Application of Bisecting K-Means Method in Grouping Earthquake Data (Case Study: Earthquakes in Indonesia 2023)

Zulkifli Rais[1*],Hardianti Hafid[1], and Shopia Risqi[1]
[3]Statistics Study Program: Faculty of Mathematics and Natural Sciences, Makassar State University, Makassar City, South Sulawesi, Indonesia
*Corresponding author: zulkifli.rais89@unm.ac.id

Received: 10 October 2024
Revised: 16 September2025
Accepted: 30 October 2025

ABSTRACT – Earthquakes are natural disasters that frequently occur in Indonesia, threatening the safety and resilience of its communities. This study aims to analyze the descriptive and clustering results of earthquake data in Indonesia. The data used in this study include various variables such as latitude, longitude, magnitude, and depth as the main features. The method used in this study is Bisecting K-means, and the Davies Bouldin Index test is used to determine the number of clusters. The study results indicate the formation of 3 groups, where cluster 1 falls into the deep earthquake category, cluster 3 falls into the intermediate earthquake category, and cluster 2 falls into the shallow earthquake category, with an average Davies-Bouldin Index value of 0.4758.
Keywords – **Clustering, Bisecting K-means, Earthquakes**

I. INTRODUCTION

Data Mining is a computational process that uses artificial intelligence techniques, machine learning, statistics, and database systems to create interesting models in large data sets [9]. Data mining is closely related to data analysis, which is used to find patterns and similarities in a set of data with certain techniques or methods [7]. There are several methods used to perform data mining, namely Association, Classification, Regression, and Clustering. However, this research uses the clustering method.

Clustering is one of the data mining techniques that aims to identify a group of objects that have certain similar characteristics that can be separated from other groups, so that objects that are in the same group are relatively more homogeneous than different objects in different groups, the number of groups that can be identified depends on the number and variety of data objects [16]. In data mining, there are two types of clustering methods used in grouping, namely hierarchical clusters and non-hierarchical clusters [18]. Hierarchical cluster methods group data in layers based on similarity, forming a tree-like structure. Non-hierarchical cluster methods determine the number of groups first, then place the data into them without a layered structure [10]. One of the algorithms contained in the non-hierarchical cluster method is the K-means algorithm [3].

The K-means algorithm is an iterative clustering that partitions a dataset into a predefined number of K clusters in the initial data [14]. The K-means algorithm quickly clusters large data with a centroid-based partitioning method. The bisecting K-means algorithm is a variation that forms a hierarchy [20]. In this study using bisecting K-means algorithm method in clustering earthquake data in Indonesia.

Earthquakes are one of the most common natural disasters in many parts of the world. Earthquakes can cause great physical damage to infrastructure and buildings, and threaten the safety of human life [11]. BMKG records daily earthquake activity in Indonesia, but different intensities and epicenter points make not all earthquakes felt [4]. The largest historic earthquakes have been greater than 9 magnitude, although there is no magnitude limit, namely in Japan in 2011 and there was a 9.1 magnitude earthquake in Aceh that caused a tsunami [1].

Related research conducted by Tania et al. (2023) analyzed earthquakes in Indonesia using a comparison method between the K-means algorithm and the bisecting K-means algorithm which resulted in cluster 1 being the cluster that had the highest number of earthquake events and had the most significant results for all magnitudes. 82.49% of the earthquakes were less than 5 magnitude, 74.84% were between 5-7 magnitude, and 71.42% were more than 7 magnitude. In addition, testing using the Silhoutte and Davies Bouldin methods, it was found that the bisecting K-means algorithm produced k=3 for the optimal number of clusters with average values of 0.7340 and 0.4081. Based on research conducted by Dwididanti & Anggoro (2022) on the comparative analysis of the bisecting K-means algorithm and fuzzy c-means on credit card user data obtained the results that the bisecting K-means value without normalization has a higher silhouette coefficient value compared to fuzzy c-means which is 0.588> 0.488.

Research conducted by Prasetio et al, (2023) on the analysis of earthquakes in Indonesia with the clustering method, it was concluded that the grouping of earthquake-prone areas in Indonesia was based on 3 clusters. Testing using the earthquake clustering method with the K-means algorithm can produce clusters that have cluster group members in accordance with manual calculations such as Cluster_0 on Rapid Miner has 209 cluster members representing the Low cluster, Cluster_1 has 863 cluster group members as a representation of the Medium cluster, and Cluster_2 has 41 cluster members corresponding to the High cluster representation. Based on the description related to earthquakes and previous research, the authors are interested in conducting research using the Bisecting K-means method and taking the title "Application of the Bisecting K-means Method in clustering earthquake data in Indonesia".

II. LITERATURE REVIEW

A. Data Mining

Data mining is a process used to extract useful information and patterns from large datasets [22]. It is defined as a technique used to find information hidden within a database. The practice of data mining involves various techniques, including artificial intelligence, machine learning, and statistical methods [23]. Muflikhah (2018) describes data mining as the process of breaking down complex datasets to uncover implicit and previously unknown information. This process is part of the broader field of Knowledge Discovery in Databases (KDD), which includes steps such as data cleaning, data integration, data selection, data transformation, data mining, pattern evaluation, and knowledge presentation [5].

B. Clustering

Clustering is a data mining technique that involves grouping a set of data points into clusters based on predefined criteria. It organizes a large volume of data within a database into groups with similar characteristics. Clustering assigns data points into groups based on their similarity without prior knowledge of class labels. According to Apriyani et al. (2023), clustering groups objects that share similarities, and the more similar the objects, the more accurate the clustering result. Santosa (2007) emphasizes that clustering aims to organize data into groups, ensuring that each group contains data that are as similar as possible.

C. K-Means Algorithm

K-Means is a widely used clustering method, especially when dealing with large datasets, due to its relatively fast performance compared to hierarchical methods. As a non-hierarchical clustering method, K-Means aims to divide data into one or more groups, or clusters, so that similar data points are grouped together, while dissimilar data are placed in separate clusters [22]. The K-Means algorithm operates by defining the number of clusters (k) and iteratively assigning data points to the closest cluster based on the Euclidean distance from the cluster centroids. It is particularly useful for numeric attributes [8]. James et al. (2017) outline the steps in the K-Means algorithm, which include initializing the centroids, assigning data points to the nearest centroid, recalculating the centroid positions, and repeating the process until convergence.

D. Bisecting K-Means

Bisecting K-Means is a variant of the K-Means algorithm that integrates hierarchical clustering concepts. It iteratively splits data into two sub-clusters using the K-Means algorithm, repeating this process until the desired number of clusters is achieved [24]. The advantage of the Bisecting K-Means algorithm is its efficiency in overcoming situations where traditional K-Means may get stuck in local optima[6]. It is particularly effective when applied to large datasets, as it reduces computational complexity by focusing on splitting individual clusters rather than the entire dataset in each iteration [20]. Below are the steps of the Bisecting K-means algorithm. First, treat the entire data as one cluster and set the value of k. Then, perform the bisecting step by splitting the data into two sub-clusters using the K-means algorithm. The steps for the K-means algorithm are as follows [6]:

a. Initialize 2 centroids for the clusters.
b. For each data point in the cluster, calculate its similarity to both centroids and assign the point to the nearest centroid using the Euclidean Distance formula (Equation 1).
c. Recalculate the centroids based on the new data assignments.
d. Repeat steps 2 and 3 until convergence. Then repeat the bisecting process until the number of clusters equals the predefined value of k, and choose the clustering result with the highest similarity.

$$D(i,j) = \sqrt{(x_{1i} - x_{1j})^2 + \ldots + (x_{ki} - x_{kj})^2} \quad (1)$$

with :

$D(i,j)$: Distance from data point i to cluster center j

x_{ki} : The i-th data point for attribute k

x_{kj} : The j-th cluster center for attribute k

E. Davies Bouldin Index

The Davies Bouldin Index (DBI) is a method used to evaluate clustering validity and to identify the optimal number of clusters in a clustering process. The DBI calculation steps are as follows [21]:

$$SSW_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_j, c_i) \quad (2)$$

with :

m_i : Number of data in cluster i

x : Data point in cluster
c_i : *Centroid* of cluster i
$d(x_j, c_i)$: Euclidean distance between each data point and centroid
Sum of square between cluster (SSB) is a formula used to measure the separation between clusters. The formula for calculating SSB is a follows:

$$SSB_{ij} = d(c_i, c_j) \quad (3)$$

with:
$d(c_i, c_j)$: represents the Euclidian distance between the centroids of clusters i and j
Once the values of separation and cohesion are obtained, the ratio measurement (R_{ij}) is then calculated to determine the comparison between cluster i and cluster j. The formula used for this calculation is:

$$R_{i,j} = \frac{SSW_i + SSW_i}{SSB_{i,}j} \quad (4)$$

The formula for calculating the Davies Bouldin Index (DBI) is as follows:

$$DBI = \frac{1}{K}\sum_{i=1}^{k} max_i \neq (R_{i,j}) \quad (5)$$

The lower the Davies Bouldin Index (DBI) value obtained, the more optimal the clustering result.

F. Earthquake

Earthquakes are geological events that result from the release of energy stored in the Earth's crust, often due to tectonic plate movements. Yuwanto (2018) defines earthquakes as natural phenomena involving ground shaking that can cause both physical and non-physical damage. Tania et al. (2023) note that Indonesia, situated on the Pacific Ring of Fire, experiences frequent earthquakes caused by tectonic activity, including spreading, collision, and transform movements between tectonic plates. These movements may occur slowly and imperceptibly, but can be detected by seismographs, which record the magnitude of the earthquakes. According to Prataopu (2013), earthquakes can be explained by two primary theories: fault movement and elastic rebound. Fault movement theory posits that the shifting of Earth's tectonic plates, which originally formed the supercontinent Pangaea, continues to cause earthquakes today. Meanwhile, elastic rebound theory explains that earthquakes occur due to the sudden release of energy from the deformation of the Earth's crust.

Various types of earthquakes exist, including tectonic, volcanic, collapse, and meteor-impact earthquakes. Tectonic earthquakes are the most common and are caused by the movement of tectonic plates, while volcanic earthquakes occur due to volcanic activity [13]. Understanding the causes and characteristics of these different types of earthquakes is crucial for analyzing seismic activity in regions such as Indonesia, where tectonic earthquakes are frequent.

III. METHODOLOGY

A. *Data Sources and Research Variables*

This research is a quantitative approach, which is research that uses analysis that emphasizes numerical data. This research categorizes earthquake data in Indonesia using the bisecting K-means method. The data in this study are secondary data, namely data on earthquakes obtained from the Meteorology, Climatology and Geophysics Agency (BMKG) from January to December 2023, which can be accessed via https://repogempa.bmkg.go.id/.

The variables in this study are magnitude, geographical coordinates (latitude and longitude), earthquake depth (Depth), earthquake time (Time). The procedures carried out in this study are retrieving data on the website of the Meteorology, Climatology and Geophysics Agency (BMKG), preprocessing data, conducting descriptive analysis on the data to be used, clustering using the bisecting K-means method to get the desired number of clusters (k), evaluating clustering results, visualizing clustering results using scatter plots and cluster maps, interpreting clustering results, making conclusions and compiling research reports. Data analysis in this study was processed with statistical software using the bisecting K-means method.

B. *Research Steps*

The analysis steps in this study can be described in detail as follows:

a. Collect data on earthquake events in Indonesia.
b. Perform data cleaning.
c. Input data into R studio.
d. Select a cluster to be divided.
e. Perform the bisecting process as follows:
    1) Find 2 sub clusters using the K-means algorithm.
    2) Cluster the data for each object in the cluster. Then, calculate the distance between the two centroids using equation 1.
    3) Recalculate the position of the centroids of each new cluster as the average of all objects in the cluster.
    4) Repeat steps 1, 2, and 3 until the centroid positions stabilize or small changes occur (convergence).
f. Choose the cluster to be split with the greatest variability or the largest number of object to be selected.
g. Repeat step e until the desired number of cluster is achieved.
h. Evaluate the clustering results using the Davies Bouldin Index to assess the quality of the generated clustering.

i. Interpret the clustering results and visualize the using scatter plots and cluster maps.

IV. RESULTS AND DISCUSSIONS

A. Descriptive Analysis

An overview of the variables studied can be seen through descriptive analysis using the bisecting K-means method. The grouping of earthquake events based on magnitude can be seen in Figure 1 to Figure 3.

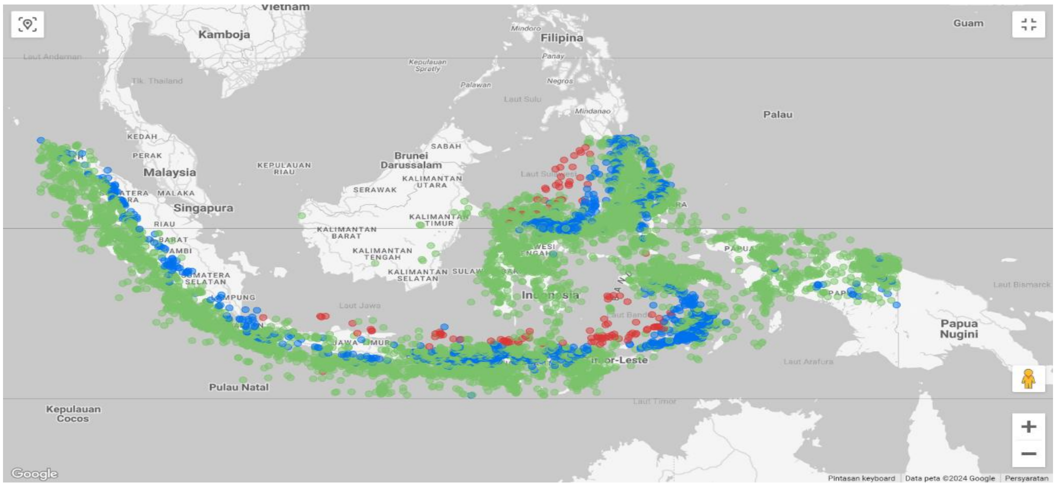

Figure 1 Map of earthquake groupings based on *magnitude* <5 SR

Based on Figure 2 is a visualization in the form of a map of the actual data of earthquake events in Indonesia which has a magnitude of less than 5 SR which is very much an earthquake event because in Indonesia every day there are micro earthquakes. Sulawesi Island is the island with the most earthquake events at magnitudes less than 5 SR, one of which is in the North Sulawesi region and its surroundings there are 899 earthquake events. As for Kalimantan Island, which only has a few earthquake events because Kalimantan Island is located far from the plate collision zone or does not have active volcanoes.

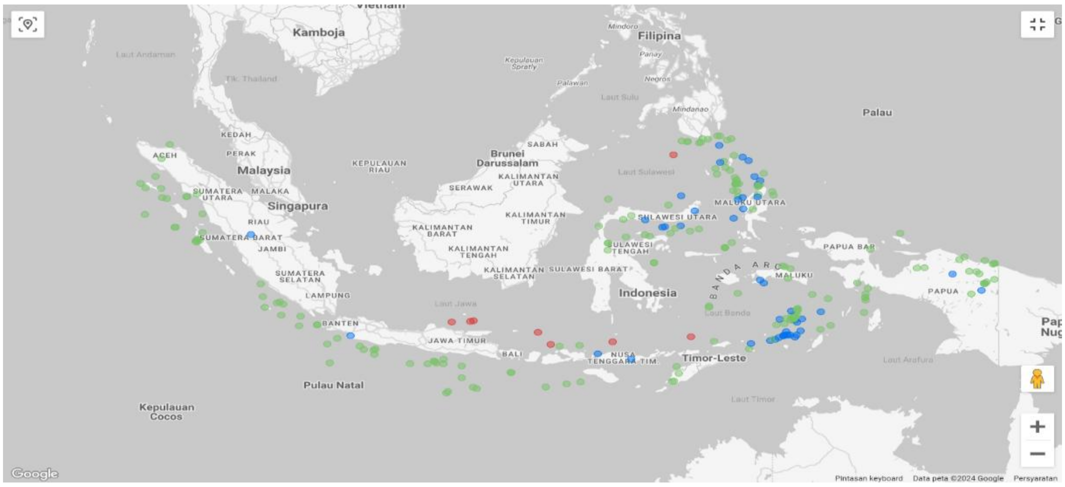

Figure 2 Map of earthquake groupings based on *magnitude* 5-7 SR

Based on Figure 3 is a visualization in the form of a map of the actual data of earthquake events in Indonesia that have a magnitude of 5 - 7 SR where every region in Indonesia has experienced an earthquake with a magnitude of 5 - 7 SR.

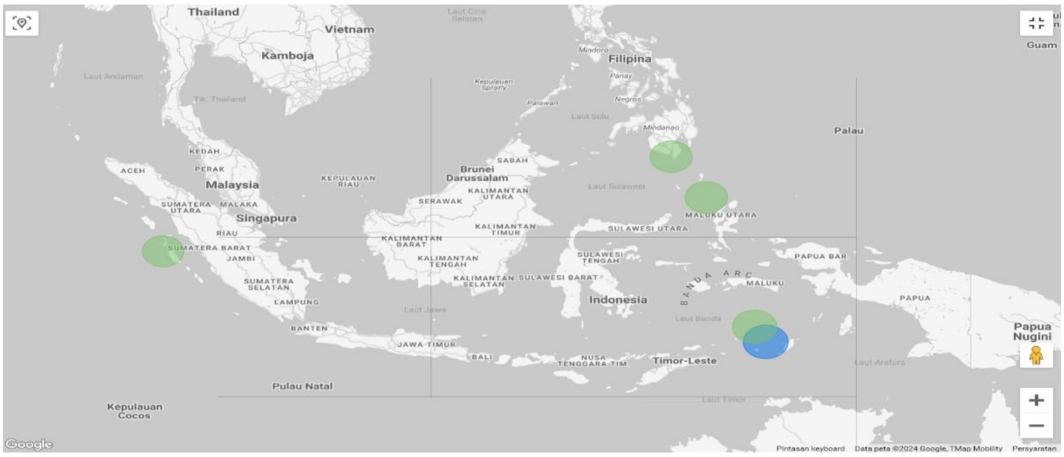

Figure 3 Map of earthquake groupings based on *magnitude* >7 SR

Based on Figure 4 is a visualization in the form of a map of the actual data of earthquake events in Indonesia that have a magnitude of more than 7 SR. It can be seen in the figure that the event only occurred several times in Indonesia during 2023, namely the Eastern Indonesia region.

Descriptive statistics of earthquake events in Indonesia can be seen in tables 1 and 2.

a. Magnitude

Table 1 Descriptive Analysis of Magnitude

No.	Description	Value (SR)
1	Min	0,8139
2	1st Qu.	2,8295
3	Median	3,3224
4	Mean	3,3833
5	3rd Qu.	3,8598
6	Max	7,7357

Based on Table 1 above, it can be seen that the minimum magnitude is 0.8139 SR and the maximum is 7.7357 SR. The median magnitude of the earthquakes had a magnitude of 3.3224 SR, meaning that half of the earthquake events had magnitudes less than or equal to 3.3224 SR and the other half were more than 3.3224 SR. The mean magnitude of the earthquakes had a magnitude of 3.3833 SR. Based on these values, the mean and median values are almost the same, indicating a fairly symmetrical distribution of magnitude. The first quartile is 2.8295 magnitude which indicates that a quarter of the earthquake events were quite low and the third quartile is 3.8598 magnitude which indicates that most of the earthquake events are still classified as mild to moderate earthquakes.

b. Depth of Earthquake

Table 2 Descriptive Analysis of Earthquake Depth

No.	Description	Value (km)
1	Min	1
2	1st Qu.	10
3	Median	17
4	Mean	48,41
5	3rd Qu.	55
6	Max	750

Based on table 2 above, it can be seen that the minimum depth is 1 km and the maximum is 750 km. The median depth occurs at a depth of 17 km where half of the earthquakes occur at depths shallower than 17 km. The mean depth is 48.41 km. The first quartile was 10 km, indicating that a quarter of the earthquakes were quite shallow, and the third quartile was 55 km, indicating that most of the earthquakes recorded occurred at relatively shallow to medium depths.

B. Clustering Analysis using Bisecting K-means

Clustering using Bisecting K-means was conducted with 10,809 earthquake event data in Indonesia. The number of clusters and centroid points obtained from the bisecting K-means process can be seen in Table 3

Table 3 *Centroid* points of each *cluster*

Cluster	*Centroid* 1	*Centroid* 2
1	4.229123	472.26164

2	3.331001	23.78306
3	3.615469	155.79556

Based on table 3 above, it can be seen that the results of bisecting K-means are 3 clusters with an SSE value of 8,490,668. So that the desired number of clusters is achieved. In addition, to assess the quality of the resulting clustering, testing can be done with the Davies Boulden Index method in determining the optimal number of clusters ranging from k = 2 to k = 6. The following can be seen in table 4.4 which is the result of the average value of Davies Bouldin Index for Bisecting K-means.

Table 4 Average value of *Davies Bouldin Index*

Number of *Clusters*	*Davies Bouldin Index*
k=2	0.7533
k=3	0.4758
k=4	0.5813
k=5	0.5735
k=6	0.6125

Table 4 above shows that the optimal number of clusters is 3 clusters with an average Davies Bouldin Index value of 0.4758. The average value produced by the Davies Bouldin Index method for k = 3 is the smallest value compared to the value of the number of other clusters, because in the Davies Bouldin Index method the smaller the value produced, it shows good results.

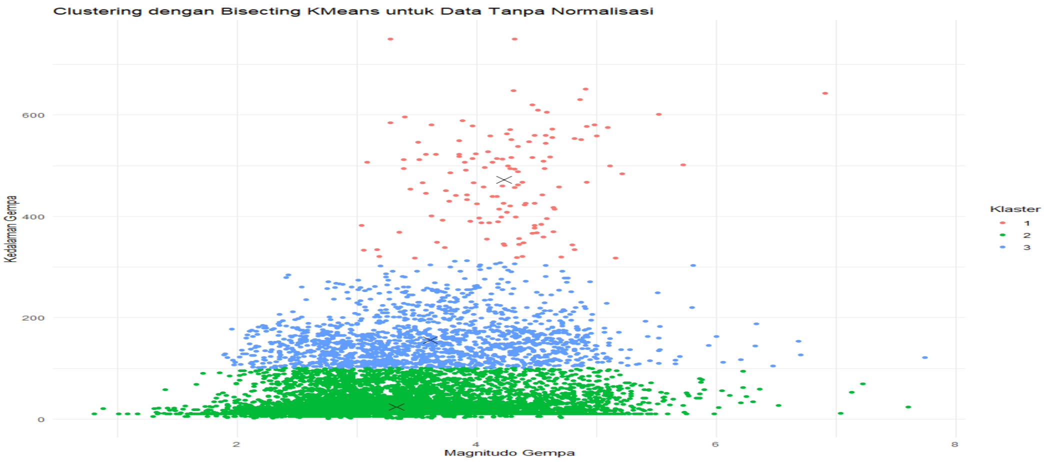

Figure 4 *Clustering* Results using *Bisecting K-means*

Based on Figure 4, it can be seen that the *clustering* results on the grouping of earthquake data in Indonesia in 2023 have 3 *clusters.* Each *cluster* represents a different combination of depth and magnitude. For example, *cluster* 1 covers depths called *Deep Earthquakes* and has *magnitudes* of around 3.03 SR to 6.91 SR. *Cluster* 3 covers depths called *Intermediate Earthquakes* and has a *magnitude of* 1.89 to 7.73. And *cluster* 2 covers a depth called *Shallow Earthquake* and has a *magnitude of* 0.81 to 7.59.

Table 5 Frequency of earthquake occurrence in each *cluster* based on *magnitude*

Magnitude	*Cluster*		
	1	2	3
< 5	122	8.916	1.528
5 - 7	8	184	46
> 7	0	4	1

Based on table 5 above, it can be seen that all magnitudes are spread across all *clusters.* However, it can be seen that *cluster* 2 is the *cluster* with the highest number of earthquake events.

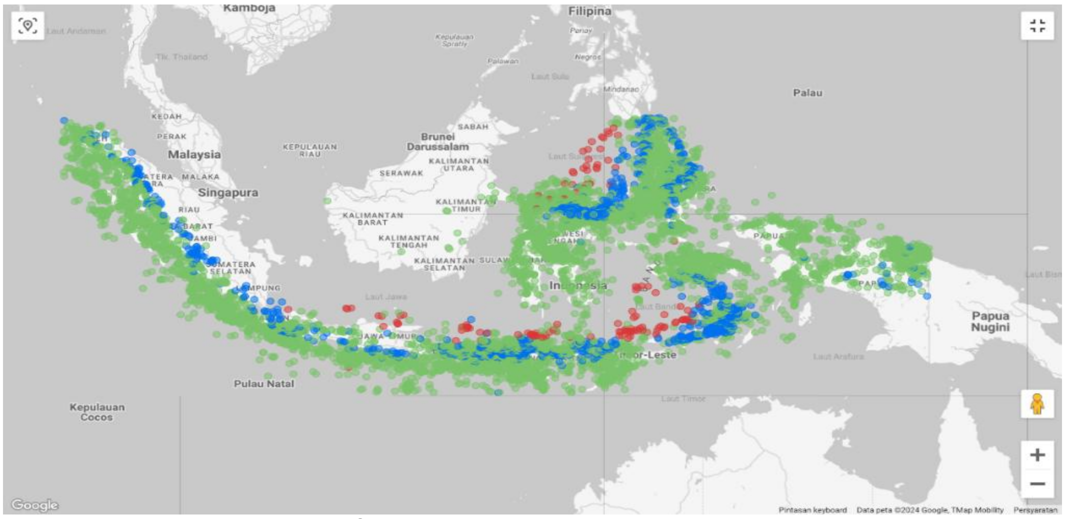

Figure 5 *Clustering* Map Results using *Bisecting K-means*

Based on 6 this seismic map of Indonesia shows the distribution of earthquakes with a classification based on depth, namely *cluster* 1 deep earthquakes, *cluster* 2 shallow earthquakes, and *cluster* 3 medium earthquakes. Light blue dots representing shallow earthquakes are widely distributed along plate tectonic boundaries, especially on the west coast of Sumatra, southern Java, Bali, Nusa Tenggara, and northern Sulawesi. Dark blue dots indicating medium earthquakes are found along subduction zones such as Sumatra, Java and Nusa Tenggara. Meanwhile, pink dots representing deep earthquakes are scattered across eastern Indonesia, including the Banda Sea, Maluku and Papua, reflecting deep subduction activity. This distribution pattern shows that Indonesia is located in an area of high seismic activity, reflecting frequent earthquakes and significant tectonic activity. Understanding this pattern is important for developing effective disaster mitigation measures.

V. CONCLUSIONS AND SUGGESTIONS

This study uses the Bisecting K-means method in clustering earthquake data in Indonesia in 2023, where the average depth of earthquakes is 48.41 km and the average magnitude of earthquakes is 3.3833 SR where the location of earthquakes often occurs in the Minahassa Peninsula region, Sulawesi by 10.3% recorded as many as 1,117 earthquake events occurred during 2023. This study uses the Bisecting K-means method in clustering earthquake data in Indonesia in 2023, the results show that the optimum k value in this study is 3 with an average value of Davies Bouldin Index without standardization of 0.4758.

The suggestions that can be given by the author based on the results of this study in order to add to the perfection of further research, namely for further research on earthquake data in Indonesia can use other clustering methods to obtain the best method in solving the same case study.

REFERENCES

[1] Anindya, F., & Oktaviana, P. P. (2022). Pemodelan Magnitude Gempa Bumi di Indonesia Menggunakan Generalized Extreme Value (GEV). Jurnal Sains Dan Seni Its, 11(6).

[2] Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. Hello World Jurnal Ilmu Komputer, 2(1), 20–33. https://doi.org/10.56211/helloworld.v2i1.230

[3] Arbaeti, E. E., Pardede, A. M. H., & Kadim, L. A. N. (2023). Application of K-Means Clustering Algorithm To Analyze Insurance Company Business (Case Study: Pt. Jasindo Insurance). Journal of Mathematics and Technology (MATECH), 2(2), 173–192.

[4] Arifin, A. (2023). Pengelompokan Titik Kejadian Gempabumi di Wilayah Asia Tenggara Menggunakan Agglomerative Hierarchical Clustering Clustering Earthquake Event Points in the Southeast Asia Region using Agglomerative Hierarchical Clustering. SISTEMASI: Jurnal Sistem Informasi, 12(September), 900–914.

[5] Ayub, M. (2007). Proses Data Mining dalam Sistem Pembelajaran Berbantuan Komputer. Jurnal Sistem Informasi, 21(1), 21–30.

[6] Dwididanti, S., & Anggoro, D. A. (2022). Analisis Perbandingan Algoritma Bisecting K-Means dan Fuzzy C-Means pada Data Pengguna Kartu Kredit. Emitor: Jurnal Teknik Elektro, 22(2), 110–117. https://doi.org/10.23917/emitor.v22i2.15677

[7] Febriyanti, L., & Zakaria, H. (2023). Implementasi Data Mining Untuk Memprediksi Produktivitas Pada Tanaman Kacang Tanah Menggunakan Metode Naive Bayes (Studi Kasus : Perkebunan Kacang Tanah Di Kota Bogor). LOGIC : Jurnal Ilmu Komputer Dan Pendidikan, 1(2), 105–118. https://journal.mediapublikasi.id/index.php/logic

[8] Firmansyah, T., Poningsih, & Andani, S. R. (2022). Analisis Clustering Algoritma K-Means Sebagai Rekomendasi Penambahan Koleksi Buku Di Perpustakaan Madrasah Tsanawiyah Negeri 2 Simalungun. Zahra: Buletin Big Data, Data Science and Artificial Intelligence, 1(1), 44–48.

[9] Hastari, D., Nurunnisa, F., Winanda, S., & Dwi Aprillia, D. (2023). Penerapan Algoritma K-Means dan K-Medoids untuk MengelompokkanData Negara Berdasarkan Faktor Sosial-Ekonomi dan Kesehatan. SENTIMAS: Seminar Nasional Penelitian Dan Pengabdian Masyarakat, 274–281. https://journal.irpi.or.id/index.php/sentimas

[10] Imro'ah, N., Ayuningtias, I., & Debataraja, N, N. (2019). Analisis Cluster Non-Hirarki Dengan Metode K-Modes. Bimaster : Buletin Ilmiah Matematika, Statistika Dan Terapannya, 8(4), 909–916. https://doi.org/10.26418/bbimst.v8i4.36633

[11] Isyfa Rhamdani, A., & Jamaludin, H. (2023). Pengelompokan Wilayah Menurut Kekuatan Gempa Bumi Menggunakan Clustering. Jurnal Media Pratama, 17(2), 149–158.

[12] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). An Introduction to Statistical Learning. Springer.

[13] K3, Tim Karakter. (2019). Pedoman K3 Gempuran: Gempa Bumi, Erupsi Gunung Merapi & Kebakaran. Universitas Negeri Yogyakarta.

[14] Lompoliuw, N. O., & Purnomo, H. D. (2023). Implementasi Algoritma K-Means untuk Pengelompokkan Lama Sembuh Pasien Covid-19. Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi), 7(2), 186–193. https://doi.org/10.35870/jtik.v7i2.706

[15] Imro'ah, N., Ayuningtias, I., & Debataraja, N, N. (2019). Analisis Cluster Non-Hirarki Dengan Metode K-Modes. Bimaster : Buletin Ilmiah Matematika, Statistika Dan Terapannya, 8(4), 909–916. https://doi.org/10.26418/bbimst.v8i4.36633

[16] Prasetio, A., Effendi, M. M., & Dwi M, M. N. (2023). Analisis Gempa Bumi Di Indonesia Dengan Metode Clustering. Bulletin of Information Technology (BIT), 4(3), 338–343. https://doi.org/10.47065/bit.v4i3.820

[17] Prataopu, R. D. (2013). Analisis Tingkat Kekerasan Tanah di Bawah Stasiun-Stasiun Seismik di Jawa Tengah Menggunakan Software Seisgram2K. Jurnal Inovasi Fisika Indonesia, 2(3), 27–36.

[18] Royal, S. (2024). Perancangan Aplikasi Data Mining Untuk Menentukan Tingkat Kelarisan Produk Menggunakan Metode Clustering Dengan Algoritma K-Means. Journal of Science and Social Research, 4307(1), 116–123. http://jurnal.goretanpena.com/index.php/JSSR

[19] Santosa, B. (2007). Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. Graha Ilmu.

[20] Tania, A., Handhayani, T., & Hendryli, J. (2023). Perbandingan Antara Algoritma K-Means Dan Algoritma Bisecting K-Means Dalam Menganalisis Gempa Bumi Di Indonesia. Simtek : Jurnal Sistem Informasi Dan Teknik Komputer, 8(2), 265–270. https://doi.org/10.51876/simtek.v8i2.205

[21] Virantika, E., Kusnawi, K., & Ipmawati, J. (2022). Evaluasi Hasil Pengujian Tingkat Clusterisasi Penerapan Metode K-Means Dalam Menentukan Tingkat Penyebaran Covid-19 di Indonesia. Jurnal Media Informatika Budidarma, 6(3), 1657. https://doi.org/10.30865/mib.v6i3.4325

[22] Wahono, R. S. (2023). Data Mining Data mining. In Mining of Massive Datasets (Vol. 2, Issue January 2013). https://www.cambridge.org/core/product/identifier/CBO9781139058452A007/type/book_part

[23] Yoliadi, D. N. (2023). Data Mining Dalam Analisis Tingkat Penjualan Barang Elektronik Menggunakan Algoritma K-Means. Insearch: Information System Research Journal, 3(01). https://doi.org/10.15548/isrj.v3i01.5829

[24] Zhou, Z., Ran, A., Chen, S., Zhang, X., Wei, G., Li, B., Kang, F., Zhou, X., & Sun, H. (2020). A Fast Screening Framework dor Second-Life Batteries Based on an Improved Bisecting K-Means Algorithm Combined with Fast Pulse Test. Journal of Energy Storage, 31, 1–32