Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th June 2025. Vol. 4. No. 3; e-ISSN: 2808-4519

Literature Review of Earthquake Clustering Algorithms in Indonesia

Abdul Hakim Prima Yuniarto[1]*, Devi Astri Nawangnugraeni[2]

[1]Prodi Teknik Elektro, Sekolah Tinggi Teknik Wiworotomo
[2]Prodi Informatika, Universitas Jenderal Soedirman
a.hakim.py@gmail.com [1]*

Abstract

This study presents a structured literature review, often referred to as a Systematic Literature Review (SLR), based on 18 articles discussing clustering methods. The primary aim of this study is to explore how clustering techniques have been applied to earthquake data in Indonesia. To achieve this, the study addresses four key research questions. First, it examines which algorithms are most commonly used for earthquake clustering in Indonesia. Second, it evaluates which algorithms demonstrate the best performance. Third, it investigates which regions within Indonesia are most frequently studied in this context. Finally, it analyzes the types of datasets that are most often utilized for earthquake clustering in the country. The findings indicate that the K-Means algorithm is not only the most frequently used but also consistently shows strong performance. In addition, earthquake clustering studies most commonly focus on Indonesia as a whole, using publicly available datasets. These insights offer valuable guidance for researchers seeking to apply or further develop clustering methods for earthquake-related studies in Indonesia.

Keywords: *Clustering; Earthquake; Indonesia; Literature Review; Machine Learning*

1. Introduction

In today's era, where the flow of information is rapid and abundant, data grouping or clustering techniques have gained increasing importance and are widely utilized across various research disciplines. Clustering allows researchers to identify patterns, relationships, and hidden structures in complex datasets more effectively [1]. As datasets become more intricate, clustering methods play an even more vital role in supporting the understanding and interpretation of information that may not be immediately apparent. One notable field that benefits from these techniques is disaster studies, particularly when studying earthquakes. Considering Indonesia's susceptibility to earthquakes, there is a need for further research to explore how clustering techniques can be employed to identify earthquake-prone zones in the country.

Several previous studies have highlighted that clustering techniques are commonly used to identify earthquake-prone zones in Indonesia [2]. Although this method has been recognized as an important tool, there has not yet been a comprehensive study that specifically discusses and evaluates its use within the context of earthquake disasters. This study is thus motivated by the need to understand how extensively clustering methods have been applied and to explore the latest trends and practices in their application for earthquake disaster analysis.

This study aims to provide a deeper understanding of the application of clustering methods in identifying earthquake-prone zones in Indonesia. To achieve this objective, several key research questions are addressed [3]. First, which algorithms are most commonly used in earthquake data clustering studies in Indonesia. Second, which algorithms have demonstrated the highest effectiveness in these studies. Third, which regions are most frequently the focus of earthquake clustering research. Finally, what types of datasets are most commonly used in earthquake clustering studies in Indonesia.

Thus, the results of this study are expected to serve as a useful reference for researchers who aim to apply or further develop clustering methods to identify earthquake-prone zones in Indonesia. Furthermore, a deeper understanding of the current trends in the use of these methods can contribute to the development of more effective and efficient approaches in the future.

2. Research Method

2.1. Systematic Literature Review

The approach used in this study is the Systematic Literature Review (SLR), which involves formulating predetermined research questions on a specific topic and then evaluating and synthesizing relevant research sources [4]. The purpose of the SLR is to collect and

summarize previous studies, identify gaps between past and current research that need to be addressed, and compile a coherent synthesis or research framework. Achieving this requires systematic and structured efforts throughout the review process.

The purpose of this literature study is to present comprehensive findings and offer a broader perspective on earthquake clustering methods in Indonesia. To achieve optimal results, the study utilized literature published in prominent journal databases, such as Google Scholar, covering the period from 2021 to 2025.

2.2. Research Question

Research questions are formulated to guide the direction of a scientific study. They serve to define the focus and objectives of the research being undertaken. Generally, the primary purpose of research questions is to provide a clear framework and direction, assist researchers in selecting appropriate methodologies, and facilitate a deeper understanding of the topic under investigation [5]. Research questions are formulated to determine the strategy employed in the study conducted by the researcher. Some of the main questions addressed in this study include:

Table 1: Research Question

Code	Research Question	Motivation
RQ1	What are the most frequently used algorithms for earthquake clustering in Indonesia?	Knowing the algorithms frequently used for earthquake clustering in Indonesia
RQ2	Which algorithm has the best performance for earthquake clustering in Indonesia?	Knowing the best performing algorithm for earthquake clustering in Indonesia
RQ3	Which areas are most frequently studied for earthquake clustering in Indonesia?	Knowing the most frequently studied areas for earthquake clustering in Indonesia
RQ4	What type of dataset is most often used for earthquake clustering in Indonesia?	Knowing the most frequently used dataset types for earthquake clustering in Indonesia

2.3. Article Quality Assesment

This study aims to conduct a systematic literature review. During the data collection process, researchers follow several stages, including identification, evaluation, and interpretation of relevant literature and reports. Related journals are reviewed and filtered through systematic procedures at each stage. Upon completion of the review process, the study primarily focuses on the findings, analytical methods, and open approaches used in the literature. The keyword used for the search was "earthquake clustering," with data sourced from the Google Scholar database covering the period from 2021 to 2025.

3. Result and Discussion

Based on a search of articles in the Google Scholar database from 2021 to 2025, several relevant articles were identified. After filtering to include only those focusing on earthquake clustering in Indonesia, 18 articles were selected. These selected articles were then reviewed and summarized in Table 2 as presented below.

Table 2: Article Title

NO	ARTICLE TITLE	AUTHOR & YEAR OF PUBLICATION	ALGORITHM
1	Analisis Gempa Bumi Di Indonesia Dengan Metode Clustering [6]	(Prasetio et al., 2023)	K-Means
2	Analisis Klasterisasi Kerawanan Gempa Bumi di Provinsi Papua Menggunakan Algoritma Invasive Weed Optimization (IWO) [7]	(Baisa et al., 2023)	Invasive Weed Optimization (IWO), K-Means, DBSCAN
3	Density-Based Clustering untuk Pemetaan Daerah Rawan Gempa Bumi di Wilayah Sumatera Barat Menggunakan Metode DBSCAN [8]	(Taufiq et al., 2024)	DBSCAN
4	Implementasi Rapidminer Pada Klasterisasi Gempa Bumi Di Indonesia Berdasarkan Kedalaman Menggunakan K-Means [9]	(Ubaidillah & Fatah, 2024)	K-Means
5	Klasterisasi Daerah Rawan Gempa Bumi di Indonesia Menggunakan Algoritma K-Medoids [10]	(Kurmiati et al., 2021)	K-Medoids
6	Klasterisasi Data Kejadian Gempa Bumi di Indonesia Menggunakan Metode K-Medoids [11]	(Inayah et al., 2024)	K-Medoids
7	Klasterisasi Wilayah Rentan Bencana Alam Berupa Gerakan Tanah Dan Gempa Bumi Di Indonesia [12]	(Setiawan et al., 2022)	DBSCAN, CNN, K-Medoids
8	Pemanfaatan Algoritma K-Means dalam Klasterisasi Gempa Sulawesi [13]	(Gunawan & Wibowo, 2024)	K-Means
9	Penerapan Metode K-Means Pada Klasterisasi Wilayah Rawan Gempa Di Indonesia [14]	(Dwitiyanti et al., 2023)	K-Means

10	Analisis Pengelompokan Gempa Bumi di Indonesia Berdasarkan Ruang-Waktu-Kekuatan Kedalaman [15]	(Anggraeni, 2024)	K-Means
11	Clustering Data Titik Gempa Dengan Metode Fuzzy Possibilistic C-Means [16]	(Putriana et al., 2021)	Fuzzy Possibilistic C-means
12	Implementasi Algoritma K-Means Pada Peristiwa Gempa Bumi di Wilayah Jawa Barat [17]	(Agustina & Mulyawan, 2023)	K-Means
13	Penerapan Metode Dbscan untuk Identifikasi Kluster Gempa Bumi di Daerah Yogyakarta [18]	(Ajitomo & Pratama, 2024)	DBSCAN
14	Pengelompokan Titik Gempa Di Pulau Sulawesi Menggunakan Algoritma St-Dbscan [19]	(Manalu et al., 2021)	ST-DBSCAN
15	Temporal Spatial Property Profiling And Identification Of Earthquake Prone Areas Using St-Dbscan And K-Means Clustering [20]	(Samsudin et al., 2024)	ST-DBSCAN dan K-Means
16	A Study of Grouping of Earthquake Damage from Magnitude Scale in Lombok Using K-Means Modeling [21]	(Kertanah et al., 2024)	K-Means
17	Implementation Of The Dbscan Method For Cluster Mapping Of Earthquake Spread Location [22]	(Bariklana & Fauzan, 2023)	DBSCAN
18	Agglomerative Clustering of 2022 Earthquakes in North Sulawesi, Indonesia [23]	(Siahaan & Rio, 2023)	K-Means

3.1. RQ1. What are the most frequently used algorithms for earthquake clustering in Indonesia?

The purpose of this research question is to identify the algorithms most frequently used for earthquake clustering in Indonesia. Based on the analysis of 18 selected articles, the algorithms applied in earthquake clustering include K-Means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), K-Medoids, Invasive Weed Optimization (IWO), Fuzzy C-Means, and Convolutional Neural Networks (CNN). Several studies employed multiple algorithms simultaneously and conducted comparative analyses.

Table 3: Number of Clustering Algorithms

Algorithm	Amount	Percentage (%)
K-Means	10	55.56
DBSCAN	7	38.89
K-Medoids	3	16.67
IWO	1	5.56
Fuzzy C-Means	1	5.56
CNN	1	5.56

Based on Table 3 above, it was found that the K-means algorithm was used in 10 articles or 55.56% of the total data. Then followed by the DBSCAN algorithm used in 7 articles or 38.89% of the total data. Then the K-Medoids algorithm used in 3 articles or 16.67% of the total data. Then for the IWO, Fuzzy C-Means, and CNN algorithms, each was used in 1 article or 5.56% of the total data.

3.2. RQ2. Which algorithm has the best performance for earthquake clustering in Indonesia?

The purpose of this research question is to identify the best-performing algorithm for earthquake clustering in Indonesia. Based on the analysis of 18 selected articles, only 10 articles (approximately 55%) reported evaluation metric calculations, while the remaining 8 did not. Among those that conducted evaluations, various metrics were used, including the Silhouette Coefficient (SC), Davies-Bouldin Index (DBI), and Sum of Squared Errors (SSE).

To facilitate the performance analysis of each algorithm, it is established that a Silhouette Coefficient (SC) value closer to 1 indicates better clustering quality, whereas lower values of the Davies-Bouldin Index (DBI) and Sum of Squared Errors (SSE) signify better performance.

Table 4: Clustering Algorithm Performance

Article no-	Algorithm	Performance
2	K-Means	Very Good
9	K-Means	Weak
12	K-Means	Very Good
16	K-Means	Very Good

2	DBSCAN	Good
3	DBSCAN	Very Good
7	DBSCAN	Weak
13	DBSCAN	Very Good
14	DBSCAN	Bad
17	DBSCAN	Very Good
6	K-Medoids	Pretty Good
7	K-Medoids	Weak
2	IWO	Very Good
7	CNN	Weak

Based on Table 4 above, it is observed that the majority of K-Means algorithm implementations demonstrate very good performance. This is followed by the DBSCAN algorithm, which shows partly very good performance. The K-Medoids algorithm exhibits only moderate performance, while the Invasive Weed Optimization (IWO) algorithm demonstrates very good performance. Lastly, the Convolutional Neural Network (CNN) algorithm shows relatively weak performance.

3.3. RQ3. Which areas are most frequently studied for earthquake clustering in Indonesia?

The purpose of this research question is to identify the regions in Indonesia that are most frequently studied for earthquake clustering. Based on the analysis of 18 selected articles, the majority of studies focused on the entire Indonesian region, while the remainder examined areas with high seismic activity.

Table 5: Clustering Research Area

Area	Amount	Percentage (%)
Entire Indonesia	7	38.89
Jawa	3	16.67
Sulawesi	3	16.67
Sumatra	2	11.11
NTB	2	11.11
Papua	1	5.56

Based on Table 5 above, it was found that the Indonesian region as a whole was studied in 7 articles, representing 38.89% of the total data. This was followed by the Java and Sulawesi regions, each studied in 3 articles (16.67%). The Sulawesi and Sumatra regions were each examined in 2 articles (11.11%), while the Papua region was studied in 1 article (5.56%).

3.4. RQ4. What type of dataset is most often used for earthquake clustering in Indonesia?

The purpose of this research question is to determine the types of datasets most frequently used for earthquake clustering in Indonesia. Based on the analysis of 18 selected articles, it was found that all of them utilized publicly available datasets, as shown in Table 6 below. These datasets were sourced from several institutions, including the Meteorology, Climatology, and Geophysics Agency (BMKG), the United States Geological Survey (USGS), and the Ministry of Energy and Mineral Resources (KESDM).

Table 6: Research Dataset

Dataset	Amount	Percentage (%)
Public	18	100.00
Private	0	0.00

4. Conclusion

Based on the results and discussion above, the following conclusions can be drawn:

1) The most frequently used algorithm for earthquake clustering in Indonesia is K-Means, which is 55.56% of the total data. This shows that the K-Means method has high popularity in clustering.
2) The best performing algorithm for earthquake clustering in Indonesia is K-Means, which has very good performance. This shows that the K-Means method has high effectiveness and accuracy in clustering.
3) Many studies on earthquake clustering in Indonesia still do not calculate the evaluation matrix, so the resulting clusters are not yet valid.

4) The most frequently studied area for earthquake clustering is the entire Indonesian region, which is 38.89% of the total data. This shows that research on earthquake clustering most often covers the entire Indonesian region.
5) The most frequently used dataset type for earthquake clustering in Indonesia is public dataset, which is 100% of the total data. This shows that the majority of research on earthquake clustering in Indonesia uses public dataset.

5. Future Research

Based on the research results above, there are suggestions that can be made as follows:
1) Using other algorithms besides K-Means as a comparison in earthquake clustering in Indonesia.
2) Conduct evaluation matrix calculations for each earthquake clustering study in Indonesia.
3) Conducting earthquake clustering for more specific regions in Indonesia.

Acknowledgement

Appreciation is extended to Sekolah Tinggi Teknik Wiworotomo for their invaluable support in facilitating this research. Their commitment and assistance ensured the study proceeded smoothly and fostered a collaborative and intellectually engaging environment.

References

[1] D. R. Adji *et al.*, "Metode dan Algoritma Dalam Data Clustering : Systematic Literature Review," *Sci. Technol. Manag. J.*, vol. 5, no. 1, pp. 1–4, 2025.
[2] A. Wahyu and R. Rushendra, "Klasterisasi Dampak Bencana Gempa Bumi Menggunakan Algoritma K-Means di Pulau Jawa," *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 1, p. 174, 2022.
[3] I. M. Faiza, G. Gunawan, and W. Andriani, "Tinjauan Pustaka Sistematis: Penerapan Metode Machine Learning untuk Deteksi Bencana Banjir," *J. Minfo Polgan*, vol. 11, no. 2, pp. 59–63, 2022.
[4] Sekar Setyaningtyas, B. Indarmawan Nugroho, and Z. Arif, "Tinjauan Pustaka Sistematis: Penerapan Data Mining Teknik Clustering Algoritma K-Means," *J. Teknoif Tek. Inform. Inst. Teknol. Padang*, vol. 10, no. 2, pp. 52–61, 2022.
[5] P. Sari, E. Efan, and R. Syahri, "Algoritma K-Means Clustering : Sebuah Studi Literatur K-Means Clustering Algorithm : a Literatur Study," *J. Inform.*, vol. x, no. x, pp. 1–7, 2023.
[6] A. Prasetio, M. M. Effendi, and M. N. Dwi M, "Analisis Gempa Bumi Di Indonesia Dengan Metode Clustering," *Bull. Inf. Technol.*, vol. 4, no. 3, pp. 338–343, 2023.
[7] L. Y. Baisa, D. Manongga, and Y. Nataliani, "Analisis Klasterisasi Kerawanan Gempa Bumi di Provinsi Papua Menggunakan Algoritma Invasive Weed Optimization (IWO)," *J. Edukasi dan Penelit. Inform.*, vol. 9, no. 2, p. 176, 2023.
[8] R. M. Taufiq, R. Firdaus, F. Handayani, P. F. Muarif, and R. R. Rizqy, "Density-Based Clustering untuk Pemetaan Daerah Rawan Gempa Bumi di Wilayah Sumatera Barat Menggunakan Metode DBSCAN," *J. FASILKOM*, vol. 14, no. 3, pp. 817–822, 2024.
[9] M. Ubaidillah and Z. Fatah, "IMPLEMENTASI RAPIDMINER PADA KLASTERISASI GEMPA BUMI DI INDONESIA BERDASARKAN KEDALAMAN MENGGUNAKAN K-MEANS," *J. Ilm. MULTIDISIPLIN ILMU*, vol. 1, no. 6, pp. 84–91, 2024.
[10] D. Kurmiati, M. Zakiy Fauzi, and A. Falegas, "Klasterisasi Daerah Rawan Gempa Bumi di Indonesia Menggunakan Algoritma K-Medoids," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. April, pp. 47–57, 2021.
[11] J. Inayah, A. Fanani, and W. D. Utami, "Klasterisasi Data Kejadian Gempa Bumi di Indonesia Menggunakan Metode K-Medoids," *JUSTIN (Jurnal Sist. dan Teknol. Informasi)*, vol. 12, no. 2, pp. 271–276, 2024.
[12] I. N. Setiawan, D. Krismawati, S. Pramana, and E. Tanur, "Klasterisasi Wilayah Rentan Bencana Alam Berupa Gerakan Tanah Dan Gempa Bumi Di Indonesia," *Semin. Nas. Off. Stat.*, vol. 2022, no. 1, pp. 669–676, 2022.
[13] W. Gunawan and A. Wibowo, "Pemanfaatan Algoritma K-Means dalam Klasterisasi Gempa Sulawesi," *Fakt. Exacta*, vol. 17, no. 3, pp. 228–240, 2024.
[14] N. Dwitiyanti, S. Ayu Kumala, and S. Dwi Handayani, "PENERAPAN METODE K-MEANS PADA KLASTERISASI WILAYAH RAWAN GEMPA DI INDONESIA Implementation of K-Means Method in Classterization of Earthquake Prone Areas in Indonesia," *Pros. Semin. Nas. UNIMUS*, vol. 6, pp. 1029–1037, 2023.
[15] W. Anggraeni, "Analisis Pengelompokan Gempa Bumi di Indonesia Berdasarkan Ruang-Waktu-Kekuatan Kedalaman," *Nucleus*, vol. 4, no. 2, pp. 136–160, 2024.
[16] S. Putriana, E. Ernawati, and D. Andreswari, "Clustering Data Titik Gempa Dengan Metode Fuzzy Possibilistic C-Means (Studi Kasus: Titik Gempa Pulau Sumatera Tahun 2013- 2018)," *Rekursif J. Inform.*, vol. 9, no. 1, 2021.
[17] M. Agustina and Mulyawan, "Implementasi Algoritma K-Means pada Peristiwa Gempa Bumi di Wilayah Jawa Barat," *J. Wahana Inform.*, vol. 2, no. 2, pp. 257–264, 2023.
[18] W. Ajitomo and I. Pratama, "Penerapan Metode Dbscan untuk Identifikasi Kluster Gempa Bumi di Daerah Yogyakarta," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 40–46, 2024.
[19] D. J. Manalu, R. Rahmawati, and T. Widiharih, "PENGELOMPOKAN TITIK GEMPA DI PULAU SULAWESI MENGGUNAKAN ALGORITMA ST-DBSCAN (Spatio Temporal-Density Based Spatial Clustering Application with Noise)," *J. Gaussian*, vol. 10, no. 4, pp. 554–561, 2021.
[20] A. R. Samsudin, D. H. Fudholi, and L. Iswari, "Temporal Spatial Property Profiling and Identification of Earthquake Prone Areas Using St-Dbscan and K-Means Clustering," *J. Tek. Inform.*, vol. 5, no. 3, pp. 917–929, 2024.
[21] K. Kertanah, A. Chintyana, C. Chandrawati, M. Rosiana, and N. Putri, "A Study of Grouping of Earthquake Damage from Magnitude Scale in Lombok Using K-Means Modeling," *Kappa J.*, vol. 8, no. 3, pp. 399–403, 2024.
[22] M. Bariklana and A. Fauzan, "Implementation of the Dbscan Method for Cluster Mapping of Earthquake Spread Location," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 2, pp. 0867–0878, 2023.
[23] M. B. Siahaan and A. R. Rio, "Agglomerative Clustering of 2022 Earthquakes in North Sulawesi, Indonesia," *Buana Inf. Technol. Comput. Sci. (BIT CS*, vol. 4, no. 2, pp. 76–84, 2023.