KPJ 8(3) (2024)

KAPPA JOURNAL
Physics & Physics Education

https://e-journal.hamzanwadi.ac.id/index.php/kpj/index

A Study of Grouping of Earthquake Damage from Magnitude Scale in Lombok Using K-Means Modeling

Kertanah[1*], Alissa Chintyana[2], Chandrawati[3], Basirun[4], Mutia Rosiana Nita Putri[5], Nasibatul Mahmudah[6]

[1,2,3,4] Department of Statistics, Universitas Hamzanwadi, Lombok Timur, NTB, Indonesia
[5] Tadris IPA, Institut Studi Islam Sunan Doe, Lombok Timur, NTB, Indonesia
[6] Dinas Pertanian, Lombok Timur, NTB, Indonesia

Received: 24 September 2024
Revised: 17 December 2024
Accepted: 20 December 2024

Corresponding Author:
Kertanah
kertha90@gmail.com

DOI:
https://doi.org/10.29408/kpj.v8i3.27563

Abstract: This study aims to group earthquake damage from its magnitude scale and visualize it on a geographical map. The magnitude of the earthquake was grouped using the K-means model. It is one of the most popular and effective clustering models in grouping data, such as earthquake data. The dataset used in this study is earthquake data for the last ten years on Lombok Island. The optimal number of clusters was used which is 2 in this case, based on the highest Silhouette score of 0.930. The highest Silhouette score shows the optimal number of clusters. The cluster on the geographical map shows most earthquakes' distribution in Northen Lombok Island with cluster 1 consisting of 145 earthquakes, while cluster 2 consists of 3 earthquakes. In addition, the earthquake's damage based on its magnitude scale, there were four different kinds of earthquake damage: slight, limited, minor, and severe damage that have occurred for the last ten years in Lombok Island. Minor and Slight damages were dominant, respectively. However, severe damage occurred in the northern part of Lombok Island due to an earthquake in 2018.

Keywords: K-means model; earthquake; grouping, Lombok

Introduction

Earthquake is one of the threats to human life around the world. Indonesia is a country that has a level of vulnerability to earthquakes because Indonesia is a maritime country located on three plates of the world commonly called the Pacific Ring of Fire (Setiawan, I. N., Krismawati, D., Pramana, S., Tanur, E., 2022; Utomo & Purba, 2019; Kirana et al., 2019). West Nusa Tenggara (NTB) is one of the provinces in Indonesia that is often affected by earthquakes. Almost every day there are earthquakes, but August 2018 was the last major earthquake in Lombok, NTB. The earthquake itself occurs due to plate motion that causes sudden vibrations on the earth's surface (Rifa & Pratiwi, Hasih, 2019). The occurrence of earthquakes causes a lot of damage such as infrastructure, buildings, highways And can also cause death. Therefore, a disaster mitigation effort needs to be carried out by all parties, from the government, academics, and researchers according to their respective expertise. About these mitigation efforts, the clustering technique in data mining is very useful for grouping earthquake data based on the same characteristics so that it can be used as a basis for predicting future earthquake events (Murdiaty, Angela, Sylvia, 2020).

The activity of seismic refers to the frequency, type, and size of earthquakes experienced over some time. The size of earthquakes can be determined by using the Richter scale. Richter scale determines earthquakes' magnitude based on amplitude. The following is a class

How to Cite:
Kertanah, K., Chintyana, A., Chandrawati, C., Basirun, B., Putri, M. R. N., & Mahmudah, N. (2024). A Study of Grouping of Earthquake Damage from Magnitude Scale in Lombok Using K-Means Modeling. *Kappa Journal, 8*(3), 399-403. https://doi.org/10.29408/kpj.v8i3.27563

of damage based on the magnitude scale of an earthquake (Irawan et al., 2020).

Table 1: The classification of earthquake's magnitude scale

No	Magnitude	Class of damage
1	3.0 – 3.9	Limited
2	4.0 – 4.9	Minor
3	5.0 – 5.9	Slight
4	6.0 – 6.9	Severe
5	7.0 – 7.9	Serious
6	8.0 or more	Great

Data mining is a popular and widely used approach in accessing data today. Data mining is the process of searching for important forms or information in data using certain methods or techniques (Mardi, 2016). The ease of filtering data and information in the form of data grouping can use one of the data mining approaches, namely clustering. Clustering is the process of grouping objects based on similarity into different groups (Madhulatha, 2012). There are many types of clustering algorithms, but the K-Means clustering method is a popular algorithm and is known as a more efficient clustering method, especially on large enough data, such as earthquake data (Raghuwanshi & Arya, 2012; Kertanah, Rahadi, et al., 2022; Wijaya et al., 2024). In addition, many studies applying the K-means model were conducted by researchers either in earthquakes or different study cases. K-means was applied to clustering earthquake epicenter (Novianti et al., 2017). Dwitiyanti et al.,(2023) implemented K-means for clustering earthquake-prone areas in Indonesia. This study used the Silhouette index as an indicator to determine an optimal number of clusters. Applying the K-means model was studied by Kertanah et al., (2022) for clustering analysis in West Nusa Tenggara. This study determined an optimal number of clusters by utilizing the Silhouette index. The k-means model was implemented in clustering for the province of the risk of COVID-19 in Indonesia. This study applied three different methods to determine an optimal number of clusters that are Elbow method, the Silhouette index, and Gap statistics, respectively. Another study was also conducted by applying the K-means model to cluster COVID-19 cases and deaths in Southeast Asia (Abdullah et al., 2022).

As stated by the backdrop above, this study proposes a K-means model to group the level of earthquake damage based on its magnitude scale. Therefore, this study also visualizes and interprets the estimation of earthquake damage on a geographical map to show the distribution of earthquake damage on Lombok Island.

Method

This study used earthquake data for the last ten years from September 2014 to September 2024 in Lombok Island. This data is obtained from the United States Geology Survey (USGS) global public seismic catalogs (https://earthquake.usgs.gov/earthquakes/search/). This study implements the K-means model to group earthquake damage. Hence, the number of optimal clusters is determined by using the Silhouette index. The following is the flowchart of the data processing and interpretation stages in this study.

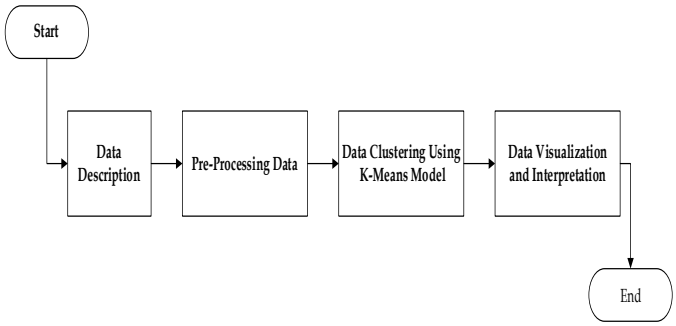

Figure 1. Data processing and interpretation stages

This study utilized Python as a tool to do those stages. In the first stage, earthquake data was collected with a minimum magnitude of 3.0. There are 22 variables in the earthquake data. In addition, the pre-processing stage was implemented to clean the data from missing values, and also selected variables that would be used - time, latitude, longitude, magnitude, depth, and place. Clustering data was carried out by using the K-means model. The result of clusters was evaluated by using the Silhouette index to find out the optimal number of clusters.

Result and Discussion

Data description was carried out to show the frequency of earthquakes. Figure 2 presents the frequency of the earthquakes that occurred in Lombok.

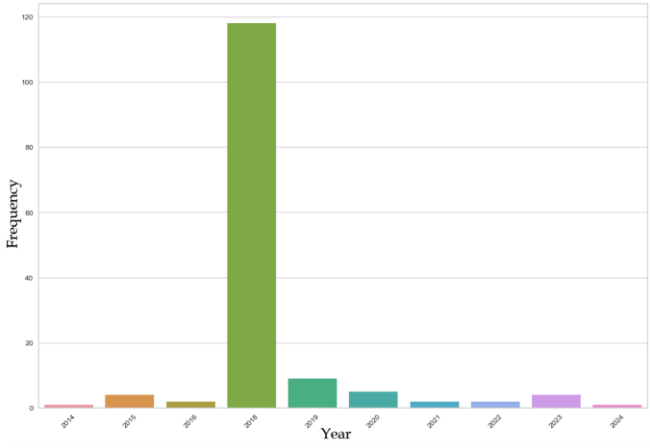

Figure 2. Lombok's earthquakes for the last ten years

For the last ten years, from 2014 to 2024, most earthquakes occurred in 2018. Moreover, in 2019, there were still quite a lot of aftershocks.
Earthquake's dataset was randomly generated by the K-means model. The Silhouette index was implemented to find out the optimal number of clusters. The following is a graph to find out the number of clusters (k).

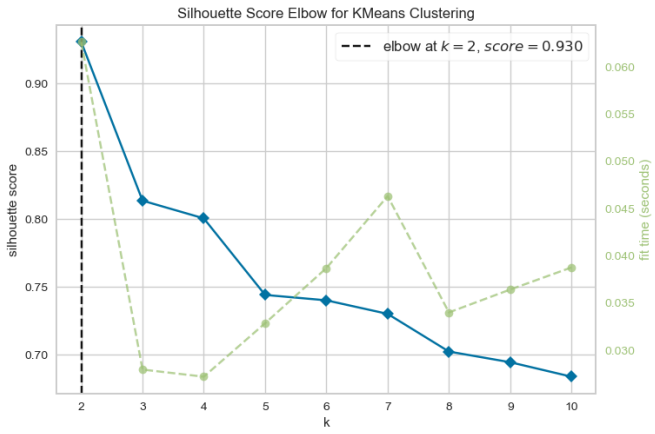

Figure 3. The optimal number of clusters (k)

The Figure above presents the optimal number of clusters. It suggests 2 clusters.

Table 2: The number of clusters against the Silhouette index

No	Number of clusters (k)	Silhouette score
1	2	0.930
2	3	0.813
3	4	0.800
4	5	0.744
5	6	0.740
6	7	0.730
7	8	0.702
8	9	0.694

Table 2 depicts the Silhouette score with the different number of clusters. Cluster number 2 shows the highest Silhouette score. It presents the optimal number of clusters. Therefore, two clusters will be used in this clustering of earthquake damage in the next stage. Plotting various earthquakes' magnitudes on the geographical map provides more context. It shows how different magnitude of earthquakes is distributed across Lombok Island based on their source.

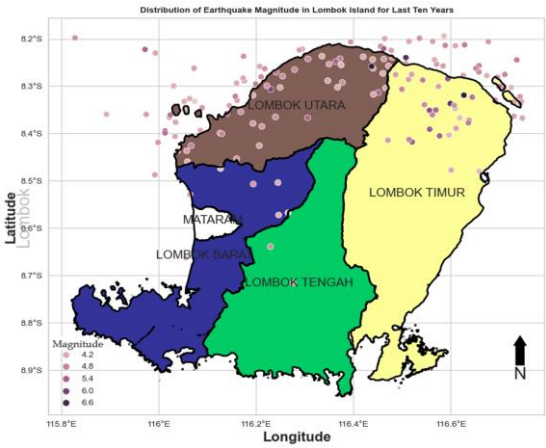

Figure 4. Distribution of earthquake's magnitude 2014 to 2024

Figure 4 presents the distribution of earthquake magnitude that has occurred in Lombok for the last ten years. Most earthquakes have occurred in the northern Lombok. The following Figure 5 shows the clustered earthquake's magnitude for the last ten years in Lombok Island.

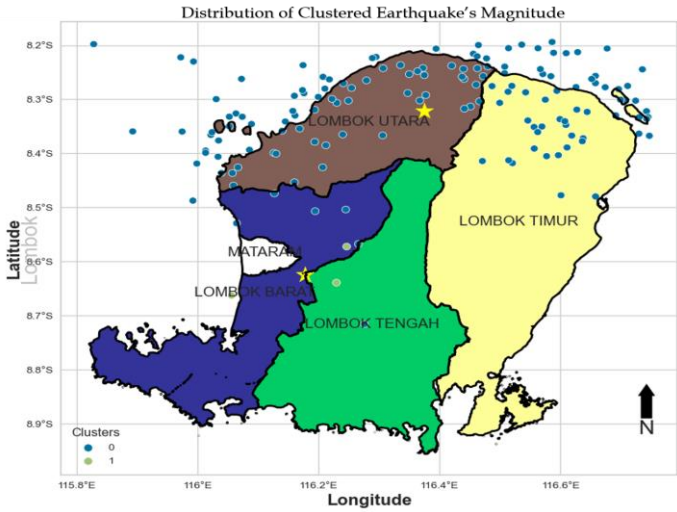

Figure 5. Distribution of clustered earthquake's magnitude

Out of 148 earthquakes in the dataset which have four different classes of damage as shown in Figure 6. Cluster 1 consists of 145 earthquakes (blue color), while the remaining 3 earthquakes (green color) are in Cluster 2 as shown in Table 3. The Euclidian distance between the centroids (yellow star) and data was used to generate the clusters. The geographical map shows the earthquake's damage was pretty spread in northern Lombok Island in Figure 6.

Table 3: The number of clusters in earthquake damage.

No	Cluster	Member of clusters
1	0	145
2	1	3

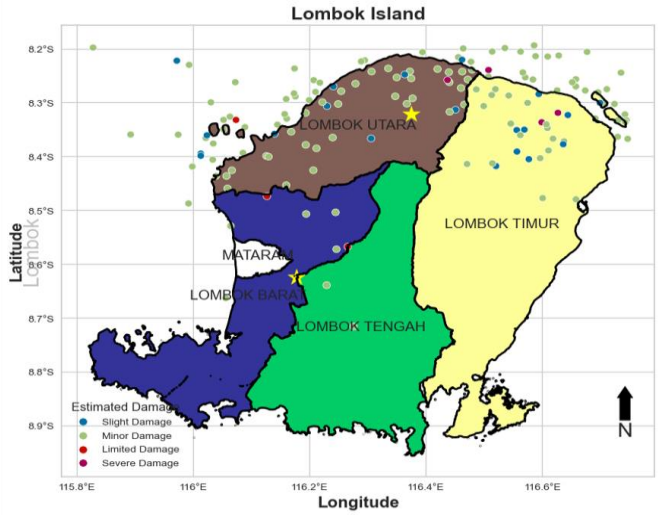

Figure 6. Distribution of earthquake damage for the last ten years

Figure 6 shows mostly earthquakes based on magnitude scale that occurred in the last ten years. Both Minor damage and slight damage are dominant, respectively. However, most of the earthquake's severe damage occurred in both the north Lombok Regency and the northern area of the east Lombok Regency. Most of them occurred in 2018.

Conclusion

Based on clustering using the K-means model, two clusters were carried out. The optimal number of clusters is based on the highest Silhouette score. The higher the Silhouette score the more optimal number of clusters. The result of clustering shows most earthquakes were spread out in the northern part of Lombok Island. Furthermore, four different damages occurred in the last ten years in Lombok Island from 2014 to 2024. Minor damage was dominant. However, Severe damage occurred the most in the northern area of Lombok Island in 2018.

Acknowledgments

We greatly acknowledge P3MP for the financial support of the grant for publication funding.

References

Abdullah, D., Susilo, S., Ahmar, A. S., & Rahmat, R. R. (2022). The application of K - means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data. *Quality & Quantity*, *56*(3), 1283–1291. https://doi.org/10.1007/s11135-021-01176-w

Dwitiyanti, N., Kumala, S. A., & Handayani, S. D. (2023). *Penerapan metode k-means pada klasterisasi wilayah rawan gempa di indonesia*. *6*, 1029–1037.

Irawan, L., Hasibuan, L. H., & Fauzi. (2020). *Analisa Prediksi Efek Kerusakan Gempa Dari Magnitudo (Skala Richter) Dengan Metode Algoritma Id3 Menggunakan Aplikasi Data*. *14*(2), 189–201. https://doi.org/https://doi.org/10.47111/JTI

Kertanah, Rahadi, I., Novianti, B. A., Syahidi, K., Putra, M., Gazali, M., & Hirzi, R. H. (2022). Indonesian Physical Review. *Indonesian Physical Review*, *5*(3), 197–207. https://doi.org/https://doi.org/10.29303/ipr.v5i3.148

Kirana, M. C., Perkasa, N. P., Zainuddin, M., Maidel, L., Informatika, T., & Batam, P. N. (2019). Visualisasi Kualitas Penyebaran Informasi Gempa Bumi di Indonesia Menggunakan Twitter. *Journal of Applied Informatics and Computing (JAIC)*, *3*(1), 23–32.

Madhulatha, T. S. (2012). AN OVERVIEW ON CLUSTERING METHODS. *IOSR Journal of Engineering*, *2*(4), 719–725. https://doi.org/https://doi.org/10.48550/arXiv.1205.1117

Mardi, Y. (2016). Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5. *Jurnal Edik Informatika*, *2*, 213–219. https://doi.org/https://doi.org/10.22202/ei.2016.v2i2.1465

Murdiaty, Angela, Sylvia, C. (2020). Pengelompokkan Data Bencana Alam Berdasarkan Wilayah , Waktu , Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma K-Means. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, *4*, 744–752. https://doi.org/10.30865/mib.v4i3.2213

Novianti, P., Setyorini, D., & Rafflesia, U. (2017). *K-Means cluster analysis in earthquake epicenter clustering*. *3*(2), 81–89.

Raghuwanshi, S. S., & Arya, P. (2012). Comparison of K-means and Modified K-mean algorithms for Large Data-set Abstract : *International Journal of Computing, Communications and Networking*, *1*(3), 106–110.

Rifa, I. H., & Pratiwi, Hasih, R. (2019). Implementasi algoritma clara untuk data gempa bumi di indonesia 1 1,2,3. *SEMINAR NASIONAL PENELITIAN PENDIDIKAN MATEMATIKA (SNP2M)*, *2006*, 161–166.

Setiawan, I. N., Krismawati, D., Pramana, S., Tanur, E. (2022). Klasterisasi Wilayah Rentan Bencana Alam Berupa Gerakan Tanah. *Seminar Nasional Official Statistics*, 669–676.

Utomo, D. P., & Purba, B. (2019). Penerapan Datamining

pada Data Gempa Bumi Terhadap Potensi Tsunami di Indonesia. *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, *September*, 846–853.

USGS. (2024). *Search Earthquake Catalog*. https://earthquake.usgs.gov/earthquakes/search/. [Accessed September 10, 2024]

Wijaya, T. A., Utami, E., & Fatta, H. Al. (2024). Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Data Gangguan PT . PLN UID Kalselteng. *INNOVATIVE: Journal Of Social Science Research*, *4*, 8846–8854.