Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14120
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Comparison of K-Means and Self-Organizing Map Algorithms for Ground Acceleration Clustering

Siska Simamora[1], Muhammad Iqbal[2], Andysah Putera Utama Siahaan[3], Khairul[4], Zulham Sitorus[5]
[1,2,3,4,5)] Prodi Magister Teknologi Informasi, Universitas Panca Budi, Sumatera Utara, Indonesia
[1)]siskasimamora36@gmail.com, [2)] wakbalpb@yahoo.co.id, [3)] andiesiahaan@gmail.com
[4)] khairul@dosen.pancabudi.ac.id , [5)] zulhamsitorus@gmail.com

Submitted : Sept 1, 2024 | **Accepted** : Sept 30, 2024 | **Published** : Oct 7, 2024

Abstract: This study evaluates earthquake-induced ground acceleration in Indonesia, which is located in the Pacific Ring of Fire zone, using Donovan's empirical method and comparing two clustering algorithms, the organizing Map (SOM) and K-Means. The main problem faced is the high risk of earthquakes in Indonesia and the need for effective methods to predict potential damage to buildings and infrastructure. The research objective is to evaluate earthquake-induced ground acceleration and identify acceleration distribution patterns using clustering techniques. The solution methods used include the application of the Donovan method to calculate ground acceleration based on BMKG data, as well as the use of SOM and K-Means algorithms to cluster the ground acceleration data. GIS and Python applications are used to visualize the clustering results. The results show that the Donovan method integrated with SOM and K-Means provides significant insights into the distribution of ground acceleration, thus assisting in risk evaluation, disaster mitigation planning, and the development of more effective earthquake-resistant infrastructure development strategies in Indonesia

Keywords: Ground acceleration; Donovan;clustering, SOM, K-Means

INTRODUCTION

Due to its location in the Pacific Ocean's fifth zone, which is marked by ongoing seismic activity and active volcanic activity, Indonesia is a country that frequently suffers earthquakes. (Hermon, 2015). Based on data from the Bureau of Meteorology, Climate, and Geophysics (BMKG), the year 2021 saw approximately 11.177 major earthquakes that have occurred in Indonesia. These include the Aceh earthquake in 2004 with a magnitude of 9.1, which caused a tsunami that caused numerous human casualties, and the Yogyakarta earthquake in 2006 with a magnitude of 6.3, which also caused numerous human casualties and building failures. (Tjandra, 2018), As an example, the 2018 Lombok earthquake had a 7 scalar Richter earthquake, which resulted in many people becoming uneasy and numerous infrastructure failures. (Azmiyati & Rancak, 2023).

The impact of an earthquake causes damage to buildings, infrastructure, and causes casualties also has significant economic and social impacts (Purwanto et al., 2023). With these events, communities in Indonesia need to increase awareness and prepare for earthquakes by conducting simulations and evacuation training (Paramesti, 2011), also affects significant ground acceleration, which is one of the important parameters in evaluating the potential damage to building structures and other infrastructure during an earthquake. Ground acceleration is the change in ground velocity at a location during an earthquake, and is measured in units of acceleration such as meters per second squared (m/s^2) or gal (1 gal is equivalent to 1 cm/s^2) (Amanullah, 2022).

Due to high ground accelerations during an earthquake, building structures and other infrastructure can be damaged or even collapse. (Kusumaningrum, 2017). High ground acceleration can also cause cracks and damage to the ground around buildings, and can trigger landslides and ground shifts. (Hidayat & Munir, 2018). (Artati et al., 2020) discusses the determination of maximum ground acceleration, seismic vulnerability index, and soil shear strain in the Jayapura City area. Peak ground acceleration (PGA) is the value of ground acceleration that has occurred in a certain area and within a certain period of time due to earthquake vibrations. (Pasau et al., 2018). The maximum ground acceleration parameter or PGA is influenced by three factors, namely the seismic source, raypath and possible geological conditions or regional characteristics that become local factors. This shows that there are differences in PGA values in an area related to differences in geological characteristics. (Donovan & Mickey, 2018).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14120
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Earthquakes have a significant impact on ground acceleration, which is a key parameter in assessing potential damage to buildings and infrastructure. Maximum ground acceleration (PGA) reflects the strength of vibration and is a key indicator of earthquake intensity in a region. Variations in ground acceleration between regions are influenced by local geological characteristics, so careful evaluation is necessary for risk mitigation, designing earthquake-resistant buildings, and understanding the potential impacts of earthquakes. To evaluate the ground acceleration, the researcher utilized the Donovan empirical method, which is one of the techniques to determine the PGA value in an area with a subduction tectonic pattern. Donovon's empirical method has been used by researchers such as (Suhada et al., 2023) said that Donovan's empirical equation is close to the PGA value from the BMKG Shakemap results by providing important information related to the potential for building damage due to earthquakes using the PGA value as an indicator . (WINDA, 2023) mapping areas with the risk of maximum ground acceleration values and analyzing the compatibility between empirical values calculated from accelerograph data. (Kapojos et al., 2015) provides important information related to mapping the distribution of maximum ground acceleration that can be used in the planning and development of resistant infrastructure.

Based on these studies, Donovan's empirical method is a useful tool in earthquake risk evaluation, providing critical information for planning and building infrastructure that is safe and resistant to ground shaking. After obtaining the results of the ground acceleration evaluation, the next step in this research is to cluster the ground acceleration by comparing two algorithms related to clustering. The basic problem of clustering is how to divide a set of data that has as similar as possible into one cluster. (Imani et al., 2023). The algorithms to be compared in clustering are SOM (Self Organizing Map) and K-Means. The K-Means Clustering algorithm provides an overview of the mapping of parameters of visitor satisfaction levels based on sample data from around 35 visitors with the results The clustering analysis reveals that the level of satisfaction at Cikundul Hot Spring is a weak point, especially in parameters such as bathrooms, prices, swimming pools, and facilities. (Hasugian et al. 2021) analyzed the village status groups using the K-means algorithm and provided the best cluster information using the Elbow Method by finding the SSE (sum of squared errors) value by utilizing the cluster closest to the elbow on the graph. The result obtained from this research is a comparison of data from the village ministry with cluster information provided by the K-Means algorithm.

LITERATUR RIVIEW

Research on maximum ground acceleration (PGA) and clustering methods in Indonesia shows a variety of approaches in evaluating earthquake potential and social patterns. Mahendra Taruna et al. found the Kanno attenuation equation to be the most reliable for predicting PGA in Mataram City, although it is recommended to develop a more region-specific model for Indonesia (Mahendra Taruna et al., 2020). (Sungkowo, 2018) and (Saputra et al., 2020) used accelerograph data and the PSHA method to analyze earthquake risk in Surakarta and Riau, whereas (Romadiana et al., 2018) mapping PGA in Manado and West Sumatra using empirical methods such as Donovan, McGuire, and Midorikawa. Other research, such as by (Romadiana et al., 2018) using clustering methods (SOM, K-Means) to cluster social data in East Nusa Tenggara and research themes at Universitas Sebelas Maret. (Mulyani et al., 2022) applied clustering to assess the spread of COVID-19 and voter data, while (Munawar & Kunci, 2015) grouping students based on project performance. Overall, this study highlights the importance of complete and quality data for more accurate earthquake and social analysis, with contributions to risk mitigation and policy making in Indonesia.

METHOD

This chapter explores the methodological framework used to evaluate maximum ground acceleration and its clustering characteristics as a result of seismic activity. A well-conceived and rigorous research method is of key importance in order to collect accurate data, and ensure the validity of the interpretation of this information. This research aims to gain a deeper understanding of the distribution of earthquake-induced ground accelerations, which is one of the critical parameters in geotechnical analysis and earthquake engineering. Utilizing an innovative and multidisciplinary approach, the research will identify and assess the variables that affect ground acceleration, including surface geological characteristics and underlying ground conditions. This will enable researchers and practitioners to design systems that are more resilient to the damaging effects of earthquakes, as well as contribute to the development of more effective seismic risk mitigation strategies. In order to improve the effectiveness and quality of research results, systematic and methodological work steps are essential. The research work steps taken are outlined in Figure 1. which is the procedure carried out. This allows for transparency in the research process as well as ensuring that each phase can be perfectly tracked and analyzed.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14120
e-ISSN : 2541-2019
p-ISSN : 2541-044X

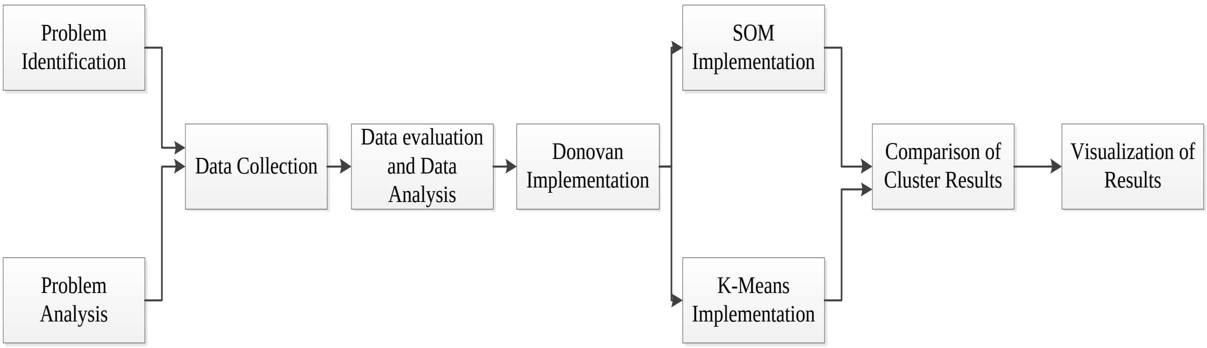

Fig 1. Research Work Steps

This study evaluates ground acceleration due to earthquakes in Indonesia using Donovan's empirical method and compares Self-Organizing Maps (SOM) and K-Means clustering algorithms. The main focus of the research is on analyzing the impact of ground acceleration on structural safety and regional development. Secondary data from BMKG and USGS were used to obtain information about the earthquake, such as magnitude and epicenter location. The data were then evaluated to ensure consistency and completeness before applying the Donovan method to calculate the maximum ground acceleration. Next, the SOM algorithm is used to analyze the ground acceleration pattern by mapping the data into clusters, while K-Means is applied for efficient clustering. The results of both clustering methods were compared to assess the advantages and disadvantages of each. Testing and validation are conducted to ensure the accuracy of the results, and the visualization results in the form of thematic maps provide in-depth insights into the distribution of ground acceleration and support earthquake risk mitigation strategies and the development of safer areas.

RESULT

Evaluation of Ground Acceleration with Donovan

After the earthquake data is generated, the ground acceleration calculation process is carried out using the Donovan emperical method, this process produces a ground acceleration value called PGA in Gal Units. PGA is an important parameter in earthquake characterization, indicating the magnitude of the maximum ground acceleration that occurs during an earthquake. PGA is measured in Gal units (1 Gal = 1 cm/s²) or in g, where 1 g is the acceleration due to the Earth's gravity (approximately 980 Gal). These PGA values are important in earthquake engineering and are used to assess the design requirements and resistance of structures to earthquake shaking. With the ground acceleration results in table 1

Table 1. Donovan's Empirically Based PGA Values

No	Earthquake location	Depth	Mag	Dmin	PGA
1	175 km W of Bengkulu	10187	5,6	2366	85,8272392
2	89 km W of Ambon	11326	4,6	2683	45,21331755
3	153 km S of Boyolangu	29,51	4,5	2634	302,7949117
4	99 km E of Kendari	9221	4,8	3743	61,4175169
5	206 km WSW of Abepura	54874	4,9	1945	6,79338988
…	…	…	…	…	…
23	224 km S of Trenggalek	10	4,5	3189	235,7811422
24	Banda Sea	9756	4,4	1989	50,25142279
25	96 km NE of Tobelo	37369	4,6	5,66	9,714682105

Based on the PGA calculation using the Donovan method, adjustments were made to the BMKG rules for data grouping. The test data consisting of 1549 samples is then visualized in the following graph. This visualization provides a clearer picture of the distribution and characteristics of the PGA data, facilitating further analysis in understanding patterns and trends.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14120
e-ISSN : 2541-2019
p-ISSN : 2541-044X

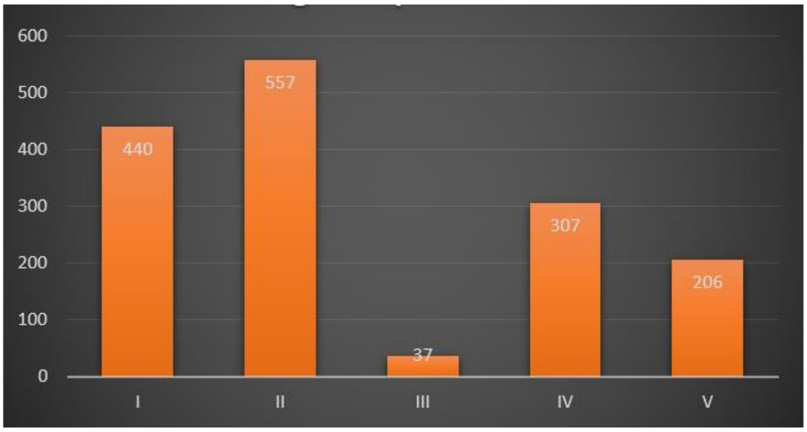

Fig 2. Ground Acceleration Grouping Chart

The clustering of ground acceleration based on figure 2 above shows the variation in the amount of data spread across the five clusters.Cluster 1 has the largest amount of data at 440, followed by Cluster 2 with 557 data, while Cluster 3 has the smallest amount of data at 36. Cluster 4 and Cluster 5 consist of 307 and 206 data respectively.

Data Clustering with SOM

In this stage, the clustering of ground acceleration is performed using the Self-Organizing Map (SOM) algorithm.The SOM algorithm, which is known as one of the methods in artificial neural networks, is used to identify hidden patterns in the ground acceleration data. The use of SOM Algorithm with description

Initialize as many input neurons x1, x2, x3, ..., xn

As input neurons in this clustering process, 25 training data are used which are expressed by x1 to X25 with the following description:

Table 2. Input neuron initialization

No	Earthquake location	Y1	Y2	Y3	Y4
1	175 km W of Bengkulu	10187	5,6	2366	85,8272392
2	89 km W of Ambon	11326	4,6	2683	45,21331755
3	153 km S of Boyolangu	29,51	4,5	2634	302,7949117
4	99 km E of Kendari	9221	4,8	3743	61,4175169
5	206 km WSW of Abepura	54874	4,9	1945	6,79338988
…	…	…	…	…	…
23	224 km S of Trenggalek	10	4,5	3189	235,7811422
24	Banda Sea	9756	4,4	1989	50,25142279
25	96 km NE of Tobelo	37369	4,6	5,66	9,714682105

Initialize as many output neurons as y1, y2, y3, ..., yn

Initialization of 5 output neurons randomly selected according to the needs in clustering, aims to ensure that each initial cluster formed has the same opportunity to develop and avoid initial bias in the artificial neural network learning process, in this test 5 data clusters were determined with the following description:

Table 3. Output neuron initialization

Selected Data	Earthquake location	Depth	Mag	Dmin	PGA	No
6	15 km N of Lubuklinggau	204591	4,2	1202	0,843758	I
1	175 km W of Bengkulu	10187	5,6	2366	85,82724	II
22	133 km SE of Bitung	35	4,5	5244	122,7774	III
16	150 km NNW of Kendari	10	5	3235	297,1229	IV
21	94 km WNW of Sinabang	18,06	5,9	2486	657,6668	V

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14120
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Determination of the Nearest Distance of Output Neuron to Input Neuron

The Clustering Formation process using SOM requires the closest distance from each output neuron to the input data. The scenario used in determining the distance by testing each data against the output neuron or in this case is expressed as a weight. Based on the distance calculation, the first data (X1) with BMU (Best Matching Unit) which is the neuron whose weight is closest to the input vector given in this case BMU on d2 with the lowest value of 8.00×10-7 so that x_(1 (data1)) enters the second position category. Distance calculations will continue until the entire data occupies the cluster provided. The results in table 4 below

Table 4. Clustering Results Using the SOM Algorithm

NO	Distance_n1	Distance_n2	Distance_n3	Distance_n4	Distance_n5	Posisi
1	194407,50	8E-07	10552,12559	10216,21927	10185,712	2
2	193270,67	1182,98795	11578,05762	11332,25567	11326,226	2
3	204566,7	10730,17261	2616,206499	601,3435472	384,67014	5
4	195386,53	1960,234533	9308,026758	9228,008565	9282,4381	2
5	149718,84	44901,02456	54938,2633	54879,93155	54855,705	2
23	315643,57	133702,8152	2058,256506	72,82997905	744,76647	4
24	305892,29	123938,7712	10251,73771	9819,278516	9763,5003	5
25	278280,93	96318,61181	37699,8746	37493,38634	32559,781	5

Establishment of Weight Update for BMU (Neuron 2)

After determining the BMU, the neuron's weight is updated to be closer to the input vector. This process helps the network learn patterns from the input data which are the weights after the First iteration with the results of each phase with the following description:

Table 5. Weight Determination Results for Each Iteration

NO	Before				After			
	W1	**W2**	**W3**	**W4**	**W1**	**W2**	**W3**	**W4**
1	10187	5,6	2366	85,827	133686,18	4,5375976	530,588	2,90300
2	10187	5,6	2366	85,8272	133686,18	4,5375976	530,588	2,90300
3	18,06	5,9	2486	657,66	21128,730	4,7875	1130,83	159,657
4	10756,5	5,1	2524,5	65,5202	133686,18	4,5375976	530,588	2,90306
5	9988,75	4,95	3133,75	63,4688	133686,18	4,5375976	530,588	2,90306
23	10	4,9	3145	293,816	10	4,7	3167	264,795
24	20,922	5,55	2523	568,948	21128,730	4,7875	1130,83	159,654
25	4888,46	4,975	2256	309,600	21128,730	4,7875	1130,83	159,657

Data Clustering with K-Means

At this stage, soil acceleration clustering is performed using the K-Means Algorithm. K-Means algorithm is one of the methods in machine learning that aims to divide data into a number of groups (clusters) based on similar characteristics.

Determination of Number of Clusters and Center Point

The process of forming clustering with the K-Means algorithm begins with determining the number of clusters, in this case it will be grouped into 5 clusters. Determining the number of clusters is one of the working steps of the method that has been determined, the next step is to randomly select the initial centroid from the dataset. These points will be the center of each cluster that will be formed. This process is crucial because randomly selected centroids can affect convergence and the final result of clustering which is outlined in the following table:

Table 6 Selected Cluster Centers

No Data	Earthquake location	Depth	Mag	Dmin	PGA	No
6	15 km N of Lubuklinggau	204591	4,2	1202	0,843758	I
1	175 km W of Bengkulu	10187	5,6	2366	85,82724	II
22	133 km SE of Bitung	35	4,5	5244	122,7774	III
16	150 km NNW of Kendari	10	5	3235	297,1229	IV
21	94 km WNW of Sinabang	18,06	5,9	2486	657,6668	V

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14120
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Distance Determination

After the initial center point is determined, the K-Means algorithm will work in an iterative manner. At each iteration, the data will be clustered based on its proximity to the nearest center point, which is calculated using the Euclidean distance. Once the clustering is complete, a new center point will be calculated as the average of all data points within each cluster. This process will continue to repeat until the center point no longer changes significantly or the maximum number of iterations is reached. The end result of this process is the formation of optimal clusters, where each data point is in the cluster with the closest center, thus minimizing the variance within each cluster.

Table 7. Clustering at Iteration I with K-Means

No	C1	C2	C3	C4	C5	Closest Distance
1	194407,5	8E-07	10552,13	10216,22	10185,71261	C2
2	193270,7	1182,988	11578,06	11332,26	11326,22695	C2
3	204566,7	10163,34	2616,206	601,3435	384,6701439	C5
4	195386,5	1682,225	9308,027	9228,009	9307,506007	C2
5	149718,8	44689,05	54938,26	54879,93	54862,4687	C2
23	204590,8	10211,32	2058,257	76,67504	819,9167225	C4
24	194836,6	573,7221	10251,74	9828,427	9769,51586	C2
25	167226,3	27284,39	37699,87	37499,41	37438,81207	C2

Based on the rules in the K-Means Algorithm, iterations are performed until the clustering is the same as the previous iteration group by determining the new center point, describing the center point with the scenario of calculating the average of each attribute (Depth, Mag, Dmin, PGA). The description of the center point is outlined below:

Table 8.. New Center Point in the Second Iteration of K-Means

Cluster	**Depth**	**Mag**	**Dmin**	**PGA**
C1	201963,3	4,5625	807,5225	1,406724
C2	35941,73	4,772727	2062,33	26,77119
C3	35	4,5	5244	122,7774
C4	10	4,7	3099,667	272,2665
C5	23,785	5,2	2560	480,2309

Once determined, the clustering formation process is carried out in the following second iteration:

Table 9. Clustering Results on the Second Iteration with Dataset 25 data

No	**C1**	**C2**	**C3**	**C4**	**C5**	Nearest Distance
1	191.782.650.992	25.756.587.917	10.552.125.593	10.205.114.241	10.172.714.954	C5
2	190.646.530.223	24.623.560.563	11.578.057.618	11.325.944.576	1.131.125.252	C5
3	201.942.274.917	35.917.830.406	2.616.206.498	467.074.318	192.335.123	C5
4	1.927.646.618	26.773.555.307	9.308.026.758	9.235.846.234	9.282.438.126	C4
5	147.093.698.229	18.932.644.105	549.382.633	54.876.791.341	54.855.705.753	C2
23	201.967.476.954	35.949.997.097	2.058.256.504	9.649.668	674.972.015	C4
24	192.210.937.379	26.185.843.205	10.251.737.708	9.811.594.568	9.758.428.808	C5
25	164.596.253.435	2.503.454.033	37.699.874.604	37.487.820.606	37.435.426.031	C2

Clustering is an important technique in data analysis that aims to group data based on similarities. Two popular methods often used for clustering are Self-Organizing Map (SOM) and K-Means. Although these two methods have the same goal, which is to group data into different clusters, they use different approaches in achieving this goal. Self-Organizing Map (SOM) is a type of artificial neural network that uses unsupervised learning to generate a two-dimensional representation of high-dimensional data. SOM updates the weights of its neurons based on the given input data, using learning rates and neighbor functions to ensure that the network learns gradually and regularly. This method not only considers the data that is most similar to a particular neuron (Best Matching Unit or BMU), but also its neighboring neurons, thus creating a topological map that maintains the relationship between the data. K-Means is a simple and effective clustering method that works by grouping data into K clusters based on the Euclidean distance to the nearest centroid. In each iteration, the centroid is updated

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14120
e-ISSN : 2541-2019
p-ISSN : 2541-044X

by calculating the average of all the data in the cluster. Unlike SOM, K-Means does not use learning rate or neighbor function. The centroid update in K-Means is direct and fast, making this method very efficient for large datasets. This section will compare the clustering results obtained from the SOM and K-Means methods. This comparison will include an analysis of how the two methods cluster the data, the performance in producing meaningful clusters, and the advantages and disadvantages of each method in the context of data clustering. Through this comparison, it will help to better understand the characteristics and applications of both methods in various data analysis situations.

Visualization Comparison of SOM and K-Means Clustering Results

Comparison of clustering results of SOM and K-Means Algorithms Based on the dataset that has been provided as much as 1547 and the iteration process as much as 5. Testing is done using the Python Programmer with the results in the following figure

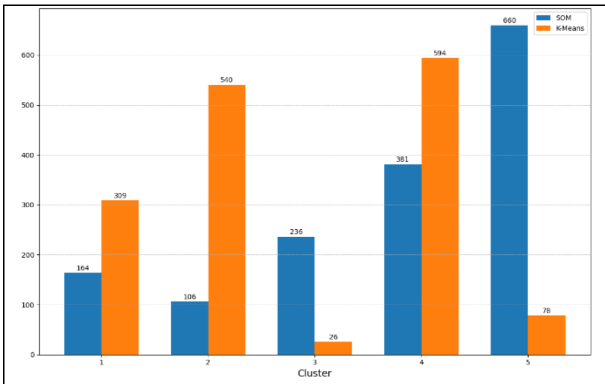

Fig. 3 Iteration of Cluster I Comparison

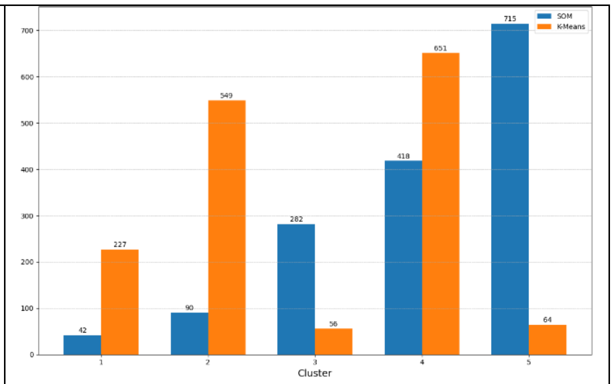

Fig. 4 Iteration of Cluster II Comparison

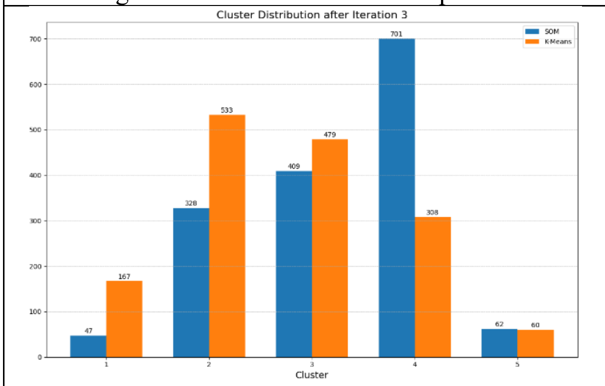

Fig. 5 Iteration of Cluster III Comparison

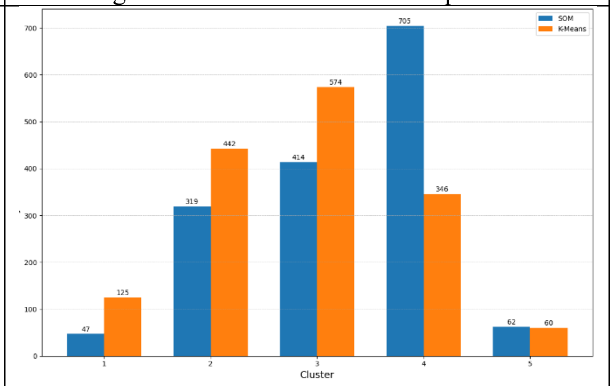

Fig. 6 Iteration of Cluster IV Comparison

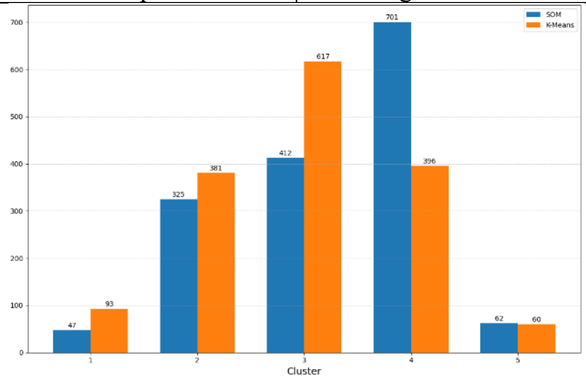

Fig. 7 Iteration of Cluster I Comparison

The visualizations in Figures 3 to 7 compare the data distribution across five clusters between two clustering algorithms, SOM (Self-Organizing Map) and K-Means. The results reveal distinct patterns for each algorithm. K-Means generally groups more data in the initial clusters (clusters 1, 2, and 3), showing a tendency to focus on these clusters. SOM, on the other hand, tends to distribute data more evenly, with significant dominance in cluster 4 and, in some cases, cluster 5. In clusters 1 and 2, K-Means consistently clusters more data than SOM. However, in clusters 3 and 4, SOM often has more data, indicating its preference for more extensive clusters. Cluster 5 shows similar data distribution between the two algorithms, reflecting a balanced clustering pattern. Overall, these

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14120
e-ISSN : 2541-2019
p-ISSN : 2541-044X

differences illustrate the distinct characteristics of the two algorithms: K-Means focuses on specific clusters, while SOM shows a more balanced or distributed approach across clusters, especially favoring larger ones.

Comparison of Data Group Suitability with SOM and K-Means Algorithms

To compare the suitability of data grouping on the SOM and K-Means algorithms which are defined as appropriate and inappropriate in this case data that has the same clustering place that has been provided with a description in the following table:

Table 10. Data Group Comparison of SOM and K-Means Data

Iterasi	Retrieved	Invalid
Iterasi 1	279	1268
Iterasi 2	162	1385
Iterasi 3	537	1010
Iterasi 4	481	1066
Iterasi 5	472	1075

In the first iteration, the algorithm showed that 279 data were placed according to the identified pattern, while 1268 data were not. This process continued in subsequent iterations with varying amounts of conforming and non-conforming data. In the second iteration, only 162 data were placed accordingly, while 1385 data were not. In the third iteration, there was an increase in the number of suitable data to 537, and the non-conforming data decreased to 1010. In the fourth iteration, the number of suitable data was 481, while 1066 data were not suitable. In the fifth iteration, the suitable data slightly decreased to 472, while the unsuitable data increased to 1075. This iterative process shows the dynamics of the algorithm in grouping data, with variations in the number of suitable and unsuitable data in each iteration. This reflects the algorithm's adjustment and refinement efforts to achieve more accurate and optimal data clustering.

CONCLUSION

The SOM and K-Means algorithms show different patterns of data distribution into five clusters. The significant difference in the amount of data in each cluster in the first iteration shows that the two algorithms have different approaches in clustering the data, which affects the final clustering result. The second iteration onwards showed significant changes in the distribution of data in the clusters, with both the SOM and K-Means algorithms showing a large redistribution of data in the second iteration, indicating adaptation and changes in clustering patterns. The third to fifth iterations show further adjustments, with the data distribution starting to stabilize in each cluster, especially with the SOM algorithm showing increasingly consistent data distribution patterns. The comparison between matched and unmatched data shows varying degrees of success in placing data in the right clusters. The first iteration shows 279 matched and 1268 unmatched data, while the fifth iteration shows an increase in suitability with 472 matched and 1075 unmatched data. This reflects the algorithm's dynamic and iterative efforts to achieve more accurate and optimal clustering, although there is still a significant amount of unmatched data at each iteration**.**

REFERENCES

AMANULLAH, F. F. (2022). *Analisis Potensi Likuifaksi Akibat Gempa Bumi Menggunakan Data Spt (Standar Penetration Test)*.

Artati, H. K., Pawirodikromo, W., & Purwanto, E. (2020). *15933-Article Text-37504-42123-10-20201024*. *XXV*(2).

Azmiyati, U., & Rancak, G. T. (2023). Kerentanan Masyarakat Lombok Terhadap Gempabumi; Belajar Dari Gempabumi Lombok 2018 Earthquake Vulnerability; Learning from Lombok. *JISIP (Jurnal Ilmu Sosial Dan Pendidikan)*, *7*(2).

DONOVAN, M. C. G., & MICKEY, D. (2018). ANALISIS PENDEKATAN EMPIRIS PGA (PEAK GROUND ACCELERATION) PULAU BALI MENGGUNAKAN METODE. *Jurnal Ilmu Dan Inovasi Fisika*, *2*(02), 87–93.

HASUGIAN, P. M., HUTAHAEAN, H. D., SINAGA, B., SRIADHI, & SILABAN, S. (2021). Big data technology for village status classification based on village index building involving k-means algorithm in programs to support the work of the ministry of the village. *Journal of Theoretical and Applied Information Technology*, *99*(19), 4658–4673.

Hermon, D. (2015). *Geografi bencana alam*. PT. RajaGrafindo Persada-Rajawali Pers.

Hidayat, R., & Munir, D. (2018). *Longsor Bendawuluh-Banjarnegara dan Penanganannya*.

Imani, N., Alfassa, A. I., & Yolanda, A. M. (2023). ANALISIS CLUSTER TERHADAP INDIKATOR DATA SOSIAL DI PROVINSI NUSA TENGGARA TIMUR MENGGUNAKAN METODE SELF ORGANIZING MAP (SOM). *Jurnal Gaussian*, *11*(3), 458–467. https://doi.org/10.14710/j.gauss.11.3.458-

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

467

Kapojos, C. G., Tamuntuan, G., & Pasau, G. (2015). Analisis Percepatan Tanah Maksimum Dengan Menggunakan Rumusan Esteva Dan Donovan (Studi Kasus Pada Semenanjung Utara Pulau Sulawesi). *Jurnal Ilmiah Sains*, 99–104.

Kusumaningrum, E. (2017). *Evaluasi Kriteria Kerusakan Bangunan Rumah Tinggal Sederhana Akibat Gempa Bumi*.

Mahendra Taruna, R., Azhar, T., & Setiadi, P. (2020). *PENENTUAN RUMUS PERCEPATAN TANAH AKIBAT GEMPA BUMI DI KOTA MATARAM MENGGUNAKAN METODE EUCLIDEAN DISTANCE*.

Mulyani, H., Setiawan, R. A., Musawarman, & Romadloni, A. (2022). Clustering Area Covid-19 Indonesia With K-Means (Case study : Kaggle Dataset). *Journal of Information Technology and Its Utilization*, *5*(2), 41–46. https://doi.org/10.56873/jitu.5.2.4894

Munawar, G., & Kunci, K. (2015). *Implementasi Algoritma Self Organizing Map (SOM) untuk Clustering Mahasiswa pada Matakuliah Proyek (Studi Kasus : JTK POLBAN)*.

Paramesti, C. A. (2011). Kesiapsiagaan masyarakat kawasan Teluk Pelabuhan Ratu terhadap bencana gempa bumi dan tsunami. *Jurnal Perencanaan Wilayah Dan Kota*, *22*(2), 113–128.

Pasau, G., Bobanto, M. D., & Pandara, D. P. (2018). Model Percepatan Tanah Maksimum Di Kota Manado Menggunakan Metode Donovan dan McGuire. *Jurnal MIPA*, *7*(1), 52–55.

Purwanto, D., Damas Permadi, D., Sukmayu Saputri, U., Umar Aditiawarman, Dini Oktarina Dwi Handayani, & Deshinta Arovva Dewi. (2023). Kajian Tentang Bencana Dan Dampak Gempa Bumi di Cianjur Bersama Mahasiswa Teknik Sipil Nusaputra. *Jurnal Abdi Nusa*, *3*(1), 36–39. https://doi.org/10.52005/abdinusa.v3i1.130

Romadiana, D., Syafriani, & Sabarani, A. Z. (2018). Dwi Romadiana1. *Pillar of Physics*, *11*(1), 9–16.

Saputra, E., Makrup, L., & Nugraheni, F. (2020). *ANALISIS PERCEPATAN TANAH PERMUKAAN DI WILAYAH RIAU DENGAN METODE PSHA*.

Suhada, A. S., Kardoso, R., Zuhdi, M., Syahrial, A., & Syamsuddin, S. (2023). Penghitungan Peak Ground Acceleration (PGA) di Desa Kuta Menggunakan Persamaan Empiris pada Gempa Bumi 19 Agustus 1977 (Mw= 8.3). *Jurnal Pendidikan, Sains, Geologi, Dan Geofisika (GeoScienceEd Journal)*, *4*(1), 9–13.

Sungkowo, A. (2018). *Perhitungan Nilai Percepatan Tanah Maksimum Berdasar Rekaman Sinyal Accelerograph di Stasiun Pengukuran UNSO Surakarta Ari*.

Tjandra, K. (2018). *Empat bencana geologi yang Paling Mematikan*. UGM PRESS.

WINDA, E. (2023). *STUDI LITERATUR KETERKAITAN GEMPA BUMI BERDASAR KAN ILMU FISIKA DAN PERSPEKTIF AL-QUR'AN*. UIN RADEN INTAN LAMPUNG.

*name of corresponding author