao, “Discernibility matrix simplification for constructing attribute reducts”, Information Sciences, vol. 179, no. 5, 2009, pp.867-882.

[27] D. Dua and C. Graff, “UCI Machine Learning Repository”, University of California, School of Information and Computer Science, 2020, http://archive.ics.uci.edu/ml.

[28] R. Zhao, L. Gu, and X. Zhu, "Combining Fuzzy C-Means Clustering with Fuzzy Rough Feature Selection", Applied Sciences, vol. 9, 2019 pp.679.