

## Taxonomy of Artist and Art Works Using Hybrid TF-IDF Fuzzy C-Means Clustering

<sup>1</sup>Bambang Nurdewanto, <sup>2</sup>Elta Sonalitha, <sup>3</sup>Anis Zubair, <sup>4</sup>Salnan Ratih  
Asriningtias, <sup>5</sup>Irfan Mujahidin

<sup>1,3,5</sup> Information Systems Faculty of Information Technology, Merdeka Malang University, <sup>2</sup>Electrical Engineering, Merdeka Malang University, <sup>4</sup>Information Technology, Brawijaya University  
Malang City, Indonesia  
elta.sonalitha@unmer.ac.id

### Abstract

*The Indonesian government has just promulgated a Law on the promotion of culture (Law No.5 2017). The Government, through the Tourism Office, determines the method of classification of cultural data using the Taxonomy method. The purpose of this study is taxonomy or mapping of big data art activists in the city of Malang, East Java, Indonesia based on the expertise of each person, so that it will facilitate the search for data for reference decision making. This research tests the calculation based on word linguistics and multi tagging from the data that the artist fills in online instruments. This study proposes the TF-IDF Fuzzy C-Means hybrid as a method of resolving these problems. TF-IDF is used as a feature extraction while Fuzzy C-Means as a clustering method. To find out the performance of the proposed method, this study uses the Variant cluster (V) technique. Based on the research analysis, the level value of  $V = 0.0000163$  is getting smaller. This shows that all cluster variants are getting better.*

**Keywords:** Taxonomy of Artist, TF-IDF, Fuzzy C-Means Clustering.

### 1. Introduction

The Government of Indonesia has just promulgated a Law on the promotion of culture Law No. 5 of 2017. To implement the law, cities in Indonesia are required to carry out protection by means of inventory, security, rescue, publication, and maintenance in their cities. One of the supports in realizing cultural promotion based on the Act is in the form of research publications[1]. This research can be in the form of collection, classification, organizing or data management in the form of information systems.

The Government, through the Department of Tourism determines the method of classification of cultural data using the Taxonomy method, where a cultural product can be categorized into several domains. For example, the cloth "batik kawung" entered in the domain of fashion, philosophy, and fine arts. Another example is that an artist has some expertise in music, dance, fine arts and others[2][3]. This research focuses on one of the objects of cultural advancement, namely Art. Art is the cultural heritage or creativity of new creation in the form of individual, communal or collective artistic expression that is manifested in various forms of activities and or mediums.

The sociological integration conditions of Malang in sociology are salad bowl and melting pot. Cultural conditions in the city of Malang based on the characteristics of sociological theory indicate that the arts in Malang tend to salad bowl. This condition is a process of cultural assimilation that is fused but does not lose its identity in the fusion. Evidently the existence of art galleries or traditional artists with the flow of East Java, Mataraman, and branches of art from other regions are also developing together[4][5]. This causes diversity and tendencies to influence each other, so that Malang city is rich in resources, ideas and works. If diversity and wealth are not well identified and structured, this will be the difference. Comparability will prevent economic and tourism effects from

optimizing. In addition, the main problem is how to map and analyze data that can provide a complete picture of the demography of art as a branch of culture in the city of Malang, East Java[6][7].

This research proposes the hybrid method of TF-IDF Text mining and Fuzzy C-Means Clustering as a means of decision support system for the government to make a policy. This decision support system is in the form of an art taxonomy application. This hybrid method was chosen because it can improve the accuracy of data calculations so that it is easy to extract information using patterns that are important to pull from data in a large database[8][9]. The analysis of this research can be used for data collection of artists as a basis for mapping the wealth of art and culture.

One of the determinants of the classification of art expertise depends on the answers that are filled in the instrument regarding the suitability of the art field with the expertise possessed by each artist. One art activist can have a lot of expertise in the arts, for example the performer master's gamelan musical instruments, as a puppeteer, and also masters the classical dance "wayang orang". These multi-talented people will be classified taxonomically and the results will be ranked using Fuzzy Clustering to determine the weight of expertise from the most mastered until it can only be done[10][11]. Research on the classification of the arts based on the expertise of artists has not yet been developed. Therefore, this classification is needed for the renewal of research. Fuzzy clustering was chosen as a method because it has the advantage of being simpler and easier to implement. This research includes exploratory research and experimental software engineering logic is *Fuzzy* used to enhance the accuracy of the linguistic subjectivity of artist data. While Fuzzy Clustering is used as a research analysis which includes the comparison of the results of the classification of experts with the application of taxonomic expert recommendations[12].

## 2. Data Set

This study uses the data of the actors of art obtained in the city area of Malang, East Java, Indonesia. This data was obtained using a questionnaire technique. This research tests the calculation based on word linguistics and multi tagging from the data that the artist fills in online instruments[13][14]. The data analysis in this research were 29 artist data shown in Table 1.

This research also uses several variables that are used as keywords in the process *term weighting* and to calculate the weight of fuzzy clustering (explained in the next chapter). These variables are shown in Table 2.

**Table 1.**  
**Record of Artist Data in Malang**

No	Name	Description (In Bahasa)
1	Katijan	Provincial Education Office East Java # 1980 Batu City Art Ambassador # Sapto Darmo 2014 # trainer # Karawitan Sapto Darmo 1987, Larasati, Arjuno Middle School, Revelation Culture Trainers of Batu City Art Ambassadors Active as artists and trainers in Batu City to Jombok, Laras Jati Margi, Ngantang, Regency. Karawitan Gagrag Metaraman; Karawitan Gagrag Surakarta Trainers; Ketoprak; Ludruk; Campursari Jaranan; Niyaga; Tayub Irianan Jaranan / Bantengan; Malang Wayang Wong; Painful Performers in training,
2	Mirza Baihaqie	secretary # secretary 2016 # Cempoko Wahyu Budoyo # 2017 Rogo wilis, cempoko wahyu culture APEKSI Batu City Youth activists in the Bantengan tradition; Jaranan Kidal Gambuh / Pawang / Dukun Peluh Gambuh in the show
3	Karin Mirda Yuwana	The best PSP and FLS2N photographer #Kota Batu # 2015 FKT best presenter # East Java # 2011 Batu City Dance Ambassador to South Korea Batu City Dance Ambassador to Malaysia Indonesian art ambassador to Taipei, Australia, Beijing Maheswari Dance Studio Saru Taru Maheswari Batu City Dance Ambassador to South Korea Duta Batu City Dance to Malaysia Indonesian art ambassadors to Taipei, Australia, Beijing Balinese Dance; Kedungmonggo

		Mask Dance; Remo Surabaya Dance / Jombang Style / Sawunggaling Style; Beskalan Dance; Malang Traditional Dance, Mataraman, Choreographed Cosmetology Dance; Dancers: Artists in the field of traditional arts
4	Alvianitaa	Matta fair Malaysia # 2016 2019 dance masterpieces There is no Maheswari dance Sanggr, Malang dance company Dancers Can go abroad only from Balinese dance; Jaipong dance; Kedungmonggo mask dance; Remo Surabaya dance / Jombang style / Sawunggaling style; Gandrung dance; Beskalan Jaranan dance; Reog; Singo Ulung clothes and art accessories; creator's make-up; coach; choreographer; fashion; stage; penguins; dancer; doers do not have
5	Yusak Santoso	best theater director student east java PentasTunggal "Visit of the Old Lady" in 2002 10 "Los Bagados De Los Pencos in 2003 5 Best Non Ranking Ranking PSP Madiun" Monument of Justice "in 2005 5 Best Non PSP Ranking Director Madiun "Monument of Justice" in 2005 Second Place National Theater Festival Love Festival "Semar Sues Semar Ruwat" Year 2007 5 Best Presenters Non Ranking Fragments of East Java History "Coup of Brahmins" 2008 Drama Dance Director "Sandya Nuswantara" Anjungan Java eastern 2016 TMII Champion 1 PSP Level Batu City "Srikandi Edian" 2016 Director "Joko Tarub Single Parents" Performing Together World Theater Day at the MbatuAJi Arts Building 2017 Director of Drama Dance Duta Batu Art "Semar Mendem" Batu East Java TMII 2018 Theater Pandu # Guidance Counselor # 1997 - Now Accurate Cultural Institution # Secretary # 2018 - Now Theater Pandu Guidance Abstinence Prostitutes themselves in the applause of Theater / Dr spectators ama Musical / Parody; Wayang Wong; Tayub Accompaniment Jaranan / Bantengan Literary Director; Manuscript Writers; Coaches; Stage Performances; Film or Soap Opera; Ketoprak; Ludruk; Kentrung Bantengan; Opera; Jaranan; Performers Videography, Photography, Multimedia Editor
...	...	...
29	Suwandi	there has been no stage from event to turonggo seto member sanduk besta mulyo sanduk besta mulyo members involved in the sanduk city rock festival event dance reog sanduk jaranan accompaniment accompaniment reog karawitan jaranan bantengan gagrag surakarta dancers perpetrators mastering several other traditional arts that have been cultivated such as arts Reog and Jaranan

**Table 2.**  
**Variables as a Keywords**

No	Keywords
1	campursari
2	puppeteer
3	gambuh
4	network
5	musical instrument
6	ketoprak
7	ludruk
8	painting
9	music
10	opera
11	statue
12	poetry
13	literature
14	dance
15	theater
16	puppet

### 3. Hybrid TF-IDF Fuzzy C-Means Clustering

This study proposes the TF-IDF Clustering Fuzzy C-Means hybrid as a method. TF-IDF is one of the methods of unsupervised learning feature selection. This method is used when data feature extraction, while Fuzzy C-Means Clustering is used as a taxonomic method. In addition, fuzzy was chosen because of its easy and simple implementation[15][16]. The following describes the stages of research using a hybrid TF-IDF Fuzzy C-Means Clustering.

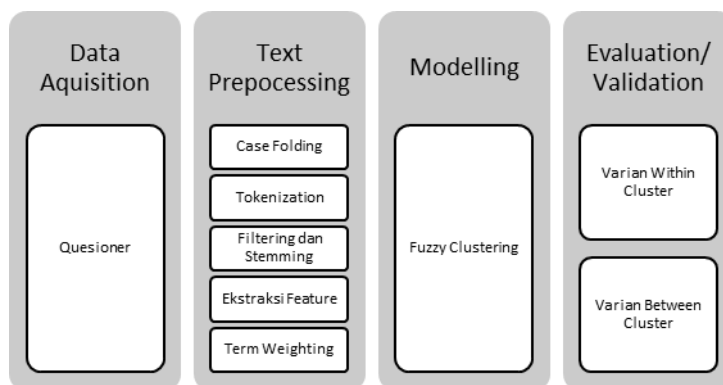


Figure 1: Flow chart of the TF-IDF Fuzzy C-Means Clustering method

This section is in the form of data collection. Data obtained by capturing artist data through a web-based online questioner. Data in the form of text or paragraph data obtained from the questionnaire form.

Text mining is an intensive knowledge process where users interact and work with a group of documents using several analysis tools. Text mining aims to get useful information from data sources in the form of documents consisting of unstructured text through identification, and exploration of a pattern[17][18]. The stage in text mining is *text preprocessing*, namely changing unstructured data into structured data. The rare steps in preprocessing text include:

- 1) *Case folding*: removes characters other than letters and changes all letters to lowercase letters
- 2) *Tokenizing*: cutting a sentence into words of its constituent words
- 3) *Filtering dan Stemming*: Taking important words from the results of tokenizing. This step can be done with 2 techniques namely stop list (discarding words that are not important) and word list (save important words). Stemming is a search for basic words from the results of filtering by removing affixes.
- 4) *Term Weighting*

Term Weighting aims to determine the value or weight in a term based on the level of interest in the document. Term Weighting calculation uses the Term Frequency-Inverse Term Frequency (TF-IDF) method. TF is the frequency of occurrence of a term in the document. IDF is the frequency of occurrence of a term in all documents. The more frequently the term appears in a document, the greater the TF value and the smaller IDF value[19]. The equation for calculating the TF-IDF value is as follows.

$$w_{ij} = tf_{ij} \times idf_j = tf_{ij} \times \log \left( \frac{N}{df_j} \right)$$

Where,  $w_{ij}$  = weight of the  $j_{th}$  term against document  $i$ ;  $tf_{ij}$  = number of occurrences of  $j$  term in document  $i$ ;  $N$  = number of documents;  $df_j$  = number of documents containing term  $j$ .

Modeling in this study uses Fuzzy clustering which is one technique for determining data clusters in which the existence of each data point in a cluster is determined by the

degree of membership. Determination of data clusters is based on the form *Euclidean Distance* to measure the proximity between data. The concept of fuzzy clustering is to decide the cluster center in progress which is the normal area for each cluster center. Within the beginning condition, the cluster center esteem is still wrong, monotonous advancements are made to the cluster center and the degree of enrollment in each information point so that the cluster center moves to the proper position[20]. This emphasis is based on the minimization of objective capacities that depict the removal from the information point to the center of the cluster weighted by the degree of enrollment of the information point. The output results are in the form of a row of cluster centers and some degree of membership for each data.

The testing process is carried out using cluster variance. The equation for calculating cluster variance is as follows. Where,  $v_c^2$  = variance in cluster  $c$ ;  $c = 1 \dots k$ ,  $k$  is the number of clusters;  $n_c$  = amount of data in cluster  $c$ ;  $d_i$  = data  $i$  on a cluster;  $\bar{d}_i$  = average of data in a cluster.

$$v_c^2 = \frac{1}{n_c - 1} \sum_{i=1}^{n_c} (d_i - \bar{d}_i)^2$$

There are two types of cluster variance namely *Varian within cluster* ( $V_w$ ) and *Varian between cluster* ( $V_b$ ), as follows.

#### *Varian within cluster* ( $V_w$ )

This variant is used to see the results of the variance of data distribution in a cluster. The smaller the value of  $V_w$ , the better the cluster. The equation for calculating  $V_w$  is as follows, where  $N$  is the whole of all information,  $k$  is the number of clusters, whereas  $n_i$  is the number of information individuals in cluster  $i$ .

$$v_w = \frac{1}{N - k} \sum_{i=1}^k (n_i - 1) v_i^2$$

#### *Varian between cluster* ( $V_b$ )

This variant is used to see the results of the variance of data spread between clusters. The greater the value of  $V_b$ , the better the cluster. The equation for calculating  $V_b$  is as takes after, where  $k$  is the number of clusters, speaks to the information to  $j$  in a cluster to  $i$ , and is the average of.

$$v_b = \frac{1}{k - 1} \sum_{i=1}^k n_i (d_{ij} - \bar{d})^2$$

### 3. Result and Discussion

In this segment will clarify the comes about of a few tests ranging from preprocessing namely text mining to evaluation.

#### A. Text Preprocessing "Text Mining"

At this stage the data obtained using a web-based questionnaire will be processed first using Text Mining. The initial stage of the text mining process is case folding cases, which removes characters other than letters and changes all letters to lowercase as shown in Table 3. The next stage is Tokenizing. This stage is the stage of cutting sentences into words according to the number of the constituent words. The following shows the results of the Tokenizing process in Table 4.

**Table 4.**  
**Result of Tokenizing Process in Text Mining**

No.	Name	Description
1	Karin Mirda Yuwana	psp psp flsn best Batu city best presenter fkt east java city ambassador for stone city south korea ambassador for city of stone malaysia art ambassador for taipei australia beijing dance studio maheswari sanggar taru maheswari dance ambassador for stone south korea ambassador for dance city for stone malaysia city ambassador for malaysia arts ambassador for indonesia taipei australia

		beijing dance studio maheswari taru maheswari dance ambassador for Batu city south korea ambassador for dance city for stone malaysia beijing bali dance mask dance kedungmonggo remo dance surabayan jombangan style sawunggaling dance beskalan dance tradition malangan mataraman creations make up choreographers dancers performers expert performers in traditional arts
2	Alvianitaa	The dancer country dance dance bali dance jaipong dance mask dance kedungmonggo remo dance surabayan style jombangan style sawunggaling dance gandrung dance beskalan jaranan reog singo ulung costume accessory cosmetology creator dancer mask kedungmonggo dance remodeling surabayan style jombangan style sawunggaling gandrung dance beskalan jaranan reog singo ulung costume accessory cosmetology creator dance choreographer kedungmonggo dance remodeling surabayan style jombangan style sawunggaling gandrung dance beskalan jaranan reog singo ulung dance accessory pengrias actors dancers
3	Yusak Santoso	best director theater student east java single stage visit old lady best presenter non ranking best presenter non rank psp madiun monument of justice best director non rank psp madiun monument justice fair champion ii theater festival homeland national love semar contugant semar ruwat best presenter non ranking fragments of history east java coup brahmana director dance drama sandya nuswantara anjungan Timur dance east java tmii champion city level stone srikandi edian director joko tarub single parents performing with the world theater theater mbatuaji dance director semar drama dance mendem ambassador art city stone jatrical east tmii theater guide builder now cultural institutions get along well theater secretary guide guide abstinence prostitute applause theater audience musical drama parody opera film soap opera puppet ketoprak ludruk kentrung bullengan jaranan tayub accompaniment jarengan bantengan literary director script writer script layout trainer the practitioners of videography photography multimedia editors
...	...	...
29	Suwandi	turonggo event seto members sanduk besta mulyo members sanduk besta mulyo members are involved in the city sanduk festival festival sanduk jaranan dance reog accompaniment jaranan bantengan accompaniment reog karawitan gagrag surakarta dancers perpetrators mastering the traditional arts of reog jaranan art

Next is the Filtering and Stemming process. This process is carried out after the word cutting results are obtained as shown in Table 4. So that in the filtering and stemming stages a basic word search of the filtering results will be obtained by removing the affixes. The results of the process are shown in Table 5.

Furthermore, the final process of Text Mining is processing data with Term Weighting or weighting with the technique Term Frequency - Inverse Document Frequency (TF-IDF). TF-IDF is an algorithm for analyzing the relationship between a phrase / sentence and a collection of documents. In this study, this technique was chosen because this technique can calculate the weight of each word most commonly used in information retrieval. Before the term weighting process is done, it is necessary to determine the keywords first. The determination of keywords is adjusted to the number of works of art in the city of Malang as shown in Table 6. Furthermore, the TF-IDF process can be carried out so as to produce a number of data in the form of keywords appearing in each art actor, as shown in Table 6.

At this stage, fuzzy C-Means clustering has begun to play a role. Fuzzy C-Means clustering at this stage is to group the results of the normalized weight calculation into 16 variables (according to the number of arts in Malang). Data clustered at this stage are the results of Term Weighting, the data shown in Table 7. While the results of clustering using fuzzy C-Means Clustering are shown in Table 8. At this stage a clustering test was conducted 10 times to get the right results. Tests carried out using an initial cluster of 2 clusters and an error below  $e < 0.001$ .

**Table 5**  
**The results of the Filtering and Stemming Process in Text Mining**

No.	Name	Description
1	Karin Mirda Yuwana	ksp psp flsn good stone town serving good fkt east java city ambassador for stone city south korea ambassador for stone city malaysia art ambassador for indonesia taipei australia beijing maheswari dance studio saru taru maheswari dance ambassador for stone city south korea ambassador for rock city dance malaysia ambassador for art city malaysia art ambassador taipei australia beijing dance studio maheswari taru maheswari dance ambassador for stone city south korea ambassador for dance city stone malaysia ambassador for art indonesia malaysia taipei australia beijing bal dance mask mask kedungmonggo remo dance surabayan jombangan style sawunggaling dance dance tradition malangan mataraman creations make up choreographers dance behavior experts in traditional arts
2	Alvianitaa	matta fair malaysia festival sanggr dance maheswari dance malang dance company neger dance dance dance balai dance jaipong dance mask dance kedungmonggo remo dance surabayan jombangan style sawunggaling dance style gandrung dance jaranan reog singo hand-made costume accessory make-up dance dance practice choreographer kedungmonggo dance remo dance style surabayan jombangan style sawunggaling gandrung dance jaranan reog singo hand-made costume accessory make-up dance dance practice choreographer fashion dance layout dance behavior
3	Yusak Santoso	good director theater teach east java pentastunggal visit old lady good food non rank psp banyuwangi los bagados de los pencos fast food good non rank psp madiun monument justice justice good director non rank psp madiun monument justice champion ii theater festival love the national level mar marat mar mar ruwat good food non rank fragments history east java coup brahmana director dance drama sand nuswantara jung east java tmii champion psp city level Batu srikandi edian director joko tarub single parents performing together world theater nian mbatuaji drama director dance dance mar mendem ambassador art city Batu jung jawa east tmii pandu bina theater now cultural institutions get along well theater secretary pandu bina abstinence prostitute applause watch theater drama musical parody opera film puppet soap opera ketoprak ludruk kentrung bantengan jarengan tayub accompaniment bullengan bantengan literary director script manuscript practice play theater drama musical parody opera film puppet soap opera ketoprak ludruk kentrung bullengan jaray tayub accompaniment bullengan bullengan literary director script manuscript practice play stage drama video parody opera cinema film editor
...	...	...
29	Suwandi	stage event turonggo seto members sanduk besta mulyo members sanduk besta mulyo members libat event festival city sanduk batu dance sanduk jaranan reog ireng jaranan iring reog karawitan gagrag Surakarta dance behavior art tradition reog jaranan

**Table 6**  
**Preprocessing Term Weighting**

No	Keyword
1	campursari
2	puppeteer
3	gambuh
4	network
5	musical instrument
6	ketoprak
7	ludruk
8	painting
9	music

<b>10</b>	opera
<b>11</b>	statue
<b>12</b>	poetry
<b>13</b>	literature
<b>14</b>	dance
<b>15</b>	theater
<b>16</b>	puppet

**Table 7**  
**TF-IDF Processing in Term Weighting**

No	Name	Keyword	TF	w	w (normalization)
1	Katijan	campursari	1	0.969	1
2	Katijan	puppeteer	0	0	0
3	Katijan	gambuh	0	0	0
4	Katijan	network	2	1.456	0.666666667
5	Katijan	musical instrument	3	1.269	0.5
6	Katijan	ketoprak	1	1.421	1
7	Katijan	ludruk	1	0.969	0.25
8	Katijan	painting	0	0	0
9	Katijan	music	0	0	0
10	Katijan	opera	0	0	0
11	Katijan	statue	0	0	0
12	Katijan	poetry	0	0	0
13	Katijan	literature	0	0	0
14	Katijan	dance	0	0	0
15	Katijan	theater	0	0	0
16	Katijan	puppet	1	0.659	0.125
17	Mirza Baihaqie	campursari	0	0	0
18	Mirza Baihaqie	puppeteer	0	0	0
19	Mirza Baihaqie	gambuh	2	5.348	1
20	Mirza Baihaqie	jaranan	1	0.728	0.333333333
...	...	...	...	...	...
464	Suwandi	puppet	0	0	0

Based on the test results shown in Table 8, it can be seen that the test results at the beginning of the test show different results. However, the longer the test is carried out, the closer it is to the right results, namely in the 9th and 10th tests, the test results are the same. Therefore, the results of testing at this stage, cluster 1 has as many as 6 members, while cluster 2 has 23 members. To find out the performance of the proposed method, this study uses cluster variance testing (Variant within cluster and Variant between cluster).

Table 9 shows analysis of the performance evaluation of the proposed method (Clustering Fuzzy C-Means) using Vw and Vb techniques. Based on Table 9, it can be seen that in each test the results of Vw are smaller than Vb, meaning that the distribution of data in clusters and between clusters is getting better. Meanwhile, to find out the variants of all clusters can be fulfilled based on the value of V. If the value of V is getting smaller, the cluster value is getting better. Based on the calculation results of an average of 1-10 tests, it is known that the data distribution in all clusters has a value of 0.0000163. These results indicate that the clusters produced are getting better.

**Table 9**  
**Results of Fuzzy C-Means Clustering Performance Evaluation**

No.	Experiment ...	Vc	Vw	Vb	V
<b>1</b>	Experiment 1	Cluster_1 = 13.514639133333	0.814815	10683.15	7.63E-05
		Cluster_2 = 20.311673320158			
<b>2</b>	Experiment 2	Cluster_1 = 20.311673320158	0.185185	10683.15	1.73E-05
		Cluster_2 = 13.514639133333			



3	Experiment 3	Cluster_1 = 20.311673320158	0.185185	10683.15	1.73E-05
		Cluster_2 = 13.514639133333			
4	Experiment 4	Cluster_1 = 20.311673320158	0.185185	10683.15	1.73E-05
		Cluster_2 = 13.514639133333			
5	Experiment 5	Cluster_1 = 20.311673320158	0.185185	10683.15	1.73E-05
		Cluster_2 = 13.514639133333			
6	Experiment 6	Cluster_1 = 13.514639133333	0.814815	10683.15	7.63E-05
		Cluster_2 = 20.311673320158			
7	Experiment 7	Cluster_1 = 13.514639133333	0.814815	10683.15	7.63E-05
		Cluster_2 = 20.311673320158			
8	Experiment 8	Cluster_1 = 20.311673320158	0.185185	10683.15	1.73E-05
		Cluster_2 = 13.514639133333			
9	Experiment 9	Cluster_1 = 13.514639133333	0.114815	10683.15	1.63E-05
		Cluster_2 = 20.311673320158			
10	Experiment 10	Cluster_1 = 13.514639133333	0.114815	10683.15	1.63E-05
		Cluster_2 = 20.311673320158			
<b>Average</b>			0.5	10683.15	4.68E-05

#### 4. Conclusion

For mapping and analyzing data that can provide a complete picture of the demographics of art as a branch of culture in the city of Malang, East Java, the hybrid method is effectively used to solve these problems. The hybrid method proposed in this research is the TF-IDF Text Mining method and Fuzzy C-Means Clustering as a means of decision support systems for the government to make a policy. This decision support system is in the form of an art taxonomy application. This study uses a sample of data from a number of artists who are grouped into 16 art fields. TF-IDF is used in data preprocessing in the text mining stage. Then, fuzzy C-Means Clustering is used to group the results of normalized weight calculations into 16 variables. The analysis of this research indicate that the clusters produced using hybrid TF-IDF Fuzzy C-Means Clustering produce better clusters as evidenced by the smaller V value  $V = 0.0000468$ .

#### Reference

- [1] N. Shahid and S. Aneja, "Internet of Things: Vision, application areas and research challenges," in *Proceedings of the International Conference on IoT in Social, Mobile, Analytics and Cloud, I-SMAC 2017*, 2017.
- [2] H. C. Soong, N. B. A. Jalil, R. Kumar Ayyasamy, and R. Akbar, "The essential of sentiment analysis and opinion mining in social media : Introduction and survey of the recent approaches and techniques," in *ISCAIE 2019 - 2019 IEEE Symposium on Computer Applications and Industrial Electronics*, 2019.
- [3] I. Mujahidin, D. A. Prasetya, Nachrowie, S. A. Sena, and P. S. Arinda, "Performance tuning of spade card antenna using mean average loss of backpropagation neural network," *Int. J. Adv. Comput. Sci. Appl.*, 2020.
- [4] M. Alaa, A. A. Zaidan, B. B. Zaidan, M. Talal, and M. L. M. Kiah, "A review of smart home applications based on Internet of Things," *Journal of Network and Computer Applications*. 2017.
- [5] I. Mujahidin, D. A. Prasetya, A. B. Setywan, and P. S. Arinda, "Circular Polarization 5.5 GHz Double Square Margin Antenna in the Metal Framed Smartphone for SIL Wireless Sensor," in *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2019, pp. 1–6.
- [6] I. Hafeez, M. Antikainen, A. Y. Ding, and S. Tarkoma, "IoT-KEEPER: Detecting Malicious IoT Network Activity Using Online Traffic Analysis at the Edge," *IEEE Trans. Netw. Serv. Manag.*, 2020.
- [7] R. Yuwono and I. Mujahidin, "Rectifier using UWB microstrip antenna as electromagnetic energy harvester for GSM, CCTV and Wi-Fi transmitter," *J. Commun.*, 2019.
- [8] P. Masulli, F. Masulli, S. Rovetta, A. Lintas, and A. E. P. Villa, "Fuzzy Clustering for Exploratory Analysis of EEG Event-Related Potentials," *IEEE Trans. Fuzzy Syst.*, 2020.
- [9] I. Mujahidin, S. H. Pramono, and A. Muslim, "5.5 Ghz Directional Antenna with 90 Degree Phase

- Difference Output,” 2018.
- [10] Y. Song, J. Lu, H. Lu, and G. Zhang, “Fuzzy Clustering-Based Adaptive Regression for Drifting Data Streams,” *IEEE Trans. Fuzzy Syst.*, 2020.
  - [11] R. Yuwono, I. Mujahidin, A. Mustofa, and Aisah, “Rectifier using UFO microstrip antenna as electromagnetic energy harvester,” *Adv. Sci. Lett.*, 2015.
  - [12] F. ŞİŞİK and E. SERT, “Brain tumor segmentation approach based on the extreme learning machine and significantly fast and robust fuzzy C-means clustering algorithms running on Raspberry Pi hardware,” *Med. Hypotheses*, 2020.
  - [13] L. Zhang, M. Luo, J. Liu, Z. Li, and Q. Zheng, “Diverse fuzzy c-means for image clustering,” *Pattern Recognit. Lett.*, 2020.
  - [14] M. Charwand, M. Gitizadeh, P. Siano, G. Chicco, and Z. Moshavash, “Clustering of electrical load patterns and time periods using uncertainty-based multi-level amplitude thresholding,” *Int. J. Electr. Power Energy Syst.*, 2020.
  - [15] N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, “Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering,” *Int. J. Med. Inform.*, 2019.
  - [16] T. Lei, X. Jia, Y. Zhang, S. Liu, H. Meng, and A. K. Nandi, “Superpixel-Based Fast Fuzzy C-Means Clustering for Color Image Segmentation,” *IEEE Trans. Fuzzy Syst.*, 2019.
  - [17] H. He and Y. Tan, “Unsupervised Classification of Multivariate Time Series Using VPCA and Fuzzy Clustering with Spatial Weighted Matrix Distance,” *IEEE Trans. Cybern.*, 2020.
  - [18] K. Zhou and S. Yang, “Effect of cluster size distribution on clustering: a comparative study of k-means and fuzzy c-means clustering,” *Pattern Anal. Appl.*, 2020.
  - [19] P. D. Pantula, S. S. Miriyala, and K. Mitra, “An Evolutionary Neuro-Fuzzy C-means Clustering Technique,” *Eng. Appl. Artif. Intell.*, 2020.
  - [20] N. Bharill *et al.*, “Fuzzy knowledge based performance analysis on big data,” *Neurocomputing*, 2020.