UUM PRESS
Universiti Utara Malaysia
06010 UUM SINTOK
KEDAH DARULAMAN
MALAYSIA

Tel: 604-928 4795/4796
Faks (Fax): 604-928 4792
Laman Web (Web): http://uumpress.uum.edu.my

Reference : UUM/UUMPRESS/P-48/15
Date : 8 February 2021

TO WHOM IT MAY CONCERN

Dear Sir,

APPOINTMENT AS REVIEWER

We hereby certify that Dr. Abdullah Husin of Universitas Islam Indragiri Indonesia was invited as peer review of Journal of Information and Communication Technology (JICT) in 2021.

Academic contribution of Dr. Abdullah Husin helped to maintain the high peer review standard of this journal.

"PRIHATIN RAKYAT: DARURAT MEMERANGI COVID-19"
"BERKHIDMAT UNTUK NEGARA"
"KEDAH SEJAHTERA – NIKMAT UNTUK SEMUA"
"KNOWLEDGE, VIRTUE, SERVICE"

Upholding the principles of trust and integrity

Your Sincerely

PROF. DR. KU RUHANA KU MAHAMUD
Chief Editor, Journal of Information and Communication Technology (JICT)
Universiti Utara Malaysia

UUM/UUMPRESS/P-48/15

Kepada/ *To:*

Editor Pengurus / *Managing Editor*
Journal of Information and Communication Technology
UUM Press
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman
Tel: +604-928 4815
Faks: +604-928 4792

PELANTIKAN PENILAI / *APPOINTMENT OF REVIEWER*
JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Nama / *Name*	:	Dr. Abdullah Husin, S.Si., M.Kom
No. Kad Pengenalan / *IC No.*	:	NIDN 1008037001
Jawatan / *Position*	:	Department of Information System Universitas Islam Indragiri Indonesia
Nama Bank & No. Akaun / *Name of Bank & Account No.*	:	
Cawangan / *Branch*	:	
		(untuk tujuan penyediaan cek / *for cheque preparation purpose*)
Alamat Jabatan / *Address of the Department*	:	Campus of Universitas Islam Indragiri Jl. Provinsi Tembilahan Hulu Indragiri Hilir Riau Indonesia 29213
No. Telefon / *Tel. Number*	:	(pejabat/*office*) (rumah/*residence*) +628127580419
Emel / *E-mail*	:	abdialam@yahoo.com
Tajuk Makalah / *Title of Article*	:	A Hybrid K-Means Hierarchical Algorithm for Natural Disaster Mitigation Clustering
Bidang Kepakaran / *Expertise*	:	Computer Science

Merujuk kepada surat tuan UUM/UUM PRESS/P-48/15, saya maklumkan bahawa saya ***MENERIMA/ ~~MENOLAK~~** pelantikan sebagai penilai untuk Journal of Information and Communication Technology dalam bidang seperti yang dinyatakan di atas. Saya akan memberikan ulasan saya sebelum atau pada tarikh yang ditetapkan. *Referring to your letter UUM/UUMPRESS/P-48/15, I am hereby ****ACCEPT/~~DECLINE~~*** *the offer of being a reviewer for Journal of Information and Communication Technology in the field stated above. I will submit my review by the stipulated date mentioned.*

Tarikh/*Date* : 31 October 2021 Tandatangan/*Signature* :

* Potong mana yang tidak berkenaan/*Delete whichever is inapplicable.*

[JICT] Article Review Request

Dari: Nor Aziani Jamil (aziani@uum.edu.my)

Kepada: abdialam@yahoo.com

Tanggal: Minggu, 31 Oktober 2021 08.40 WIB

Dr Abdullah Husin, S.Si., M.Kom:

I believe that you would serve as an excellent reviewer of the manuscript, "A Hybrid K-Means Hierarchical Algorithm for Natural Disaster Mitigation Clustering," which has been submitted to Journal of Information and Communication Technology. The submission's abstract is inserted below, and I hope that you will consider undertaking this important task for us.

Please log into the journal web site by 2021-11-07 to indicate whether you will undertake the review or not, as well as to access the submission and to record your review and recommendation. The web site is http://e-journal.uum.edu.my/index.php/jict

The review itself is due 2021-11-14.

As a reviewer, you will receive a token of appreciation for the articles that have been reviewed. Please fill in this form and return with your reviewed article. Click here to download the form. Kindly attach your Appointment of Reviewer Form and *Borang Maklumat Akaun (Individu)* with the bank statement which include name and account number as our reference to Madam Nor Aziani Jamil at aziani@uum.edu.my.

If you do not have your username and password for the journal's web site, you can use this link to reset your password (which will then be emailed to you along with your username). http://e-journal.uum.edu.my/index.php/jict/login/lostPassword

Submission URL: http://e-journal.uum.edu.my/index.php/jict/reviewer/submission?submissionId=13768&reviewId=5724&key=GDfiQapf

Thank you for considering this request.

Nor Aziani Jamil
Universiti Utara Malaysia
aziani@uum.edu.my

"A Hybrid K-Means Hierarchical Algorithm for Natural Disaster Mitigation Clustering"

Abstract

Cluster methods such as k-means have been widely used to group areas with a relatively equal number of disasters to determine which areas are prone to natural disasters. However, this approach still faces the question of how the k-means method achieves good performance. This paper aims to apply the k-means algorithm and hierarchy for the anticipated clustering process of natural disaster mitigation levels. This paper proposes a hybrid approach that combines the k-means algorithm and a hierarchy that applies it to existing datasets—also added keyword and disaster type fields to provide additional information for a better clustering process. This process produced 3 clusters for the anticipation level of natural disaster mitigation, namely Cluster 1 for low anticipation level, cluster 2 for medium anticipation level, and cluster 3 for high anticipation level. In addition, we validated by experts so that there were 67 districts / cities including cluster 1 (82.7%), 9 distric ts / cities including cluster 2 (11.1%), and the remaining 5 districts / cities including cluster 3 (6.2%). From the analysis of the silhouette coefficient calculation, the hybrid algorithm can provide relatively homogeneous clustering results. Furthermore, applying the hybrid algorithm to the keyword segment and the type of disaster also produces a homogeneous clustering based on the Purity coefficient calculation. The clusters have a TPu value of 0.8271605, which is close to 1 and, hence, represents the results of a good clustering algorithm.

The following message is being delivered on behalf of Journal of Information and Communication Technology.

[JICT] Article Review Acknowledgement

Dari: Nor Aziani Jamil (aziani@uum.edu.my)

Kepada: abdialam@yahoo.com

Tanggal: Minggu, 7 November 2021 07.33 GMT+7

Dr Abdullah Husin, S.Si., M.Kom:

Thank you for completing the review of the submission, "A Hybrid K-Means Hierarchical Algorithm for Natural Disaster Mitigation Clustering," for Journal of Information and Communication Technology. We appreciate your contribution to the quality of the work that we publish.

Nor Aziani Jamil
Universiti Utara Malaysia
aziani@uum.edu.my

The following message is being delivered on behalf of Journal of Information and Communication Technology.

A Hybrid K-Means Hierarchical Algorithm for Natural Disaster Mitigation Clustering

ABSTRACT

Cluster methods such as k-means have been widely used to group areas with a relatively equal number of disasters to determine areas prone to natural disasters. However, it is troublesome to obtain a homogeneous clustering result of the k-means method because this method is sensitive to a random selection of the centers of the cluster. This paper presents the result of a study that aimed to apply a proposed hybrid approach of the combined k-means algorithm and hierarchy to the clustering process of anticipation level datasets of natural disaster mitigation in Indonesia. This study also added keyword and disaster-type fields to provide additional information for a better clustering process. The clustering process produced three clusters for the anticipation level of natural disaster mitigation. Based on the validation from the expert, 67 districts/cities (82.7%) fall into cluster 1 (low anticipation), 9 districts/cities (11.1%) is classified into cluster 2 (medium), and the remaining 5 districts/cities (6.2%) is categorized in cluster 3 (high anticipation). From the analysis of the calculation of the silhouette coefficient, the hybrid algorithm provides relatively homogeneous clustering results. Furthermore, applying the hybrid algorithm to the keyword segment and the type of disaster also produces a homogeneous clustering as indicated by the calculated purity coefficient and the total purity values. Thus, the proposed hybrid algorithm is a good clustering algorithm.

Commented [w1]: Need to add the significance of study

Keywords: clustering, hybrid, k-means, mitigation, natural disaster

INTRODUCTION

Many countries in the world are prone to natural disasters, including Indonesia. High rainfall, active tectonic and volcanic activities, and natural disasters, including floods, volcanic eruptions, earthquakes, and tsunamis, are very common occurrences in Indonesia. Consequently, disaster mitigation efforts are indispensable to minimize the impact of a disaster in many regions in Indonesia.

Many kinds of research on natural disaster mitigation have been carried out. Prihandoko and Bertalya (2016) studied some factors for natural disasters in Indonesia and found that the geographical condition is the main cause for natural disaster occurrence instead of the weather condition. Anjayani (2008) suggests that the earthquake hypocenters strongly correlate with the locations of many active volcanoes. Moreover, Supriyadi et al. (2018) revealed that floods are the most common natural disaster in Indonesia. In 2007, the Indonesian government passed the Law of the Republic of Indonesia Number 24 of 2007 concerning disaster management, the national reference (Indonesia, 24 C.E.). Rachmawati (2018) conducted a study on the community’s knowledge in the disaster area to measure people’s general awareness of areas at risk to lessen the consequences of natural disasters.

Many previous works have used data and information on natural disaster mitigation compiled by the Indonesia agency for disaster management (BNPB). Sadewo et al. (2018) conducted a clustering of disaster mitigation anticipation levels at the provincial government using the k-means method. Priatmodjo (2011) stated that disaster mitigation requires preparedness, which includes analysis of potential disasters and planning for anticipation. He also developed tools for disaster prevention and management. Atasever (2017) revealed a method to determine the level of damage due to a disaster. Han and Kamber (2001) use data mining to process large amounts of disaster data. Meanwhile, Prihandoko et al. (2017) used data mining techniques to analyze and predict disaster mitigation anticipation levels.

Various methods have been used to cluster the anticipation level of natural disaster mitigation. Ediyanto et al. (2013) described hierarchical clustering based on euclidean distances to calculate the level of similarity. Hierarchical clustering is usually shown in the form of a tree diagram (dendrogram). Whereas for large amounts of data, the k-means method is more often used (Bagirov et al., 2011).

Commented [w2]: Need to add the problem

This paper presents the results of the clustering process of datasets of mitigation activities using data mining techniques to determine the anticipation levels for natural disasters. In this study, the k-means method and hierarchy were combined with a hybrid approach to producing a hierarchical k-means hybrid clustering. The clustering process was conducted using datasets originating from various research reports of natural disaster mitigation activities conducted by local and provincial governments available in the National Scientific Repository managed by LIPI. The keyword fields and disaster types were also added as additional information for a better clustering process. Clustering is a powerful tool for data mining, which applies to virtually every field where large amounts of information are needed for data organization (Abdulsahib & Kamaruddin, 2015).

Commented [w3]: Better if this statement replace with the significance of this study.

RELATED WORKS

Cluster methods such as k-means have been widely used to group areas with relatively the same number of disaster characteristics to see which areas are prone to natural disasters (Yana et al., 2018). A study by Supriyadi et al. (2018) used k-means to classify disaster-prone areas into three clusters: high, medium, and low. In addition, Yana et al. (2018) found two regional clusters in Indonesia, namely, prone to and not prone to natural disasters. Prihandoko and Bertalya (2016) suggested the cluster correlation between natural disasters, the number of victims, and weather conditions using k-means.

Sadewo et al. (2018) classify provinces according to their mitigation efforts using k-means in a disaster mitigation study. His research results show three clusters (high, medium, and low mitigation efforts). The regions of West Java, Central Java, and East Java enter into a high level of mitigation. In another study, Kandel et al. (2014) discuss a comprehensive assessment of fuzzy techniques for mitigation. Kandel uses incremental fuzzy clustering to group mitigation data. Nevertheless, the author did not experiment with other clustering techniques on the same dataset for accuracy measures. Only one algorithm result does not necessarily make a correct analysis. The author should have experimented with some more efficient clustering techniques or comparative study to strengthen the effectiveness of the k-means algorithm against other methods.

Combined (hybrid) k-means and hierarchical clustering have been applied to studying disasters such as air pollution (Govender & Sivakumar, 2020). Hierarchical and k-means clustering are two approaches but have different strengths and weaknesses. For instance, hierarchical clustering identifies groups in a tree-like structure but suffers from computational complexity in large datasets. In contrast, k-means clustering is efficient but designed to identify homogeneous spherically shaped clusters (Peterson et al., 2018). Studies combine these two methods, such as Govender and Sivakumar (2020), which applied a combination of k-means and hierarchical clustering techniques to analyze air pollution. Atasever (2017) combines the k-means cluster method and BSA (hybrid) to detect damage to natural disaster areas. The data results are grouped with a hybrid approach into two classes: damaged and undamaged areas.

Moreover, some studies compare and combined k-means cluster methods with other cluster methods. Nugroho (2021) compared the kernel k-means algorithm on bipartite graphs and k-means on the term-document matrix in the COVID-19 research dataset. The result is that the k-means kernel algorithm provides slightly better validation compared to k-means. Balavand et al. (2018) combined the CSA k-means method with data envelopment analysis and compared it with the DCSPO, GCUK, ACDE algorithms.

However, their other research works use combination clustering methods for disaster or other subjects. Wen et al. (2019) developed a combination of GIS technology and the QUEST cluster algorithm, and the results showed the distribution of drought disaster areas. Ali et al. (2018) discuss disaster management with cluster techniques for emergencies, and Welton-Mitchell et al. (2018) discuss clusters of people affected by the disaster. Ng and Khor (2016) evaluated the rapid profiling with clustering algorithms for plantation stocks on Bursa Malaysia. Ng and Khor utilized expectation maximization (EM), k-means (KM), and hierarchical clustering (HC) algorithms to cluster the 38 plantation stocks listed on Bursa Malaysia. The results showed that a cluster resulting from EM had a better profile.

This study seeks to address some of the shortcomings of previous research. First, no one has explicitly used hybrid KM and hierarchies algorithms to suppress the level of disaster mitigation effort. Second, previous research only surveyed the combination of KM and hierarchies clustering studies. In this study, we present the application of the hybrid clustering approach that amalgamates the two methods to identify general-shaped the level of disaster mitigation clusters more efficiently. Specifically, we first partitioned the dataset into groups using the KM algorithm. The next stage is to combine KM and HC as a hybrid approach. The hybrid approach is used because the KM algorithm uses random observational data to determine the initial centroid. The clustering solution KM is very sensitive to a random selection of the centers of the cluster. Therefore, clustering results may vary when recomputing. A hybrid approach that combines KM and hierarchical algorithms can avoid this problem.

Commented [w4]: This statement is just suitable for conclusion

METHODOLOGY

This study clustered the natural disaster dataset from technical reports on natural disaster research compiled by the National Scientific Repository of the Indonesian Institute of Sciences (RIN LIPI). This dataset consists of 237 documents of technical research reports conducted by researchers within and outside LIPI. A total of 81 districts and cities (next named "region") in Indonesia are included in this dataset. We created a mitigation category for each of the technical reports on natural disaster research. This category consists of A, B, C, D, E, and F. Table 1 presents the dataset summary.

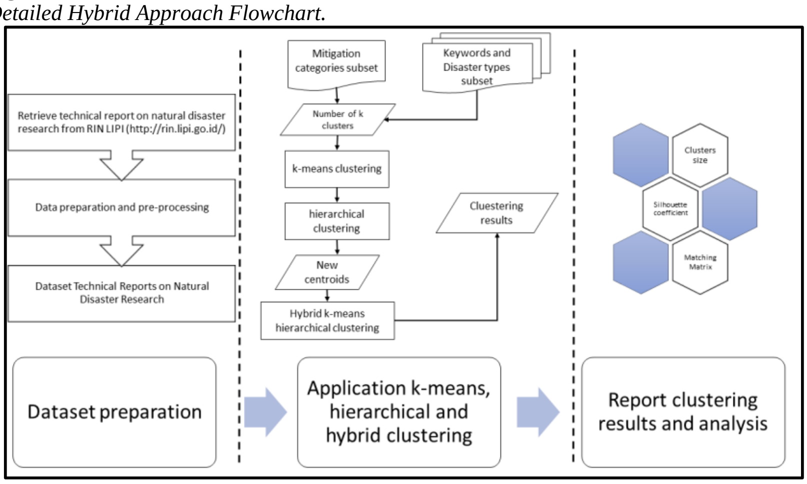
Table 1
Summary of Dataset.

No	Region	Keyword	Type	Mitigation	Mitigation Code
1	Kota Padangsidimpuan	Monograph; Volcanic; Basic Data; Andesite;	Earthquake	Assessment of Disaster Risk and Characteristics	C
2	Kab. Simeulue	Active tectonics; earthquake; Continuous GPS; Subduction zone; Megathrust; Sumatra	Earthquake; Tsunami	Preparation and Installation of Early Warning System Instruments	D
3	Kota Padang	Active tectonics; earthquake; Continuous	Earthquake; Tsunami	Assessment of Disaster Risk and Characteristics	C
4	Kab. Sumedang	Rainfall; Slope Instability; Hydrological	Landslide	Construction and Strengthening of Building Structures	A
5	Kab. Kebumen	Weathering; Residual Soil; Physical Properties	Landslide	Assessment of Disaster Risk and Characteristics	C
6	Kab. Tanggamus	Disaster, Tanggamus; Vulnerability; Mitigation; Spatial	Earthquake; tsunami	Planning and Implementation of Spatial Planning	E
7	Kota Serang	Earthquake; Geotechnical; Liquidation; Decline	Earthquake	Construction and Strengthening of Building Structures	A
8	Kota Bandung	Bandung Basin; Garut and Sumedang; Hazard Zoning; Earthquake; Active Fault	Earthquake	Planning and Implementation of Spatial Planning	E

9	Kab. Kepulauan Talaud	Talaud Regency; Border; Environment	Earthquake	Assessment of Disaster Risk and Characteristics	C
10	Kab. Kepulauan Mentawai	Mentawai Islands; Sumatran GPS Array (Sugar); Coral;	Earthquake	Preparation and Installation of Early Warning System Instruments	D

We made three clusters of anticipation levels of disaster mitigation from each category: high, medium, and low. To ensure its validity, we confirmed this cluster with experts from the Research Center for Geotechnology–Indonesian Institute of Sciences, with expertise in natural disasters. We used a hybrid approach to categorize the anticipation level that combines the KM and the HC algorithm. This method adopts Sadewo et al. (2018) concerning clustering of anticipated levels of natural disaster mitigation at the provisional level and Atasever (2017) concerning a hybrid approach to detect damage due to natural disasters. We also used the R programming language with the factoextra library in computing the application of a hybrid approach (Kassambara & Mundt, 2020). Figure 1 illustrates the exact method.

Figure 1
Detailed Hybrid Approach Flowchart.



In the first stage, we applied the KM algorithm with a numerical vector parameter of the mitigation category, the number of clusters $k = 3$. The next step applies HC to the dataset of disaster mitigation categories with the parameter number of k clusters 3 and the ward method. We used silhouette measures to find out the validity of this clustering result.

The next stage was to combine the KM and hierarchical clustering as a hybrid approach. The hybrid approach calculated the hierarchical clusters and cut the tree into several k clusters. It then calculated the centroid of each cluster. Finally, the hybrid approach calculated the KM using the cluster centroid obtained from the previous calculation as the cluster's initial centroid. Next, hierarchical and KM clustering results were compared, respectively, with hybrids using a matching matrix.

We confirmed the cluster hybrids' results from the disaster mitigation category with experts to account for their validity. Furthermore, we made these three clusters a ground truth reference for applying the hybrid KM hierarchical algorithm to the natural disaster dataset on the subset of keywords and types of disasters.

We used disaster-type and keyword subsets as the clustering base. The subsets represented the mitigation category relationship. This study carries out the clustering stage from each subset using the KM algorithm by building a TF-IDF matrix to convert the document into a TF-IDF vector. We eliminated stop words. In the computation, stop words are filtered out before or after processing natural language data (text), such as the, is, at, which, and on. We did not apply the stemming process because the terms in the sections we used are specific terms that reflect the contents' technical report documents. We applied the hybrid algorithm of KM and hierarchies with the number of clusters k = 3 according to the anticipated level of natural disaster mitigation that has been defined.

We utilized the unsupervised learning technique (e.g., clustering) to divide the input data point with some common properties. In the previous stage, we have defined prior knowledge class labels as the ground truth. In order to validate the clustering results, we then intuitively used a matching matrix method. As described by Samatova et al. (2013), the matching matrix (Figure 2) is a $V \times W$ matrix, where V is the number of class labels in P and W is the total number of resulting clusters. Each row of the matrix represents one class label, and each column represents a cluster ID. Each m_{ij} entry represents the number of points from Class i that are present in cluster g_j. The table was filled based on the prior knowledge P and clusters obtained using U.

In this paper, we employed purity as the validation metric for our hybrid algorithm. Purity (Pu) is a measure to analyze the cluster's homogeneity concerning the class labels. Equation (1) calculates purity as follows:

$$Pu_{gj} = \max_{i = 1 \; to \; V} P_{ij} \tag{1}$$

This measure takes any value in the range of 1/V to 1. A value of 1 indicates an utterly homogeneous cluster. We calculated the total purity (TPu) for the entire cluster's results. The TPu, as denoted by Eq. (2) for the whole cluster set, was calculated as the sum of each cluster's purities weighted by the number of elements in each cluster.

$$TPu = \sum_{j=1}^{W} \frac{m_j}{M} P_{ugj} \tag{2}$$

Figure 2

Matching Matrix Template (Samatova et al., 2013).

		Cluster ID				
		1	2	......	W	
Class Labels	1	m_{11}	m_{12}	......	m_{1W}	
	2	m_{21}	m_{22}	......	m_{2W}	
	⋮	⋮	⋮	......	⋮	
	V	m_{V1}	m_{V2}	......	m_{VW}	
Total Elements per Cluster m_j		$\sum_{i=1}^{V} m_{i1}$	$\sum_{i=1}^{V} m_{i2}$	......	$\sum_{i=1}^{V} m_{iW}$	$T = \sum_{j=1}^{W} m_j$

RESULTS

In the first stage, the KM algorithm applied a numerical vector parameter of the mitigation category, the number of clusters k = 3. Table 2 presents a pruned result of 15 clustered data. Next, we presented the results of KM clustering, as shown in Figure 3. To determine the validity of this clustering result, we used the silhouette size shown in Figure 4. The silhouette coefficient measures how well an observation is grouped and estimates the average distance between clusters (i.e., the average silhouette width). The negative silhouettes coefficient indicates that the observations may be grouped in the wrong cluster. Table 3 presents some data with negative silhouette coefficients., indicating that some data points in cluster 2 were grouped in the wrong cluster.

Figure 3

Cluster plot of KM on Mitigation Category.

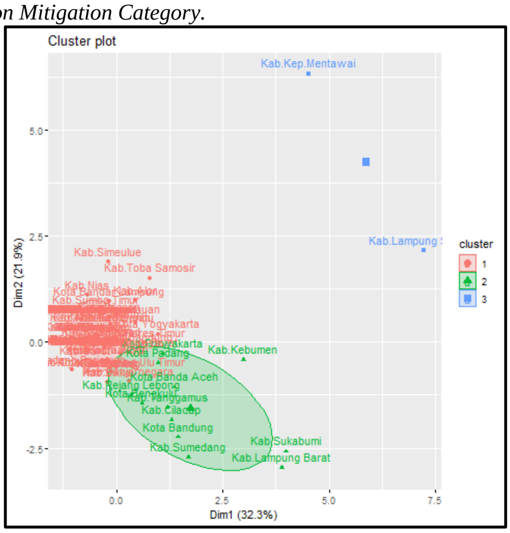


Figure 4

Cluster silhouette plot of KM on Mitigation Category.

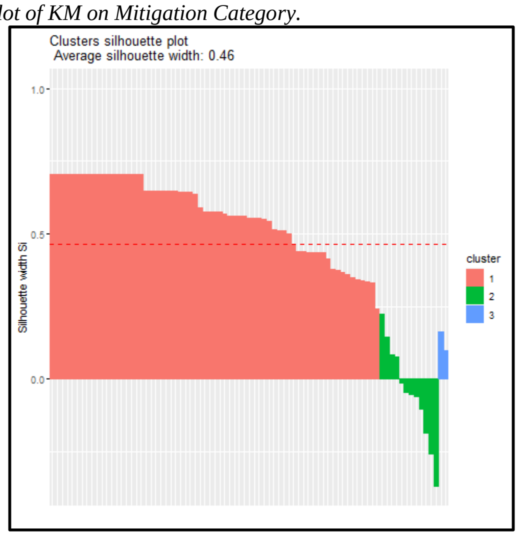


Table 2
KM Clustering Result (pruned).

Region	Cluster
Kab. Aceh Besar	1
Kab. Adm Kep. Seribu	1
Kab. Alor	1
Kab. Badung	1
Kab. Bandung	1
Kab. Bandung Barat	1
Kab. Banggai	1
Kab. Banggai Kep.	1
Kab. Banjarnegara	1
Kab. Bantul	1
Kab. Banyuwangi	1
Kab. Bekasi	1
Kab. Bengkulu Selatan	1
Kab. Bengkulu Tengah	1
Kab. Biak Numfor	1

Table 3
Data Point with Negative Silhouette Coefficient in KM Clustering.

Region	Cluster	Neighbor	Sil_width
Kota Bandung	2	1	-0.01
Kab. Cilacap	2	1	-0.04
Kab. Tanggamus	2	1	-0.05
Kota Banda Aceh	2	1	-0.06
Kota Bengkulu	2	1	-0.10
Kab. Purwakarta	2	1	-0.18
Kab. Rejang Lebong	2	1	-0.25
Kota Padang	2	1	-0.36

In the next stage, we applied the HC algorithm to the dataset of disaster mitigation categories with the parameter number of k clusters 3 and the ward method. Table 4 presents a summary of the 15 clustered data, and Table 5 presents the data points with negative silhouette coefficients. The results of HC are also presented in a dendrogram graph, as shown in Figure 5. As shown in Table 5, there are only two data points with negative silhouette coefficients. Figure 6 shows that the averaged silhouette width of HC is 0.53, which is higher than that of KM (see Figure 4) for the same clustering category. Thus, HC may produce a better result than the KM algorithm in clustering the dataset.

Table 4
Hierarchy Clustering Result (pruned).

Region	Cluster
Kab. Aceh Besar	1
Kab. Adm Kep. Seribu	1
Kab. Alor	1
Kab. Badung	1
Kab. Bandung	1
Kab. Bandung Barat	1
Kab. Banggai	1

Kab. Banggai Kep.	1
Kab. Banjarnegara	1
Kab. Bantul	1
Kab. Banyuwangi	1
Kab. Bekasi	1
Kab. Bengkulu Selatan	1
Kab. Bengkulu Tengah	1
Kab. Biak Numfor	1

Table 5

Data Points with Negative Silhouette Coefficient in Hierarchy Clustering.

Region	Cluster	Neighbor	Sil_width
Kota Banda Aceh	2	1	-0.02
Kab. Kebumen	2	1	-0.02

Figure 5

Cluster Dendrogram of Hierarchical Clustering on Mitigation Category.

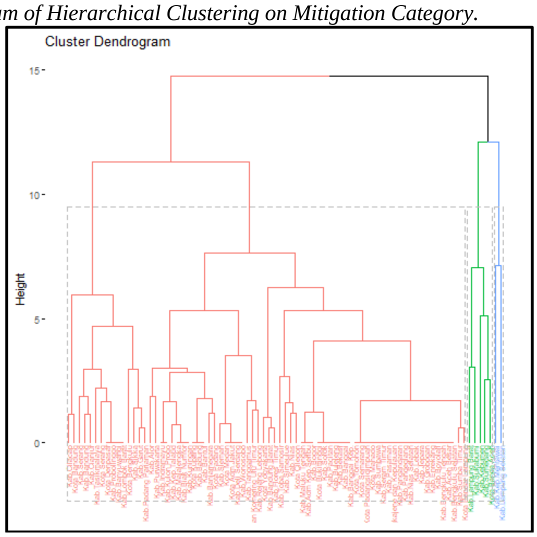
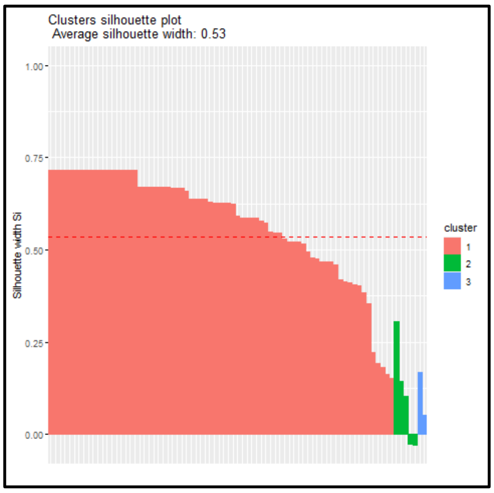


Figure 6

Cluster Silhouette of Hierarchical Clustering on Mitigation Category.



The next stage was to combine KM and HC as a hybrid approach. We used the hybrid approach because the KM algorithm uses random observational data to determine the initial centroid. The clustering solution of KM is very sensitive to a random selection of the centers of the cluster. Therefore, clustering results may vary when recomputing.

The hybrid approach calculated the HC and cut the tree into several k clusters. It then calculated the center of each cluster and calculated the KM using the cluster center obtained from the previous calculation as the cluster's initial center. The new centroids were defined as the mean of the variables in the cluster. Table 6 summarizes the result of the calculation of the new centroid.

Table 6

New Centroids for Hybrid KM Hierarchy Clustering.

Cluster	A	B	C	D	E	F
1	−0,19	−0,14	−0,18	−0,12	−0,12	0
2	2,86	1,37	0,74	−0,02	1,62	0,16
3	0,13	1,86	5,11	4,77	0,66	−0,46

Next, KM clustering was applied using the cluster's center above to obtain the cluster results, as presented in Table 7. Table 8 presents the negative values of the silhouette. Next, the hierarchical and hybrid clustering results were compared using the match matrix from Tables 4 and 7. Table 9 shows that the hybrid algorithm produced better clustering results than the standard KM and HC algorithm. The data points are clustered homogeneously into each cluster.

Table 7

Hybrid KM Hierarchy Clustering Result (pruned).

Region	Cluster
Kab. Aceh Besar	1
Kab. Adm Kep. Seribu	1
Kab. Alor	1
Kab. Badung	1
Kab. Bandung	1
Kab. Bandung Barat	1
Kab. Banggai	1

Kab. Banggai Kep.	1
Kab. Banjarnegara	1
Kab. Bantul	1
Kab. Banyuwangi	1
Kab. Bekasi	1
Kab. Bengkulu Selatan	1
Kab. Bengkulu Tengah	1
Kab. Biak Numfor	1

Table 8
Data Points with Negative Silhouette Coefficient in Hybrid Clustering.

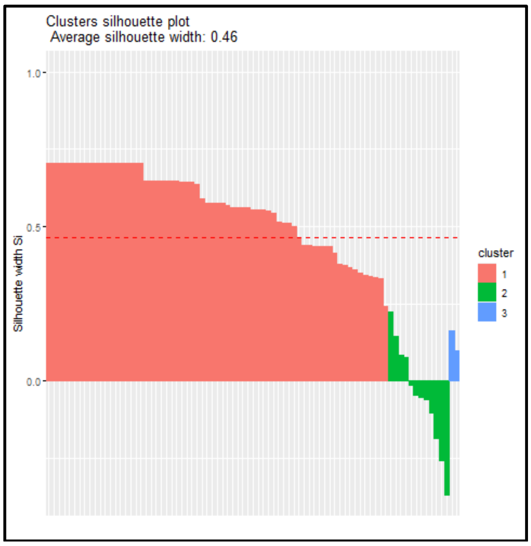
Region	Cluster	Neighbor	Sil_width
Kota Bandung	2	1	-0.01
Kab. Cilacap	2	1	-0.04
Kab. Tanggamus	2	1	-0.05
Kota Banda Aceh	2	1	-0.06
Kota Bengkulu	2	1	-0.10
Kab. Purwakarta	2	1	-0.18
Kab. Rejang Lebong	2	1	-0.25
Kota Padang	2	1	-0.36

Table 9
Matching Matrix of Standard Hierarchy and Hybrid Clustering.

		Hybrid results		
		1	2	3
Hierarchy results	1	67	0	0
	2	7	5	0
	3	0	0	2

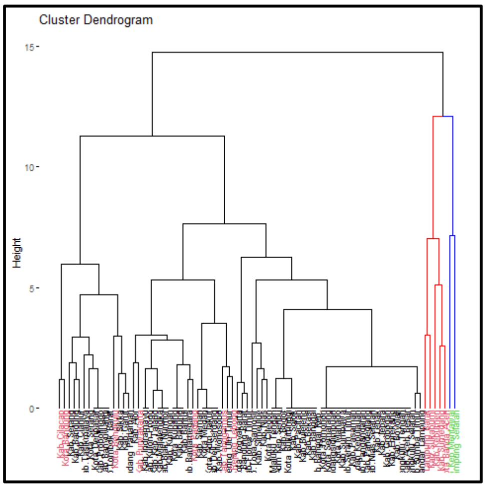
Figure 7 shows that most data points have positive silhouette values, which means that the data points are clustered into the correct cluster. However, we can also see that in cluster two, eight data points have negative values. The negative value indicates that there is a possibility that the data points are not clustered correctly in Cluster 2. This finding is confirmed by the silhouette values earlier. The final clustering solution, KM, regrouped some data.

Figure 7
Cluster silhouette of Hybrid Clustering on Mitigation Category.



From Figure 8 below, we can observe that most data points are clustered homogeneously into a predetermined cluster. However, some data do not appear to fit into the cluster. Cluster 2, marked in red, contains 7 data points included in Cluster 1. The mis-clustered data points were confirmed by calculating the silhouette value, as illustrated in Figure 7 and Table 8.

Figure 8
Cluster Dendrogram of Hybrid KM Hierarchy Clustering on Mitigation Category.



In the same way, using the matching matrix, we compared the clustering of standard KM in Table 2 with the hybrid approach in Table 7, as shown in Table 10. Table 10 describes a matching matrix that consolidates the results of the standard and hybrid KM clustering. Each Cluster 1, 2, and 3 shows the clustered data points correctly by the two types of clustering algorithms applied.

Table 10

Matching Matrix of Standard KM and Hybrid Clustering.

		Hybrid result		
		1	2	3
KM result	1	67	0	0
	2	0	12	0
	3	0	0	2

We consulted the results of this clustering with related experts. The clustering results show that there are two regions, namely, Kab. Mentawai and Kab. South Lampung, in the high anticipation category. Based on expert justification, several other cities, Banda Aceh, Padang City, and Bengkulu City, could be highly anticipated. This difference is due to the lack of research in category A that discusses building structures' construction and strengthening.

We also confirmed the cluster hybrids' results from the disaster mitigation category with experts to account for their validity. The expert comes from the LIPI Geotechnology Research Center with expertise in the field of natural disasters. Furthermore, we define three clusters as a ground truth reference. The ground truth was for applying the hybrid KM hierarchical algorithm. The hybrid algorithm was then used to keyword and disaster types keyword. Table 11 presents the ground truth, which consolidates to validate the clustering results of applying the hybrid algorithm and validation from experts to validate the clustering results. Cluster 1 represents areas with low anticipation levels, cluster 2 for medium anticipation levels, and cluster 3 for high anticipation levels.

Table 11

Consolidated Clustering Result as Ground Truth.

Cluster (Count)	Region
1 (67)	Kab. Aceh Besar, Kab. Adm Kep. Seribu, Kab. Alor, Kab. Badung, Kab. Bandung, Kab.

	Bandung Barat, Kab. Banggai, Kab. Banggai Kep., Kab. Banjarnegara, Kab. Bantul, Kab. Banyuwangi, Kab. Bekasi, Kab. Bengkulu Selatan, Kab. Bengkulu Tengah, Kab. Biak Numfor, Kab. Bogor, Kab. Boyolali, Kab. Buleleng, Kab. Cianjur, Kab. Deli Serdang, Kab. Enrekang, Kab. Flores Timur, Kab. Garut, Kab. Grobogan, Kab. Indramayu, Kab. Jepara, Kab. Kep. Talaud, Kab. Kulon Progo, Kab. Kuningan, Kab. Lebak, Kab. Lombok Barat, Kab. Majalengka, Kab. Maluku Tengah, Kab. Nias, Kab. Nias Selatan, Kab. Ogan Komering Ulu Timur, Kab. Pacitan, Kab. Padang Pariaman, Kab. Pandeglang, Kab. Pangandaran, Kab. Pangkajene dan Kepulauan, Kab. Pasaman, Kab. Probolinggo, Kab. Seram Timur, Kab. Serang, Kab. Sikka, Kab. Simeulue, Kab. Sleman, Kab. Solok, Kab. Sumba Timur, Kab. Sumbawa, Kab. Tasikmalaya, Kab. Toba Samosir, Kab. Wonosobo, Kota Adm. Jakut, Kota Bandar Lampung, Kota Bukittinggi, Kota Cilegon, Kota Denpasar, Kota Kupang, Kota Manado, Kota Medan, Kota Padangsidimpuan, Kota Semarang, Kota Serang, Kota Tomohon, Kota Yogyakarta
2 (9)	Kab. Cilacap, Kab. Kebumen, Kab. Lampung Barat, Kab. Purwakarta, Kab. Rejang Lebong, Kab. Sukabumi, Kab. Sumedang, Kab. Tanggamus, Kota Bandung
3 (5)	Kab. Kep. Mentawai, Kab. Lampung Selatan, Kota Banda Aceh, Kota Bengkulu, Kota Padang

We applied the hybrid algorithm of k-mean and hierarchies with the number of clusters $k = 3$ according to the anticipated level of natural disaster mitigation that has been defined. The clustering results, shown in Figure 9 and Table 12, indicate that most of all data are grouped in cluster 1. There are only two data in cluster 2 and one data in cluster 3.

Figure 9

Cluster Dendrogram of Hybrid KM Hierarchy Clustering on Keywords.

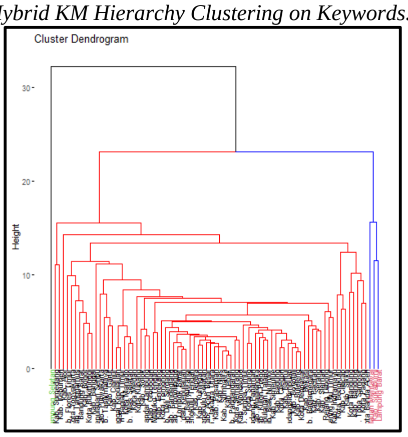


Table 12

Result of Hybrid KM Hierarchy Clustering on Keywords.

Cluster (Count)	Region
1 (77)	Kota Padangsidimpuan, Kab. Simeulue, Kota Padang, Kab. Sumedang, Kab. Kebumen, Kab. Tanggamus, Kota Serang, Kota Bandung, Kab. Kepulauan Talaud, Kab. Bandung, Kab. Purwakarta, Kab. Cilacap, Kab. Badung, Kota Semarang, Kab. Garut, Kab. Toba Samosir, Kab. Biak Numfor, Kab. Cianjur, Kab. Buleleng, Kab. Probolinggo, Kab. Lombok Barat, Kab. Pangkajene dan Kepulauan, Kota Banda Aceh, Kota Adm. Jakarta Utara, Kab. Tasikmalaya, Kab. Banyuwangi, Kota Denpasar, Kab. Wonosobo, Kab. Lebak, Kab. Bandung Barat, Kab. Ogan Komering Ulu Timur, Kab. Pandeglang, Kab. Banjarnegara, Kab. Kulon Progo, Kota Cilegon, Kab. Serang, Kab. Indramayu, Kab. Maluku Tengah,

	Kab. Flores Timur, Kab. Administrasi Kepulauan Seribu, Kab. Sumba Timur, Kota Manado, Kab. Solok, Kab. Nias Selatan, Kab. Bogor, Kota Tomohon, Kab. Pacitan, Kab. Alor, Kab. Majalengka, Kota Kupang, Kab. Aceh Besar, Kab. Kuningan, Kota Bandar Lampung, Kab. Pangandaran, Kota Bengkulu, Kab. Bantul, Kab. Sumbawa, Kab. Nias, Kota Yogyakarta, Kab. Boyolali, Kab. Rejang Lebong, Kab. Sleman, Kab. Bengkulu Selatan, Kota Bukittinggi, Kab. Sikka, Kab. Banggai Kepulauan, Kab. Banggai, Kab. Grobogan, Kab. Jepara, Kota Medan, Kab. Bengkulu Tengah, Kab. Pasaman, Kab. Seram Bagian Timur, Kab. Enrekang, Kab. Padang Pariaman, Kab. Bekasi, Kab. Deli Serdang
2 (3)	Kab. Kepulauan Mentawai, Kab. Sukabumi, Kab. Lampung Barat
3 (1)	Kab. Lampung Selatan

Figure 10 illustrates the keywords that represent each mitigation category. For category A, which is construction and strengthening of building structures, the keywords that appeared most often include earth movements, pressure, and earthquakes. Category B is for mapping of disaster-prone areas, and the keyword that appeared the most is fault. Then category C is for assessment of disaster risk and characteristics, and the keywords that appeared are earthquakes and tectonic plates. While in category D, which is for preparation and installation of early warning system instrumentation, keywords such as deformation and earthquake fault appeared. Category E is for planning and implementation of spatial planning, and the keyword that most often appeared is earthquakes. Finally, in the F category, which is for = outreach and information dissemination, the keyword that often appears is disaster.

Figure 10
Keyword Word Cloud on Mitigation Category.

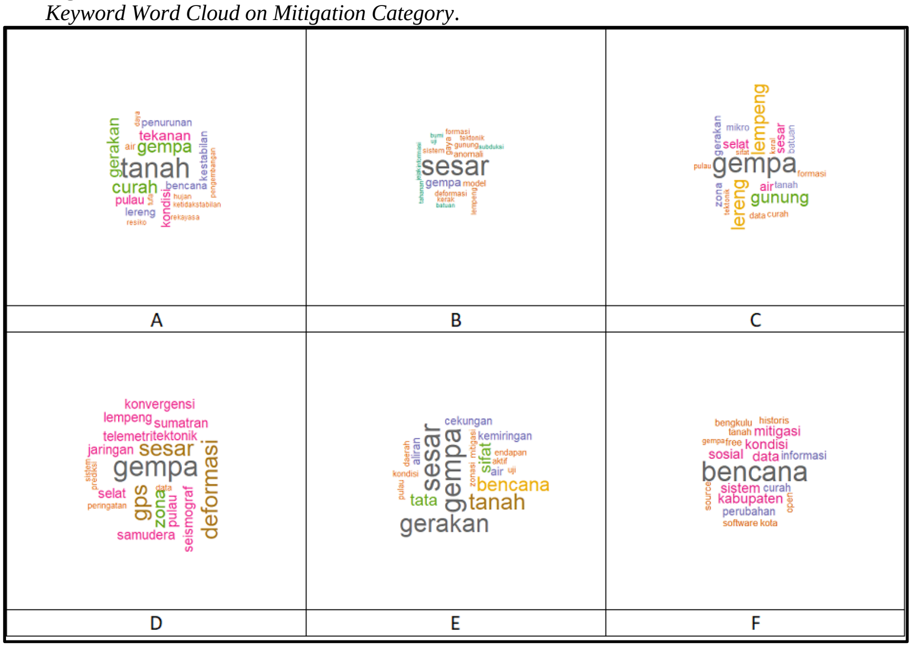


Figure 11 and Table 13 show the result of hybrid KM HC on disaster types. In contrast with the previous clustering results on the keywords, the clustering on the disaster types resulted in more regions fall into clusters 2 and 3.

Figure 11

Cluster Dendrogram of Hybrid KM Hierarchy Clustering on Disaster Types.

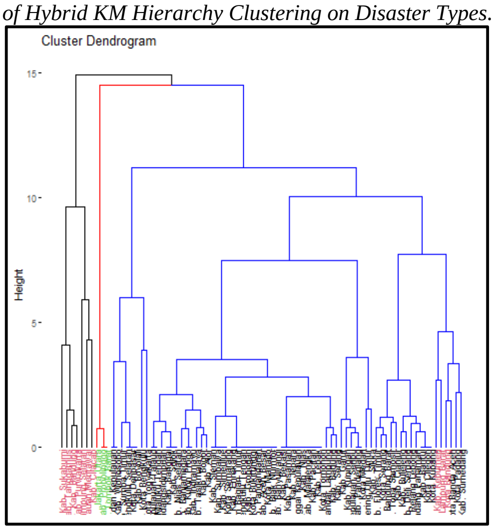


Table 13

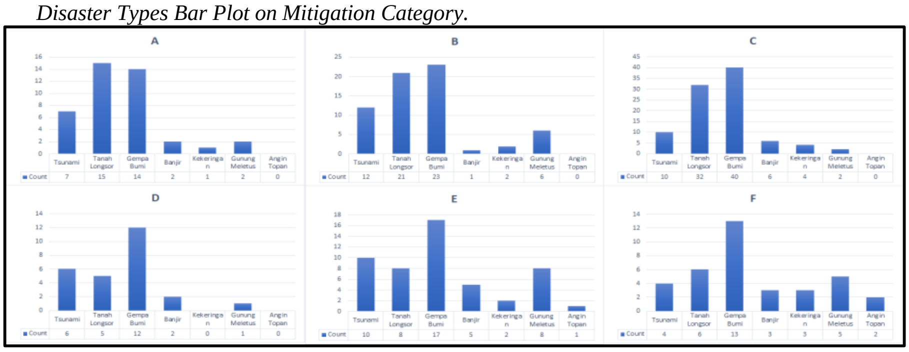
Result of Hybrid KM Hierarchy Clustering on Disaster Types.

Cluster (Count)	Region
1 (68)	Kota Padangsidimpuan, Kab. Simeulue, Kota Padang, Kab. Sumedang, Kota Serang, Kab. Kepulauan Talaud, Kab. Badung, Kota Semarang, Kab. Garut, Kab. Toba Samosir, Kab. Biak Numfor, Kab. Cianjur, Kab. Buleleng, Kab. Pangkajene dan Kepulauan, Kota Banda Aceh, Kab. Tasikmalaya, Kab. Banyuwangi, Kota Denpasar, Kab. Wonosobo, Kab. Lebak, Kab. Bandung Barat, Kab. Ogan Komering Ulu Timur, Kab. Pandeglang, Kab. Banjarnegara, Kab. Kulon Progo, Kota Cilegon, Kab. Serang, Kab. Indramayu, Kab. Maluku Tengah, Kab. Flores Timur, Kab. Administrasi Kepulauan Seribu, Kab. Sumba Timur, Kota Manado, Kab. Solok, Kab. Nias Selatan, Kab. Bogor, Kota Tomohon, Kab. Pacitan, Kab. Alor, Kab. Majalengka, Kota Kupang, Kab. Aceh Besar, Kab. Kuningan, Kota Bandar Lampung, Kab. Pangandaran, Kota Bengkulu, Kab. Bantul, Kab. Sumbawa, Kab. Nias, Kota Yogyakarta, Kab. Boyolali, Kab. Rejang Lebong, Kab. Sleman, Kab. Bengkulu Selatan, Kota Bukittinggi, Kab. Sikka, Kab. Banggai Kepulauan, Kab. Banggai, Kab. Grobogan, Kab. Jepara, Kota Medan, Kab. Bengkulu Tengah, Kab. Pasaman, Kab. Seram Bagian Timur, Kab. Enrekang, Kab. Padang Pariaman, Kab. Bekasi, Kab. Deli Serdang
2 (10)	Kab. Kebumen, Kab. Tanggamus, Kota Bandung, Kab. Kepulauan Mentawai, Kab. Bandung, Kab. Purwakarta, Kab. Cilacap, Kab. Sukabumi, Kab. Lampung Selatan, Kab. Lampung Barat
3 (3)	Kab. Probolinggo, Kab. Lombok Barat, Kota Adm. Jakarta Utara

Figure 12 shows the correlation of each category of mitigation anticipation with different types of disasters. For example, in the mitigation category A, which is construction and strengthening of building structures, landslides are the most anticipated. While in category D, which is for preparation and installation of early warning system instrumentation, earthquake, and tsunami disasters are the most anticipated.

Figure 12

Disaster Types Bar Plot on Mitigation Category.



To validate the clustering results, we used a matching matrix on Keywords, Disaster Types, and Mitigation Code to determine purity to validate the clustering results. As shown in Table 14, the results show that the clusters have an Averaged TPu value of 0.88 for the Hybrid Clustering Algorithm, 0.84 for HC, and 0.86 for KM. Hence, the TPu value is close to 1, representing the acceptable results of the hybrid clustering algorithm. From this Table, we conclude that the hybrid clustering outperformed Standard KM and HC since the TPu value is the highest.

Table 14

Matching Matrix Validation on Keywords, Disaster Types, and Mitigation Code.

	Standard KM	HC	Hybrid
TPu on Keywords	0.81	0.81	0.82
TPu on Disaster Type	0.76	0.79	0.82
TPu on Mitigation Code	1	0.91	1
Average TPu	0.85	0.83	0.88

CONCLUSION

This study has researched to cluster the natural disaster literature dataset. We do the clustering process by applying the KM, hierarchical, and hybrid algorithms. This process produced three clusters for the anticipation level of natural disaster mitigation: Cluster 1 for low anticipation level, cluster 2 for medium anticipation level, and cluster 3 for high anticipation level. In addition, from validation by experts, the clustering results indicate that 67 districts/cities (82.7%) fall into cluster 1, 9 districts/cities (11.1%) are classed into cluster 2, and the remaining five districts/cities are categorized in cluster 3 (6.2%). From the analysis of the silhouette coefficient calculation, the hybrid algorithm can provide relatively homogeneous clustering results.

Furthermore, we used a matching matrix on keywords, disaster types, and mitigation code to determine purity to validate the clustering results. The clusters have a TPu close to 1, representing acceptable results of the hybrid clustering algorithm. We conclude that the hybrid clustering outperformed Standard KM and HC since the TPu value is the highest.

A further study that aims to compare the hybrid clustering algorithm with other algorithms is recommended. The method for determining the disaster mitigation level also needs improvement.

ACKNOWLEDGMENT

The authors gratefully acknowledge the use of service and facilities of the Research Centre for Informatics, Indonesian Institute of Sciences (LIPI). The authors would like to thank Dr. Eng. Budi Nugroho and Dr. Adrin Tohari for their mentorship, fruitful discussions, and valuable feedback.

REFERENCES

Ali, K., Nguyen, H. X., Vien, Q.-T., Shah, P., & Chu, Z. (2018). Disaster management using d2d communication with power transfer and clustering techniques. *IEEE Access*, *6*, 14643–14654. https://doi.org/10.1109/ACCESS.2018.2793532

Anjayani, E. (2008). *Indonesia di pertemuan lempeng tektonik*. Cempaka Putih.

Atasever, U. H. (2017). A new unsupervised change detection approach with hybrid clustering for

detecting the areal damage after natural disaster. *Fresenius Environmental Bulletin*, *26*(6), 3891–3896.
Bagirov, A. M., Ugon, J., & Webb, D. (2011). Fast modified global k-means algorithm for incremental cluster construction. *Pattern Recognition*, *44*(4), 866–876. https://doi.org/10.1016/j.patcog.2010.10.018.
Balavand, A., Kashan, A. H., & Saghaei, A. (2018). Automatic clustering based on crow search algorithm-kmeans (CSA-Kmeans) and data envelopment analysis (DEA). *International Journal of Computational Intelligence Systems*, *11*(1), 1322–1337. https://doi.org/10.2991/ijcis.11.1.98
Ediyanto, Mara, M. N., & Satyahadewi, N. (2013). Classification of characteristics using the k-means cluster analysis method. *Buletin Ilmiah Matematika Statistik dan Terapannya*, *2*(2), 133–136.
Govender, P., & Sivakumar, V. (2020). Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). *Atmospheric Pollution Research* (Vol. 11, Issue 1, pp. 40–56). https://doi.org/10.1016/j.apr.2019.09.009
Han J., & Kamber M. (2001). *Data mining: Concepts and techniques*. Morgan Kaufmann Publishers.
INDONESIA, P. R. (24 C.E.). *Undang-undang republik indonesia nomor 24 tahun 2007 tentang penanggulangan bencana*.
Kandel, A., Tamir, D., & Rishe, N. D. (2014). Fuzzy logic and data mining in disaster mitigation. *Improving Disaster Resilience and Mitigation - IT Means and Tools. NATO Science for Peace and Security Series C: Environmental Security*, *July 2014,* 167-186. https://doi.org/10.1007/978-94-017-9136-6_11
Kassambara, A., & Mundt, F. (2020). *Factoextra: Extract and visualize the results of multivariate data analyses*. https://cran.r-project.org/package=factoextra
Khazaal Abdulsahib, A., & Sakira Kamaruddin, S. (2015). Graph based text representation for document clustering, *Journal of Theoretical and Applied Information Technology*, *76*(1), 1-13.
Ng, K.-H., & Khor, K.-C. (2016). Evaluation on rapid profiling with clustering algorithms for plantation stocks on bursa Malaysia. *Journal of Information and Communication Technology*, *15*(2). https://doi.org/10.32890/jict2016.15.2.4
Nugroho, B. (2021). Perbandingan aplikasi algoritma kernel k-means pada graf bipartit dan k-means pada matriks dokumen-istilah dalam dataset penelitian covid-19 ristekbrin. *Jurnal Teknologi Informasi dan Ilmu Komputer*, *8*(2), 411–418.
Peterson, A. D., Ghosh, A. P., & Maitra, R. (2018). Merging k-means with hierarchical clustering for identifying general-shaped groups. *Stat*, *7*(1). https://doi.org/10.1002/sta4.172
Priatmodjo, D. (2011). Penataan kota bermuatan antisipasi bencana. *Nalars*, *10*(2), 83–104.
Prihandoko, & Bertalya. (2016). A data analysis of the impact of natural disasters using k-means clustering algorithm. *Jurnal Ilmiah KURSOR*, *8*(4), 169–174. www.bnpb.go.id.
Prihandoko, Bertalya, & Ramadhan, M. I. (2017). An analysis of natural disaster data by using k-means and k-medoids algorithm of data mining techniques. *15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering*, 221–225.
Rachmawati, L. (2018). People’s knowledge on hazard map and Merapi hazard mitigation. *Hinga*, *13*(2), 143–156.
Sadewo, M. G., Perdana Windarto, A., & Wanto, A. (2018). Penerapan algoritma clustering dalam mengelompokkan banyaknya desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam menurut provinsi dengan k-means. *Konferensi Nasional Teknologi Informasi dan Komputer*, *KOMIK 2018*, 311–319. http://ejurnal.stmik-budidarma.ac.id/index.php/komik
Samatova, N. F., Hendrix, W., Jenkins, J., Padmanabhan, K., & Chakraborty, A. (2013). *Practical graph mining with R*. CRC Press.
Supriyadi, B., Windarto, A. P., Soemartono, T., & Mungad. (2018). Classification of natural disaster-prone areas in Indonesia using k-means. *International Journal of Grid and Distributed Computing*, *11*(8), 87–98. https://doi.org/10.14257/ijgdc.2018.11.8.08
Welton-Mitchell, C., James, L. E., Khanal, S. N., & James, A. S. (2018). An integrated approach to mental health and disaster preparedness: a cluster comparison with earthquake affected communities in Nepal. *BMC Psychiatry*, *18*. https://doi.org/10.1186/s12888-018-1863-z

Wen, L.-H., Shi, Z.-H., & Liu, H.-Y. (2019). Research on risk assessment of natural disaster based on cloud fuzzy clustering algorithm in Taihang Mountain. *Journal of Intelligent & Fuzzy Systems*, *37*(4), 4735–4743. https://doi.org/10.3233/JIFS-179308

Yana, M. S., Setiawan, L., Ulfa, E. M., Rusyana, A., Statistika, J., Kuala, S., & Aceh, B. (2018). Penerapan metode k-means dalam pengelompokan wilayah menurut intensitas kejadian bencana alam di Indonesia tahun 2013-2018. *Journal of Data Analysis, 11(2)*.

Web of Science™

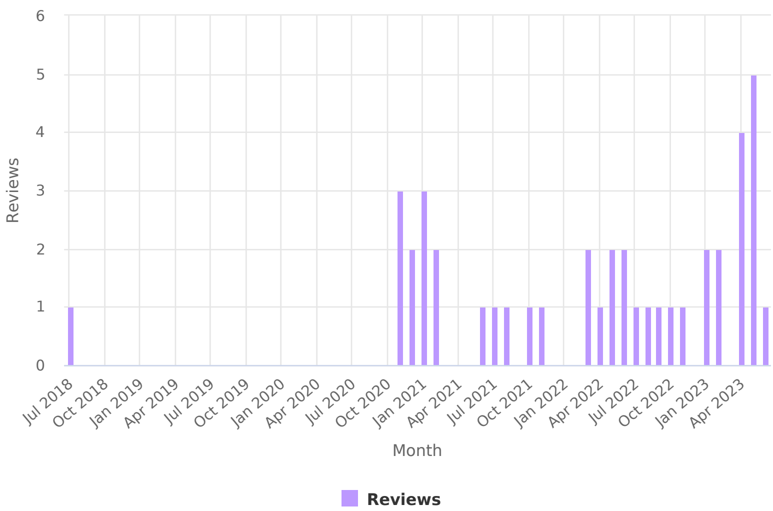
Web of Science CV
Prepared on June 10th 2023

Abdullah Bin Husin

https://www.webofscience.com/wos/author/rid/AAK-6598-2020
Web of Science ResearcherID: AAK-6598-2020
ORCiD: 0000-0001-5169-7678

Verified reviews

Review Summary



Reviewer Summary

For manuscripts reviewed from date range June 2018 - June 2023

(38) Baghdad Science Journal	(2) Journal of Information and Communicati…
(1) Journal of Basic and Applied Research In…	(1) Journal of Advances in Mathematics and …

42 REVIEWS OF 35 MANUSCRIPTS
For manuscripts published from date range June 2018 - June 2023

Yolo Yolo: A Competitive Analysis Of Modern Object Detection Algorithms For Road Defects Detection Using Drone Images
Reviewed: Jun 2023 for Baghdad Science Journal

A Carbon Monoxide Detector Safety System Design and Analysis of Vehicle Using IoT
Reviewed: May 2023 for Baghdad Science Journal

Alexnet Convolutional Neural Network Architecture With Cosine And Hamming Similarity/distance Measures For Fingerprint Biometric Matching
Reviewed: May 2023 for Baghdad Science Journal

Hetero-associative Memory Based New Iraqi License Plate Recognition
Reviewed: May 2023 for Baghdad Science Journal

Flamingo Search Algorithm for Aircraft Landing Scheduling
Reviewed: May 2023 for Baghdad Science Journal

Wavelet Scattering Network for Classifying Three Stages of Cataract Disease
Reviewed: May 2023 for Baghdad Science Journal

العراق -كربلاء - الوند
Reviewed: Apr 2023 for Baghdad Science Journal

Fraudulent Taxi Driver Detecting Based on Priority Factors and Density Based Clustering Approach
Reviewed: Apr 2023 for Baghdad Science Journal

Development of an Algorithm for Diagnosing Heart Disease Based on Cardiac Signals
2 rounds from Feb 2023 to Apr 2023 for Baghdad Science Journal

Efficient Approach for the Localization of Copy-move Forgeries Using Pointrend with Regnetx
2 rounds from Feb 2023 to Apr 2023 for Baghdad Science Journal

Artificial Neural Network and Latent Semantic Analysis for Adverse Drug Reaction Detection
2 rounds from Jan 2023 to Jan 2023 for Baghdad Science Journal

Evaluating the Influence of Majority Voting – Arithmetic Average for Feature Selection in Detecting Network Intrusion Utilizing Machine Learning Approaches
Reviewed: Nov 2022 for Baghdad Science Journal

Improving Wireless Sensor Network Security Using Quantum Key Distribution
2 rounds from Sep 2022 to Oct 2022 for Baghdad Science Journal

Proposed Framework for Official Document Sharing and Verification in E-government Environment Based on Blockchain Technology
Reviewed: Aug 2022 for Baghdad Science Journal

Resolving Signal Transmission, Coverage and Connectivity Issues in 3d-plane Using Airborne Wireless Network
Reviewed: Jul 2022 for Baghdad Science Journal

Impact of Denial-of-Service Attack on Directional Compact Geographic Forwarding Routing Protocol in Wireless Sensor Networks
Reviewed: Jun 2022 for Baghdad Science Journal

Using VGG Models with Intermediate Layer Feature Maps for Static Hand Gesture Recognition
Reviewed: Jun 2022 for Baghdad Science Journal

A Multi-hop Routing Protocol for an Energy-efficient in Wireless Sensor Network
Reviewed: May 2022 for Baghdad Science Journal

Modeling the Power Grid Network in Iraq
2 rounds from Apr 2022 to May 2022 for Baghdad Science Journal

A Security and Privacy Aware Computing Approach on Data Sharing in Cloud Environment
Reviewed: Mar 2022 for Baghdad Science Journal

Honeyword Generation Using a Proposed Discrete Salp Swarm Algorithm
Reviewed: Mar 2022 for Baghdad Science Journal

A Hybrid K-Means Hierarchical Algorithm for Natural Disaster Mitigation Clustering
Reviewed: Nov 2021 for Journal of Information and Communication Technology

Covid-19 Recognition Using Spectral and Statistical Analysis of Cough Recordings Based on The Combination of Svd and Dwt
Reviewed: Oct 2021 for Baghdad Science Journal

Distributed Heuristic Algorithm for Migration and Replication of Self-organized Services in Future Networks
2 rounds from Jul 2021 to Aug 2021 for Baghdad Science Journal

Influence of Cold Plasma on Sesame Paste and the Nano Sesame Paste Based on Co-occurrence Matrix
Reviewed: Jun 2021 for Baghdad Science Journal

Advanced GIS-based Multi-Function Support System for Identifying the Best Route
Reviewed: Feb 2021 for Baghdad Science Journal

Developing French Teaching Method Based on Google Classroom for Class Xi at Sma Negeri 21 Medan
Reviewed: Feb 2021 for Journal of Basic and Applied Research International

Tourism Companies Assessment via Social Media Using Sentiment Analysis
2 rounds from Nov 2020 to Jan 2021 for Baghdad Science Journal

WOAIP: Wireless Optimization Algorithm for Indoor Placement Based on Binary Particle Swarm Optimization (BPSO)
Reviewed: Jan 2021 for Baghdad Science Journal

Touchscreen-based Smartphone Continuous Authentication System (SCAS) Using Deep Neural Network
Reviewed: Jan 2021 for Baghdad Science Journal

Performance Assessment of Solar-Transformer-Consumption System Using Neural Network Approach
Reviewed: Dec 2020 for Baghdad Science Journal

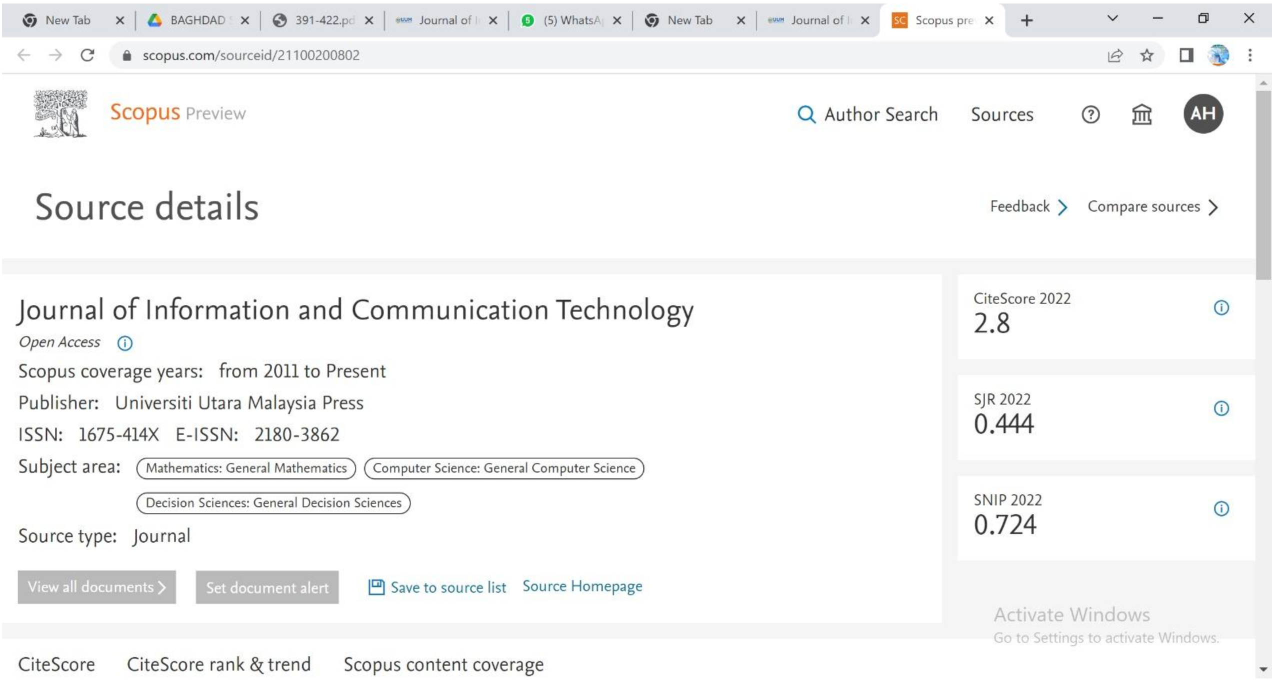
Constructing a Software Tool for Detecting Face Mask-wearing by Machine Learning
Reviewed: Dec 2020 for Baghdad Science Journal

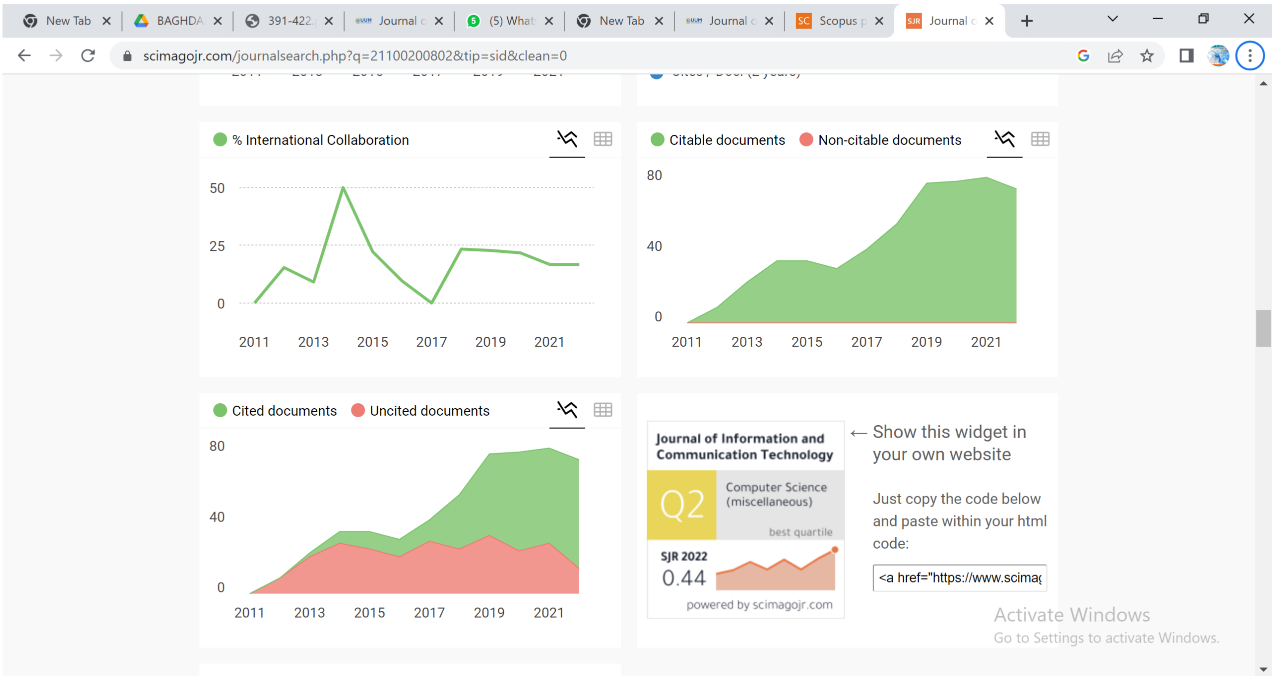
Lion Optimized K-Means Support Vector Machine for Clustering Problems in Cloud Internet of Things Environment
Reviewed: Nov 2020 for Baghdad Science Journal

Fuzzy Unordered Rule Using Greedy Hill Climbing Feature Selection Method: An Application to Diabetes Classification
Reviewed: Nov 2020 for Journal of Information and Communication Technology

The impact of standard and quality in e-learning system
Reviewed: Jul 2018 for Journal of Advances in Mathematics and Computer Science


e-journal.uum.edu.my/index.php/jict
Register
Login
Journal of ICT
Journal of Information and Communication Technology
Announcements
Forthcoming Articles
Current
Archives
About
Search
CALL FOR PAPERS: SINTOK INTERNATIONAL CONFERENCE ON SOCIAL SCIENCE AND MANAGEMENT (SICONSEM) 2023
30-04-2023
Greetings from Sintok!
We cordially invite you to participate in The 5th Sintok International Conference on Social Science and Management (SICONSEM) 2023 that will be conducted virtually on 17 & 18 October 2023.
SICONSEM 2023
Organised by
UUM PRESS
Universiti Utara Malaysia
17-18 OCTOBER 2023/ VIRTUAL CONFERENCE
WEBSITE: http://uumpress.uum.edu.my/siconsem/ | FB: https://www.facebook.com/siconsem2023/
CiteScoreTracker 2021
3.2 = 316 Citation to date / 99 Documents to date
Powered by Scopus
Scopus coverage: 2011 to Present
JICT Template
Make a Submission


scopus.com/sourceid/21100200802
Scopus Preview
Author Search
Sources
Source details
Feedback
Compare sources
Journal of Information and Communication Technology
Open Access
Scopus coverage years: from 2011 to Present
Publisher: Universiti Utara Malaysia Press
ISSN: 1675-414X E-ISSN: 2180-3862
Subject area: Mathematics: General Mathematics
Computer Science: General Computer Science
Decision Sciences: General Decision Sciences
Source type: Journal
View all documents
Set document alert
Save to source list
Source Homepage
CiteScore 2022
2.8
SJR 2022
0.444
SNIP 2022
0.724
CiteScore
CiteScore rank & trend
Scopus content coverage


scimagojr.com/journalsearch.php?q=21100200802&tip=sid&clean=0
% International Collaboration
Citable documents
Non-citable documents
Cited documents
Uncited documents
Journal of Information and Communication Technology
Q2
Computer Science (miscellaneous)
best quartile
SJR 2022
0.44
powered by scimagojr.com
Show this widget in your own website
Just copy the code below and paste within your html code:

A Hybrid K-Means Hierarchical Algorithm for Natural Disaster Mitigation Clustering

By: Prasetyadi, A (Prasetyadi, Abdurrakhman) ; Nugroho, B (Nugroho, Budi) ; Tohari, A (Tohari, Adrin)

View Web of Science ResearcherID and ORCID (provided by Clarivate)

JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY-MALAYSIA

Volume: 21 **Issue:** 2 **Page:** 175-200

DOI: 10.32890/jict2022.21.2.2

Published: APR 2022

Indexed: 2022-05-03

Document Type: Article

Abstract:

Cluster methods such as k-means have been widely used to group areas with a relatively equal number of disasters to determine areas prone to natural disasters. Nevertheless, it is difficult to obtain a homogeneous clustering result of the k-means method because this method is sensitive to a random selection of the centers of the cluster. This paper presents the result of a study that aimed to apply a proposed hybrid approach of the combined k-means algorithm and hierarchy to the clustering process of anticipation level datasets of natural disaster mitigation in Indonesia. This study also added keyword and disaster-type fields to provide additional information for a better clustering process. The clustering process produced three clusters for the anticipation level of natural disaster mitigation. Based on the validation from experts, 67 districts/cities (82.7%) fell into Cluster 1 (low anticipation), nine districts/cities (11.1%) were classified into Cluster 2 (medium), and the remaining five districts/cities (6.2%) were categorized in Cluster 3 (high anticipation). From the analysis of the calculation of the silhouette coefficient, the hybrid algorithm provided relatively homogeneous clustering results. Furthermore, applying the hybrid algorithm to the keyword segment and the type of disaster produced a homogeneous clustering as indicated by the calculated purity coefficient and the total purity values. Therefore, the proposed hybrid algorithm can provide relatively homogeneous clustering results in natural disaster mitigation.

Keywords

Author Keywords: Clustering; Hybrid; K-means; Mitigation; Natural disaster

Addresses:

1 Natl Res & Innovat Agcy, Res Ctr Informat, Jakarta, Indonesia

2 Natl Res & Innovat Agcy, Res Ctr Geotechnol, Jakarta, Indonesia

Categories/ Classification

Research Areas: Computer Science

Citation Topics: 4 Electrical Engineering, Electronics & Computer Science > 4.61 Artificial Intelligence & Machine Learning > 4.61.869 Clustering

Web of Science Categories: Computer Science, Information Systems

+ See more data fields

If you would like to improve the quality of the data in this record, please Suggest a correction