HASIL CEK_Clustering with Principal Component Analysis and Fuzzy Subtractive Clustering Using Membership Function Exponential and Hamming Distance

by Sugiyarto Sugiyarto

Submission date: 02-Nov-2020 10:37AM (UTC+0700)
Submission ID: 1433356738
File name: g_Using_Membership_Function_Exponential_and_Hamming_Distance.pdf (433.87K)
Word count: 3597
Character count: 17328

Clustering with Principal Component Analysis and Fuzzy Subtractive Clustering Using Membership Function Exponential and Hamming Distance

Abstract. The problem of dimension reduction in multivariate data is how to obtain a smaller number of variables but still be able to maintain most of the information contained in the data. One method that can be used is principal component analysis (PCA). Principal component analysis (PCA) is a technique used to reduce the dimensions of data consisting of several dependent variables while maintaining the variance in the data. PCA can be used to stabilize measurements in statistical analysis, one of which is cluster analysis. Fuzzy clustering is a method of grouping based on membership values that include fuzzy sets as a basis for weighting for grouping. One method of fuzzy clustering is Fuzzy Subtractive Clustering (FSC). The method used in this study is PCA and FSC. The purpose of this study is to compare the most optimal cluster results using PCAFSC and FSC methods. The results obtained indicate that the clustering using the PCAFSC method is better than the FSC method.

1. Introduction

The problem of dimensional reduction in multivariate data is how to get a smaller number of variables but still be able to retain most of the information contained in the data. One method that can be used is principal component analysis (PCA) [1]. This method is designed to reduce the spatial dimension of the data to find and interpret dependencies among variables or to help stabilize measurements in statistical analyzes such as regression analysis or cluster analysis [2].

Other According to [3], principal component analysis (PCA) is a technique used to reduce the dimensions of data consisting of several dependent variables while maintaining the variance in the data. This can be done by transforming the data into new variables (main components) that are not correlated to form a new coordinate system with maximum variance. Therefore, PCA can be used to reduce the dimensions of data without significantly reducing the characteristics of the data [4].

Cluster analysis is a technique used to look for patterns in a data set by grouping observations into groups. The aim is to find the optimal grouping in which the observations of objects in each cluster are similar, but the clusters are different from each other [5]. The basic principle in cluster analysis is to classify objects (observations) in a cluster that have very large similarities with other objects in the same cluster (similarity) but are not very similar to other objects in different clusters (dissimilarity). This means that a good cluster will have high homogeneity (similarity) between members in one cluster (within-cluster) and high heterogeneity (difference) between clusters (between-cluster) [6].

There are several methods that can be used in the grouping, one of which is the most frequently used is Fuzzy C-Means (FCM). In this method, the number of groups is known and uses a group membership matrix that was determined previously. However, the initial group membership matrix was initialized randomly so that the FCM method was complicated. Another grouping method is Fuzzy Subtractive Clustering. This method is used when the number of groups is unknown. Fuzzy Subtractive Clustering has more consistent results and better speed compared to FCM [7]. The concept of the FSC method is to determine the area in a variable that has the highest density of the surrounding points. The point with the most number of neighbors will be selected as the center of the cluster and its

density will be reduced. Then the selection of another point with the most neighbors will be made to become the center of another cluster. This step will be repeated until all points are tested [8].

There are several studies on clustering, including research conducted by [9] used the FSC to classify polymers according to their similarity to various chemicals in the vapor phase. The polymer that determines the center of the cluster is selected to create the sensor array. The virtual sensor array is determined based on the polymer selected by FSC using a polymer-coated SAW oscillator as a chemical sensor. Research conducted by [10] on fuzzy clustering combined with PCA to reduce seventeen variables by producing three main components. Furthermore, the three main components will be grouped using the FCM as the basis for evaluating driver service levels. The PCA and FCM algorithms not only extend the function of grouping based on principal component analysis, but also increase the algorithm speed.

In addition, research conducted by [11] used the PCA method to reduce data dimensions, where the results of the main components will be used as input to the FCM method. The distance method used is the euclidean distance. Another study was also conducted by [12] using Fuzzy C-Means and Fuzzy Subtractive Clustering to compare the planning for the placement of the UMTS900 Node B network on the existing BTS of one of the operators in the Malang area using the Fuzzy C-Means and Fuzzy Subtractive Clustering methods. The results obtained are Fuzzy Subtractive Clustering which results in a more even placement of Node B in each of the 5 districts in Malang City compared to Fuzzy C-Means. Based on some of the studies above, that the PCA method can optimize clustering performance and the FSC method is better than FCM, so this study will use the PCA and FSC methods for clustering using exponential membership functions and hamming distance.

2. Literature Reviews

2.1. Principal Component Analysis (PCA)

Suppose that the random vector m is known, namely $X_1, X_2, \dots, X_m$. The requirements for the main components of the variables $X_1, X_2, \dots, X_m$ are the vectors of $PC_1, PC_2, \dots, PC_l$ which are obtained with the following condition $l < m$, $Var(PC_1) > Var(PC_2) > \cdots > Var(PC_l)$, and $Cov(PC_i, PC_j) = 0$, where $i \neq j$. The principal component depends only on the covariance matrix $\boldsymbol{S}$ of $X_1, X_2, \dots, X_m$. In general, the main component is obtained from the form of a linear combination, where the definition of a linear combination is as follows:

Definition [13] A vector $\boldsymbol{w}$ is said to be a linear combination of vectors $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$ if there are scalars $c_1, c_2, \dots, c_r$ such that

$$\mathbf{w} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_r\mathbf{v_r} \quad (1)$$

Therefore, the form of the main components is as follows:

$$PC_1 = \sum_{j=1}^{m} a_{j1}X_j = a_{11}X_1 + \cdots + a_{m1}X_m$$
$$\vdots \quad (2)$$
$$PC_m = \sum_{j=1}^{p} a_{jm}X_j = a_{1m}X_1 + \cdots + a_{mm}X_m$$

where

$$\boldsymbol{a} = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{mm} \end{bmatrix}$$

and

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

To determine the number of main components, it can be done in 2 ways, namely:
First, through proportion of total variance (PTV)

$$PTV = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \cdots + \lambda_p} \times 100\%,\ k = 1,2,\ldots,p \quad (3)$$

If a large proportion of the total population of variants for m, say 80 to 90%, can be said to be with one, two or three components, then that component can replace the original m variable without losing much information [14].
Second, through the Cumlative Proportion of Variance (PKV)

$$PKV = \frac{\sum_{j=1}^{k} \lambda_j}{\lambda_1 + \lambda_2 + \cdots + \lambda_p} \times 100\% \quad (4)$$

If the majority of the total variance (70 - 80%), for large m can be connected by the first, second, or third principal components, then this principal component can interpret m the original variable without losing much information [15].

2.2. Fuzzy Set

Definition [16] If X is a collection of objects denoted by x, then the fuzzy set U in X is expressed as the set of consecutive pairs

$$U = \{(x, \mu_U(x)) \mid x \in X\}$$

where $\mu_U(x)$ s the degree of membership x in the fuzzy set U at interval [0,1].

2.3. Fuzzy Subtractive Clustering

Fuzzy clustering is a method of grouping based on membership values that includes fuzzy sets as a weighting basis for grouping. Each data is given a value for the probability of being able to join each existing group, which means that the data is not absolute or firm to be a member of one group only, but also has a value for the probability of being a member of another group with the largest degree of membership showing a high tendency for data to be be a member of a particular group [11].

One of the fuzzy clustering methods is Fuzzy Subtractive Clustering (FSC), where the number of clusters to be formed is unknown. The basic concept of this method is to determine each point of the data which has a high density of the surrounding points. The point with the most number of neighbors will be used as the center of the cluster. Then the density of the point that will be used as the center of the cluster will be reduced and the algorithm will select another point that has the most neighbors to become the center of another cluster. This is done until all points are tested.

3. Methods

The research method used is clustering using Principal Component Analysis (PCA) and Fuzzy Subtractive Clustering (FSC) by using hamming distance and exponential membership function. Output of the PCA will be used as input for FSC, then, the quality of the clusters formed will be seen using the Partition Coefficient. Below will be one by one regarding the methods used in this research:

3.1. Principal Component Analysis Steps

The steps of the Principal Component Analysis are as follow:

- Data Normalization

  The data normalization used in this study is Softmax normalization, defined as follows [18]:

$$x'_{ij} = \frac{1}{1 + e^{-\left(\frac{x_{ij} - \mu_j}{\sigma_j}\right)}} \tag{5}$$

- Covariance Matrix

  The sample covariance for the m and n variables is

$$s_{mn} = \frac{1}{k-1} \sum_{j=1}^{k} (x_{jm} - \bar{x}_m)(x_{jn} - \bar{x}_n) \tag{6}$$

  The covariance matrix (S) is as follows:

$$\boldsymbol{S} = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix} \tag{7}$$

- Eigen Value and Eigen Vector

  The eigen values (λ) of the matrix S measuring $n \times n$ can be found using equation (8) as follows:

$$det(S - \lambda I)x = 0 \tag{8}$$

  The eigen vector (a) corresponding to λ can be found using equation (9):

$$det(S - \lambda I)x = 0 \tag{9}$$

- Determine the Number of Main Components

  The number of major components can be determined using equations (3) and (4).

- Compiling the Main Components

  In general, the main components are obtained from the form of a linear combination using equation (2).

3.2. Fuzzy Subtractive Clustering Steps

The steps of the Fuzzy Subtractive Clustering algorithm are as follow:

- Determine the value of the parameter, i.e. r (radius), q (squash factor), accept ratio, rejecet ratio.
- Converts data to fuzzy numbers using the exponential membership function [19]:

$$\mu(x) = \begin{cases} 1 & x \le a \\ \dfrac{e^{-\left(\frac{x-a}{b-a}\right)} - e^{-s}}{1 - e^{-s}} & a \le x \le b \\ 0 & x \ge b \end{cases} \tag{10}$$

- Determine the potential of each data point $D_i; i = 1,2,3, \dots, n$ with the following steps:

  First, Calculate the distance of each data against μ_{A_j} using the hamming distance [20]:

$$Dist_{ij} = \left(\frac{|\mu_A(u_i) - \mu_B(u_i)|}{r} \right) \tag{11}$$

  where $\mu_{A_j} = \mu_{B_{ij}}; j = 1,2, \dots, m; i = 1,2, \dots, n$

  Second, determine the initial potential of each data point using :

$$D_i = \sum_{k=1}^{n} e^{-4\left(\sum_{j=1}^{m} Dist_{ij}^2(x_i)\right)} \quad (12)$$

- Search for data points with the greatest potential value:

  $M = max[D_i | i = 1,2, \ldots,]$; for the first iteration.

  $Z = max[D_i | i = 1,2, \ldots,]$; for the second, third iteration an so on.

- Calculate the ratio (R) of a candidate center cluster using

$$R = \frac{Z}{M} \quad (13)$$

  In the first iteration, the value Z = M.

- Check the eligibility of prospective cluster centers to cluster centers by using 3 conditions as follows:

  a) Condition 1: if the ratio > accept ratio, then the prospective cluster center can be accepted as a new cluster center.

  b) Condition 2: if the reject ratio < ratio ≤ accept ratio is checked the feasibility of the prospective cluster center can be accepted as a new cluster center, if not then the iteration is stopped because there are no more data considered to be prospective cluster centers. The procedure in condition 2 is as follows:

$$Md = -1$$

  For $k = 1,2, \ldots, p$, where p = number of clusters

$$Sd_k = \sum_{j=1}^{m} \left(\frac{V_j - C_{kj}}{r}\right)^2 \quad (14)$$

  If $(Md < 0)$ or $(Sd_k < Md)$, then $Md = Sd_k$,

$$Mds = \sqrt{Md};$$

  where Mds is the closest distance from the prospective cluster data center to the cluster center. If (ratio + Mds) ≥ 1; prospective cluster centers are accepted as new cluster centers. Meanwhile, if (ratio + Mds) < 1 then prospective cluster centers are not accepted and will not be reconsidered as new cluster centers (the potential data is set zero) centers are not accepted and will not be reconsidered as new cluster centers (the potential data is set to zero).

  c) Condition 3: if the ratio is reject ratio, then there are no more data points that will be considered to be prospective cluster centers and the iteration is stopped.

- If a new cluster center is obtained, a reduction in the potential data around the previous cluster center is performed using

$$D_i^t = D_i^{t-1} - D_{c_{ki}} \quad (15)$$

  where

$$D_{c_{ki}} = Z * e^{-4\left[\sum_{j=1}^{m}\left(\frac{C_{kj} - x_{ij}}{r * q}\right)^2\right]} \quad (16)$$

- Returns the cluster center to the original data using

$$x = (a - b)\ln(\mu - \mu e^{-s} + e^{-s}) + a \quad (17)$$

- Calculate the degree value of membership using

$$\mu_{k_i} = e^{-\sum_{j=1}^{m}\left(\frac{x_{ij}-c_{kj}}{\sqrt{2}\sigma_j}\right)} \quad (18)$$

where

$$\sigma_j = \frac{r * \left(X_{max_j} - X_{min_j}\right)}{\sqrt{8}} \quad (19)$$

- Calculating the validity index using the Partition Coefficient [21]

$$PC = \frac{1}{N}\left(\sum_{i=1}^{N}\sum_{j=1}^{K}\mu_{ij}^2\right) \quad (20)$$

where N is the number of research objects, K is the number of clusters, and μ_{ij} is the value of membership of the i-th object with the center of the j-th group.

4. Result and Discussion

The PCA calculation process will produce output in the form of the main component, where the main component will be used as input in the FSC calculation process. The data used in this study is hypertension data. The amount of data is 500 data with 5 variables. The variables used are age (X_1), sex (X_2), systolic pressure (X_3), diastolic pressure (X_4), and body weight (X_5). Data processing was performed using Jupyter Notebook Software with the Python programming language. Then, the cluster results from the FSC with the PCA input will be compared with the cluster results from the FSC only.

4.1. Fuzzy Subtractive Clustering with Input Principal Component Analysis

The clustering process starts from PCA calculations by normalizing the data using (1) so that the data is in a certain range. Before determining the main component, each variable must be related to each other so that the covariance matrix ($\boldsymbol{S}$) will be searched first. From the covariance matrix, a PCA analysis will be carried out by looking at the eigen values of each variable. The new variable (principal component) that is formed is based on the proportion of the total variance and the proportion of cumulates. The results of the calculation of eigen values and variance can be seen in Table 1.

Table 1. Eigen value and total variance.

Component	Eigen Value (λ)	Variance (%)	Kumulatif (%)
1	0.0689	30.45	30.45
2	0.0569	25.11	55.56
3	0.0459	20.30	75.86
4	0.0374	16.51	92.37
5	0.0173	7.62	100

In Table I, there are 3 new variables or main components that represent the analyzed variables. The ability of each component to represent the variables analyzed is indicated by the amount of variance described, which is called the eigenvalue. The value of the eigen values shows the contribution of the main components to the variance of all the original variables analyzed. In addition, Table 1 shows that the PKV value ≥ 70% is 75.86%. This means that the number of main components formed is 3 components. The amount of this variance also determines the principal component number to be selected based on the total proportion of the variance or eigen values. The eigen values graph can be seen in Figure 1.

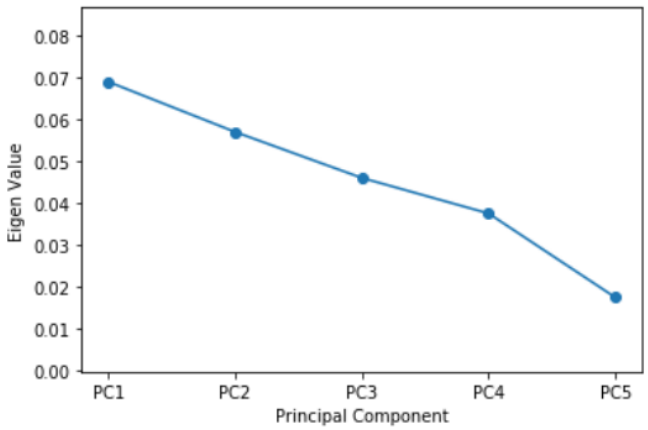

Figure 1. Eigen value.

Then, based on the previously obtained eigen vectors, it will show which variables are included in the 3 main components produced. The eigen vectors obtained from the 3 largest eigenvalues are the loading/coefficient values of the three main components formed. The loading value / coefficient of the 3 main components, namely:

Table 2. Loading/Principal Component Coefficient.

Variable	Component		
	1	2	3
1	-0.4380	**0.5087**	0.3117
2	-0.1385	-0.1556	**0.8373**
3	**0.7084**	0.6652	0.1942
4	**-0.5217**	0.5187	-0.2555
5	-0.0663	-0,0801	**0.3142**

Table 2, shows the three main components that are formed, the largest loading / coefficient value (in bold) will be chosen where this value is able to explain the variables that affect hypertension. The three new variables (main components) that were formed can be seen in Table 3.

Table 3. Summary of PCA.

Component	Variable	Loading/ coefficient	The variants describe
PC_1	X_3	0.7048	30.45 %
	X_4	-0.5317	
PC_2	X_1	0.5078	25.11 %
PC_3	X_2	0.8373	20.30 %
	X_5	0.3142	

Based on Table III, PC_1 consists of variables X_3 dan X_4 with a variance of 30.45 %, PC_2 consists of variable X_1 with a variance of 25.11 % dan PC_3 consists of variables X_2 dan X_5 with a variance of 20.30%.

Then the output from the PCA will be used as input to the FSC. The results of the comparison of the Partition Coefficient value and the number of iterations between the PCAFSC and FSC methods are shown in Table 4.

Table 4. Comparison of the results PCAFSC and FSC.

Number of clusters	PCAFSC			FSC		
	r	Partition coefficient	Number of iterations	r	Partition coefficient	Number of iterations
2	0.5	0.5874	3	0.72	0.5369	3
3	0.55	0.5301	4	0.79	0.4801	9
4	0.69	0.5974	5	0.97	0.5472	11

Based on Table IV, it is obtained that the largest Partition Coefficient value is found in the number of clusters 4, for each method of 0.5974 and 0.5472. The cluster center using PCAFSC is

$$C_{0.5} = \begin{bmatrix} 0.2250 & 0.3864 & 0.2464 \\ 0.3623 & 0.2964 & 0.6703 \\ 0.5189 & 0.5110 & 0.1633 \\ 0.2492 & 0.8169 & 0.2562 \end{bmatrix}$$

Meanwhile, the cluster center for the FSC method is as follows

$$C_{0.72} = \begin{bmatrix} 0.4340 & 0 & 0.4609 & 0.2862 & 0.4018 \\ 0.4634 & 1 & 0.4609 & 0.2862 & 0.6501 \\ 0.4196 & 1 & 0.4609 & 0.2862 & 0.1468 \\ 0.3919 & 0 & 0.5144 & 0.1597 & 1 \end{bmatrix}$$

5. Conclusion

Based on the PCA calculation results obtained 3 new variables (main components) of the 5 variables studied. The main components that are formed are PC_1, PC_2, and PC_3. The first main component (PC_1) consists of variables X_3 and X_4 with a total variance of 30.45%. The second main component (PC_2) consists of the variable X_1 with a total variance of 25.11%. The third main component (PC_3) consists of variables X_2 and X_5 with a total variance of 20.30%. The three main components formed resulted in a cumulative variance proportion of 75.86%, which means that the factors that influence hypertension can be explained only by the 3 new variables that are formed. For FSC calculations, the optimal number of clusters using the PCAFSC and FSC methods is 4 clusters, where the PC value for each method is 0.5974 with 4 iterations and 0.5473 with 11 iterations. So, it can be concluded that the clustering using the PCAFSC method is better than the FSC method.

References

[1] Mattjik A A and Sumertajaya I M 2011 *Sidik Peubah Ganda* (Kampus IPB Darmaga: IPB Press)

[2] Timm N H 2000 *Applied Multivariate Analysis* (New York: Springer)

[3] Jollife I T 2002 *Principal Component Analysis* (New York: Springer)

[4] Dafitri H, Asih M S and Astuti R I 2019 *Media Interaktif Pengenalan Angka Dengan Jari Tangan Menggunakan Metode PCA* (Journal Of Information System)

[5] Rencher A C 2002 *Methods Of Multivariate Analysis* (USA: Wiley Sons)

[6] Rahmawati, Wartono and Putri A Y 2019 *Pengklasteran Tingkat Pendidikan Pegawai Menggunakan Metode Fuzzy Subtractive Clustering (Studi Kasus: Badan Kepegawaian Dan Pelatihan Daerah Provinsi Riau* (Journal Of Mathematical Science And Statistics)

[7] Syahra Y, Syahril M and Yusnidah 2018 *Implementasi Data Mining Dengan Menggunakan Algoritma FSC Dalam Pengelompokkan Nilai Untuk Menentukan Minat Belajar Siswa SMP*

Primbana Medan (SAINTIKOM Journal)
[8] Kusumadewi S and Purnomo H 2010 *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan* (Yogyakarta: Graha Ilmu)
[9] Singh T S, Verma P and Yadava R D S 2017 *Fuzzy Subtractive Clustering For Polymer Data Mining For SAW Sensor Array Based Electronic Nose*
[10] Chen H and Li C 2019 *Comprehensive Service Level Analysis Of Online Taxi Drivers Based On Fuzzy Clustering Combined With Principal Component Analysis* (ICHSSD)
[11] Wang S, Wang J and Han Y 2019 *Pattern Matching Strategy Based On Optimized Fuzzy Clustering Algorithm For Industrial Internet*
[12] Anggraini S, Kusumawardani M and Darmono H 2017 *Perbandingan Penempatan Node B UMTS900 pada BTS Existing Menggunakan Fuzzy C-Means dan Fuzzy Subtractive Clustering* (Journal of JARTEL)
[13] Anton H and Rorres C 2005 *Elementary Linear Algebra* John (Wiley Sons)
[14] Johnson R A and Wichern D W 2002 *Applied Multivariate Statitical Analysis(*USA: Prentice Hall)
[15] Gasperz V 1995 *Teknik Analisis Dalam Penelitian Percobaan* (Bandung: Tarsito)
[16] Zimmerman H 1996 *Fuzzy Set Theory And Its Applications* (London: Kluwer Academic)
[17] Jang J S R, Sun C T and Mizutani E 1997 *Neuro Fuzzy And Soft Computing: A Computational Approach To Learning And Machine Intelligence* (USA: Prentice Hall)
[18] Priddy K L and Keller P E 2005 *Artificial Neural Networks* (USA: SPIE Press)
[19] Mahajan S and Gupta S K 2019 *On Fully Intuitionistic Fuzzy Multiobjective Transportation Problem Using Different Membership Functions* (Springer)
[20] Rezaei K and Rezaei H 2019 *New Distance And Similarity Measures For Hesitant Fuzzy Soft Sets* (IJFS)
[21] Bezdek J C 1981 *Pattern Recognition With Fuzzy Objective Algorthm* (New York: Plenum Press)

HASIL CEK_Clustering with Principal Component Analysis and Fuzzy Subtractive Clustering Using Membership Function Exponential and Hamming Distance

ORIGINALITY REPORT

18%	4%	17%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	P Laras Sakti Cahyaningrum, Nursyiva Irsalinda. "Indonesian community welfare levels clustering using the fuzzy subtractive clustering (FCM) method", Journal of Physics: Conference Series, 2019 Publication	7%
2	Submitted to Universitas Brawijaya Student Paper	3%
3	"The International Conference on ASEAN 2019", Walter de Gruyter GmbH, 2019 Publication	1%
4	Luz Stella Robles Pedrozo, M. Rodriguez-Artacho. "A cluster-based analisys to diagnose students' learning achievements", 2013 IEEE Global Engineering Education Conference (EDUCON), 2013 Publication	1%
5	Submitted to Mankato State University Student Paper	1%

6	IN Yulita, RR Julviar, A Triwahyuni, T Widiastuti. "Multichannel Electroencephalography-based Emotion Recognition Using Machine Learning", Journal of Physics: Conference Series, 2019 Publication	1%
7	link.springer.com Internet Source	1%
8	www.ukessays.com Internet Source	<1%
9	E Z Khulaidah, N Irsalinda. "FCM using squared euclidean distance for e-commerce classification in Indonesia", Journal of Physics: Conference Series, 2020 Publication	<1%
10	Kalevi Tervo, Lauri Palmroth, Vesa Hölttä, Aki Putkonen. "Improving Operator Skills with Productivity Model Feedback", IFAC Proceedings Volumes, 2008 Publication	<1%
11	"The Data-Driven Fuzzy System with Fuzzy Subtractive Clustering for Time Series Modeling", International Journal of Innovative Technology and Exploring Engineering, 2020 Publication	<1%
12	Y Arkeman, R A Rizkyanti, E Hambali. "Determination of Indonesian palm-oil-based	<1%

bioenergy sustainability indicators using fuzzy inference system", IOP Conference Series: Earth and Environmental Science, 2017
Publication

13	iopscience.iop.org Internet Source	<1%
14	www.saycocorporativo.com Internet Source	<1%
15	Sheng Wang, Jinkuan Wang, Yinghua Han. "Pattern Matching Strategy Based on Optimized Fuzzy Clustering Algorithm for Industrial Internet", 2019 Chinese Control Conference (CCC), 2019 Publication	<1%
16	Chengju Gong, Yi Su, Weiwei Liu, Yuzhen Hu, Ying Zhou. "The Distance Induced OWA Operator with Application to Multi-criteria Group Decision Making", International Journal of Fuzzy Systems, 2020 Publication	<1%
17	Atiqah Mohd Khuzi Noor Anis, V S Asirvadam, Lila Iznita Izhar. "Modeling Cardiac Signal as tomographic mapping", 2010 International Conference on Intelligent and Advanced Systems, 2010 Publication	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		