Smart Seismic Intelligence Machine Learning for Spatial Clustering and Earthquake Magnitude Prediction in Indonesia

*[1]Harry Setya Hadi, [2]Rosnita Rauf, [3]Agus Salim, [4]Kevin Maulana Firdaus
[12]Electrical Engineering, Faculty of Engineering, Ekasakti University, Padang
[3]Dinas Kominfo Kota Padang
[4]Information Systems, Faculty of Engineering, State University of Surabaya
[1]xmoensen@gmail.com, [2]ekasakti5974@gmail.com, [3]gestjava@gmail.com, [4]kevin.23100@mhs.unesa.ac.id

Abtract - Indonesia is located within the Pacific Ring of Fire, one of the most seismically active regions in the world due to the interaction of multiple major tectonic plates. Understanding the spatial distribution of earthquakes and accurately estimating their magnitudes is essential for effective disaster risk assessment and mitigation planning. This study aims to analyze earthquake distribution patterns and develop a machine learning-based approach to predict earthquake magnitude using seismic data from the Meteorology, Climatology, and Geophysics Agency (BMKG). The study employs two machine learning methods: K-Means Clustering to identify spatial groupings of earthquake events and Random Forest Regression to predict magnitude based on spatial and temporal features. The dataset consists of 67 earthquake events recorded in February 2026, including attributes such as latitude, longitude, depth, magnitude, and occurrence time. Clustering results indicate that the optimal number of clusters is k = 4, with a Silhouette Score of 0.3444, suggesting a moderate clustering structure. This implies that spatial patterns are present, although cluster separation is not yet well-defined. The Random Forest model achieved an R^2 of 0.7382 on training data and 0.0975 on testing data, indicating overfitting likely due to the limited dataset size. Feature importance analysis reveals that longitude contributes the most (43.7%), followed by depth (29.6%), latitude (20.6%), and time (6.0%). These findings highlight the dominant role of spatial factors in Indonesia's seismic activity. However, the limited dataset restricts model generalization; therefore, future studies should use larger datasets and incorporate additional geophysical parameters to improve predictive performance.

Keywords — Data Mining, K-Means Clustering, Random Forest, Gempa Bumi, Machine Learning

I. Introduction

Indonesia is located in the Pacific Ring of Fire, an area with the highest levels of volcanic and earthquake activity in the world. As an archipelagic nation, Indonesia also lies at the convergence of three major tectonic plates: the Eurasian, Indo-Australian, and Pacific Plates. Indonesia experiences an average of 5,000-6,000 earthquakes annually[1].

Earthquake data collected by the Meteorology, Climatology, and Geophysics Agency (BMKG) provides crucial information that can be used to analyze seismic activity patterns and characteristics in various regions of Indonesia. With advances in machine learning and data mining technology, earthquake data analysis can be performed more efficiently, uncovering hidden patterns undetected by conventional methods[2].

This research aims to apply machine learning algorithms to BMKG earthquake data to identify spatial and temporal patterns of earthquakes in Indonesia. The two main approaches used are unsupervised learning with K-Means Clustering for regional segmentation and supervised learning with Random Forest for earthquake magnitude prediction.

Although various studies have applied machine learning techniques to earthquake analysis, several limitations remain. Most previous research has focused on the application of deep learning methods or on large-scale global seismic datasets. Meanwhile, research utilizing national seismic data specifically to analyze the spatial patterns of earthquakes in Indonesia is still relatively limited.

Furthermore, most studies focus solely on earthquake magnitude prediction without incorporating spatial clustering analysis of seismic regions. Identifying spatial patterns of earthquake occurrence is crucial for understanding the distribution of tectonic activity in a region. Therefore, this study attempts to address this gap by applying the K-Means Clustering method for earthquake region segmentation and Random Forest Regression for magnitude prediction, using the BMKG earthquake dataset. This approach is expected to provide an initial understanding of earthquake distribution patterns in Indonesia and demonstrate the potential for applying machine learning to national seismic data analysis.

A. Problem Formulation

Based on the background presented, the problems to be addressed in this study can be formulated as follows:

a. How can we group earthquake areas in Indonesia based on magnitude, depth, and geographic location?

b. How can we predict earthquake magnitude based on location, depth, and time?

c. What features are most influential in predicting earthquake magnitude?

B. K-Means Clustering

K-Means Clustering is an unsupervised learning algorithm used to group data into k clusters based on similar characteristics or features. This algorithm works by minimizing the total squared distance between each data point and the closest cluster center (centroid).

[3]. The iterative K-Means process includes: (1) centroid initialization, (2) data assignment to the closest cluster, and (3) centroid updating until convergence.

C. Random Forest

Random Forest is an ensemble learning method that combines several decision trees to improve prediction accuracy while minimizing the risk of overfitting. This algorithm uses bagging (bootstrap aggregation) and random feature selection techniques to build a collection of independent decision trees [4]. Prediction results are obtained through voting (classification) or averaging (regression) across all trees in the forest.

D. Related Research

Several recent studies have been conducted to analyze earthquake data using machine learning [5]. A systematic review of Artificial Intelligence-based Earthquake Early Warning (EEW) systems found that a hybrid approach performed best. A study applied various machine learning techniques to earthquake seismology and demonstrated their effectiveness in seismic signal detection [6]. A study used deep learning with a dynamic loss function for earthquake prediction with promising results [2]. A study developed a deep residual network (CRED) for earthquake signal detection using recurrent neural networks [7]. A study applied deep learning to predict earthquake magnitude in the Horn of Africa with high accuracy.

Research in Asia has also shown significant progress, developing a real-time earthquake magnitude estimation method using deep learning with multiple seismometers. A study used explainable AI (XAI) models to estimate the spatial probability of earthquakes in the Arabian Peninsula [8]. A study applied machine learning to earthquake prediction in the North Zagros with competitive results.

II. Research Method

A. Dataset

The earthquake data used in this study is derived from the official earthquake catalog of the Agency for Meteorology, Climatology and Geophysics (BMKG) and is provided via the national seismic data portal. Data access is automated via an XML service at

https://data.bmkg.go.id/DataMKG/TEWS/autogempa.xml.

This service provides current earthquake information in real time and has been seismically validated by the BMKG. The system developed in this study retrieves the data automatically, so that all updates to the earthquake information are stored directly in the database of the Ekasakti University server as part of the research data repository. The stored data is then used for machine learning, analysis, and modeling. Furthermore, the stored earthquake data is visualized via a web-based application at https://app.unespadang.ac.id/Gempa, enabling more interactive earthquake monitoring.

The dataset used in this study comprises earthquake data from the BMKG (Meteorological, Climatological and Geophysical Office) covering the period from February 4 to March 1, 2026. It consists of 67 earthquake records with the following key characteristics: geographic coordinates (latitude and longitude), magnitude, depth, time of occurrence, and area information.

The data were extracted from the BMKG earthquake catalog, which is publicly accessible via the national seismic data portal.

Table 1. hows the descriptive statistics of the dataset:

Parameter	Min	Max	Mean ± SD
Magnitude	2.2	7.2	3.99 ± 1.09
Kedalaman (km)	1.0	628.0	31.81 ± 79.38
Latitude	-9.89	6.94	-2.45 ± 4.89
Longitude	95.60	140.62	114.35 ± 12.87

B. Research Flowchart

This research process was systematically conducted in several main phases to analyze earthquake distribution patterns and predict magnitude using a machine learning method. The research phase began with the collection of earthquake datasets from the earthquake catalog of the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG), which provides real-time earthquake data through its national seismic data service. The collected data were then stored in a research database for further processing and analysis.

The next step involved data preprocessing, which included data cleaning, attribute format conversion, and data normalization to adapt to the requirements of the machine learning algorithm. After preprocessing was complete, feature extraction was performed to determine the key variables for analysis, such as geographic coordinates (latitude and longitude), earthquake depth, and time of occurrence.

Subsequently, the K-means clustering method was applied to identify spatial cluster patterns of earthquake events in Indonesia. The clustering results then served as the basis for further analysis using random forest regression to predict earthquake magnitude based on the available spatial and temporal parameters.

The next step involved model evaluation, which aimed to assess the performance of the machine learning model used. This evaluation was conducted using various metrics, including the coefficient of determination (R^2), the mean squared error (RMSE), and the mean absolute error (MAE) for the regression model, as well as the silhouette score to assess the quality of the clustering results. The evaluation results were subsequently analyzed further to understand seismic activity patterns and factors influencing earthquake magnitude.

The entire research process is illustrated in Figure 1. It demonstrates the data analysis process from data acquisition to the interpretation of the results.

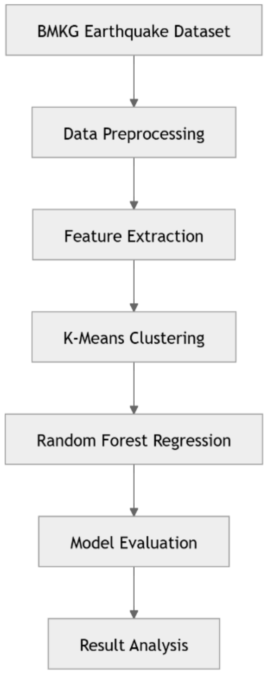

Figure 1: Research flowchart

C. Data Preprocessing

The preprocessing includes: (1) extracting coordinates from the string format, (2) converting the depth to numerical values, (3) extracting time features (hours), and (4) standardizing the features using the StandardScaler to ensure a uniform scale for each variable and to avoid the dominance of individual values in certain features.

D. K-Means Clustering

The main goal of the K-Means algorithm is to minimize the variance within the clusters. Mathematically, the objective function of K-Means can be formulated as follows:

$$J = \sum_{i=1} \sum_{\chi_j \in C_i} ||x_j - \mu_i||^2 \qquad (1)$$

Where:

- C_i is the cluster to -i
- x_j is the jth data that is in cluster C_i
- μ_i is the centroid of the i-th cluster

The K-means algorithm works iteratively by updating centroid positions and assigning data to clusters until convergence is achieved. The K-means algorithm is implemented with the following features: latitude, longitude, magnitude, and depth. The optimal number of clusters is determined using the elbow method and silhouette scoring. The parameters used are: max_iter=300, n_init=10, random_state=42.

E. Random Forest Regressor

Random Forest is an ensemble learning method that combines multiple decision trees to improve prediction stability and accuracy. The final model prediction is determined by averaging the predictions of all decision trees.

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(x) \qquad (2)$$

Where :

- T is the number of decision trees in the model
- $f_t(x)$ is the prediction function of the t–th decision tree
- $\hat{y}$ is the final predicted value

To predict magnitude, a random forest model with the following input features was used: latitude, longitude, depth, and time of day. The dataset was split into a training dataset (70%) and a test dataset (30%). Hyperparameters: n_estimators = 50, max_depth = 5, min_samples_split = 5, min_samples_leaf = 3.

F. Model Evaluation

The performance of the regression model was assessed using various evaluation metrics, including the coefficient of determination (R^2), the mean squared error (RMSE), and the mean absolute error (MAE). During the clustering process, the quality of the resulting clusters was measured using the silhouette score as an indicator of the degree of separation and compactness between the clusters.

III. Reslts and Discusion

A. Results of K-Means Clustering

Based on the elbow method and silhouette score analysis, the optimal number of clusters k = 4 was determined with a silhouette score of 0.3444. Figure 2 shows the results of the cluster optimization.

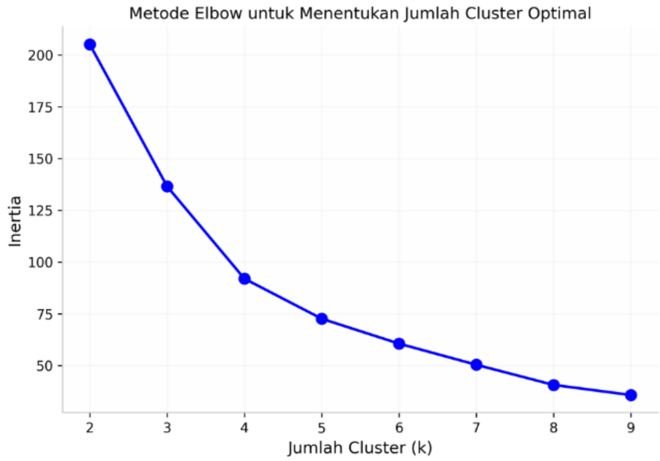

Figure 2: Determining the Optimal Number of Clusters with the Elbow Method and Silhouette Score

The results of applying the elbow method serve to determine the optimal number of clusters (k) in the K-means clustering process for earthquake data. The graph shows the relationship between the number of clusters (k) on the horizontal axis and the inertia value on the vertical axis. The so-called "elbow point" marks the optimal number of clusters [11].

Inertia is a measure of the squared distance between each data point and the nearest cluster center. Mathematically, it is calculated as the sum of the squares within the clusters (WCSS) and indicates the compactness of a cluster. The smaller the inertia value, the better the clustering, as the points within the cluster are closer to the center [12].

The graph shows that the inertia value decreases significantly with increasing cluster number from k = 2 to k = 3 to k = 4. From k = 4 onward, however, the decrease is flatter and less drastic. The inflection point of the slope (kink angle) appears to be at k = 4. This point represents the optimal balance between model complexity and cluster quality. In simpler terms: If the number of clusters is too small, data with different features are grouped together, leading to high internal cluster variance. Conversely, if the number of clusters is too large, the model becomes too complex and may suffer from oversegmentation, where small differences are treated as separate groups. Therefore, choosing k=4 offers the best compromise between pattern accuracy and model efficiency [13].

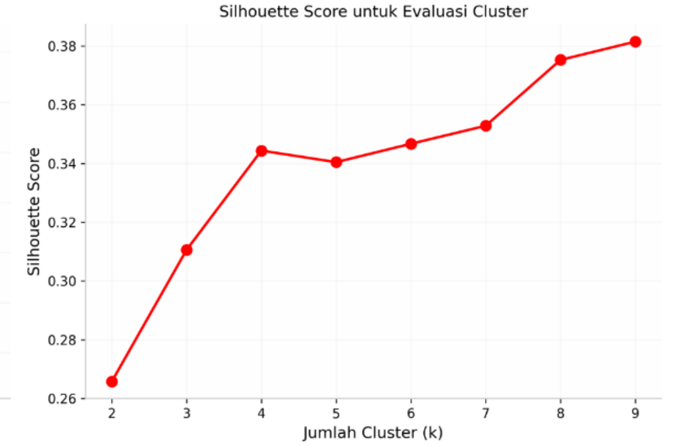

Figure 3: Silhouette Score for Cluster Evaluation

A silhouette score of 0.3444 indicates a medium quality of cluster structure. In general, silhouette scores between 0.25 and 0.50 indicate that the cluster structure is beginning to form, but the separation between the clusters is not yet fully developed. Although the cluster analysis has successfully identified the spatial grouping pattern of the earthquakes, the cluster boundaries still exhibit some overlap.

This graph illustrates the silhouette score as a method for evaluating the quality of cluster results in the K-means algorithm. It shows the relationship between the number of clusters (k) on the horizontal axis and the silhouette score on the vertical axis [13].

This graph illustrates the silhouette score as an evaluation method for assessing the quality of cluster results in the K-means algorithm. It shows the relationship between the number of clusters (k) on the horizontal axis and the silhouette score on the vertical axis [14].

The silhouette score is a metric that measures how well data fits into its cluster compared to other clusters. Mathematically, the silhouette score for each point is calculated by comparing its average distance to members of the same cluster (cohesion) and its average distance to other nearby clusters (separation). Silhouette scores range from -1 to 1. The closer the score is to 1, the better the separation between the clusters and the more compact the cluster structure. [15]

The graph shows that the silhouette score gradually increases from k=2 to k=9. A significant increase is observed between k=2 and k=4, after which it remains relatively stable with a further gradual increase. A score at k=4 is already quite good compared to k=2 and k=3, indicating that the cluster structure is becoming increasingly clear and spatially separated.

Although the highest silhouette scores occur at k=8 or k=9, the increase after k=4 is not significant. In the context of cluster analysis, small increases in the number of clusters often indicate a more detailed subdivision, but do not necessarily provide stronger interpretive information. Therefore, the choice of k=4 remains relevant, as it strikes a balance between the quality of cluster separation and the simplicity of the model [16].

Table 2. Shows the characteristics of each cluster formed:

Cluster	Region	Amount	SR	Depth	Coordinate
0	Aceh-Sumatera	13	3.61	16.6 km	3.25, 97.51
1	Papua-Maluku	23	4.53	22.7 km	-1.49, 128.11
2	Jawa-Bali-NTT	30	3.63	25.5 km	-6.39, 111.69
3	Gempa Dalam	1	7.20	628.0 km	6.94, 116.26

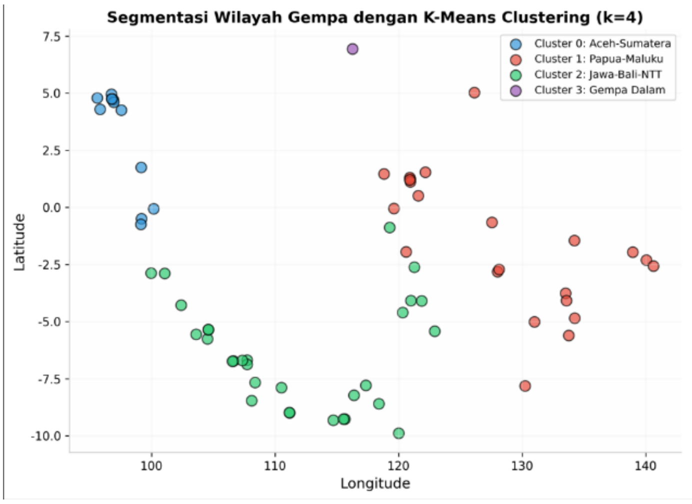

Figure 4: Earthquake Area Segmentation with K-Means Clustering (K=4)

The results of spatial segmentation of earthquake events in Indonesia using the K-means clustering algorithm with four clusters (k = 4) are presented. This visualization shows the distribution of earthquakes based on geographic coordinates (latitude and longitude), thus enabling the quantitative and visual observation of clustering patterns in seismic areas [17].

The horizontal axis represents longitude, and the vertical axis represents latitude. Each point represents a single earthquake event, and the point's color indicates the result of the K-means clustering. The K-means process minimizes the distance within the clusters (in-cluster variance), grouping points with similar spatial proximity into the same cluster.

The segmentation results reveal four main clusters of strong tectonic significance. Cluster 0 (blue) is concentrated in the Aceh–North Sumatra region. This region lies within an active subduction zone between the Indo-Australian and Eurasian plates, resulting in relatively high and spatially concentrated seismic activity. Cluster 1 (red) encompasses the Papua-Maluku region and eastern Indonesia. This area is known for its complex tectonic structure, created by the interaction of

multiple microplates. This results in a dispersed earthquake distribution, yet it forms consistent regional clusters [9].
Cluster 2 (green) represents the Java-Bali-Nusa Tenggara (NTT) region. This cluster exhibits an elongated earthquake distribution along the southern Indonesian subduction zone. This elongated pattern reflects the agreement between the K-means calculation results and the actual geological conditions, as this zone is among the most seismically active in Indonesia. Cluster 3 (purple) represents earthquake events with special features, possibly deep earthquakes or spatial anomalies that do not fully follow the dominant horizontal pattern. These points are spatially separated and form their own clusters[1]. From a scientific perspective, this visualization shows that unsupervised learning methods such as K-means are able to identify natural groupings in earthquake data without the use of initial labels. The resulting clusters are not only mathematically meaningful but also geophysically and tectonically relevant. This suggests that the earthquake distribution in Indonesia does indeed form seismic zones that can be modeled using computational methods[2].

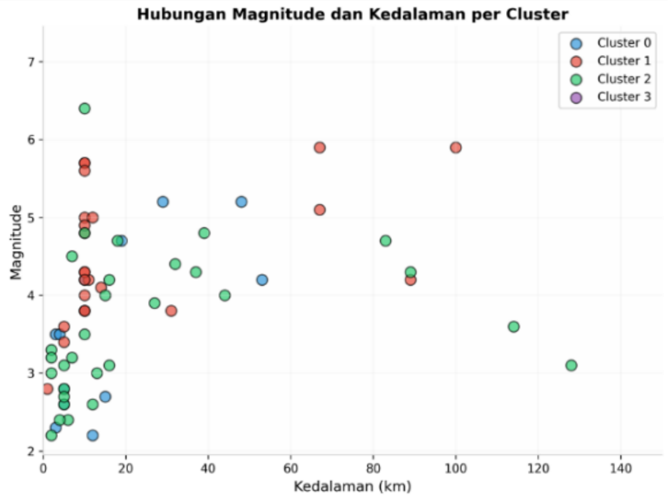

Figure 5: Relationship between magnitude and depth per cluster

This visualization serves to analyze the physical properties of earthquakes based on the previously identified seismic regions. The horizontal axis shows the earthquake depth (km), and the vertical axis the magnitude. Each point represents a single earthquake event, and the color indicates the cluster affiliation. This allows us to observe how the magnitude distribution changes with depth within each cluster.
In general, most earthquakes occur at shallow depths (< 20 km) with magnitudes between 2.5 and 5.5. This suggests that Indonesia's seismic activity is dominated by shallow earthquakes, which, from a geophysical perspective, are more likely to be felt at the surface and affect infrastructure. The concentration of points at shallow depths indicates that tectonic energy is primarily released in the upper crust and in shallow subduction zones [3].
Cluster 0 (blue) shows a distribution of earthquakes at shallow to medium depths (approximately 5–50 km) with relatively stable magnitudes in the range of 2.5–5.2. This pattern exhibits relatively uniform regional earthquake characteristics without extreme anomalies at great depths.
Cluster 1 (red) shows greater magnitude variation at shallow to medium depths. Several events with magnitudes near 6 at depths of approximately 60–100 km have been observed, indicating the potential for medium-energy earthquakes in the subduction zone. This reflects the tectonic complexity of eastern Indonesia.
Cluster 2 (green) shows the greatest depth variation, including several earthquakes at depths exceeding 100 km. However, the magnitudes at these greater depths are not always high and range between 3 and 4.5. This suggests that greater depth does not directly correlate linearly with higher magnitude. From a scientific perspective, deep earthquakes often have less impact on the surface than shallow earthquakes of similar magnitude. [18] Cluster 3 (purple), however, represents more isolated events with specific characteristics (likely deep earthquakes with relatively high magnitudes) that differ from the distribution of the other clusters.
Overall, this graph shows that there is no strong linear relationship between depth and earthquake magnitude. Most strong earthquakes occur at shallow to intermediate depths. This finding is important with regard to risk reduction, as shallow earthquakes of intermediate magnitude can have significant impacts on densely populated areas. [8]

B. Results of the Random Forest Regressor

The Random Forest Regressor model was evaluated using the training and test datasets. The results are presented in Table 3.

Table 3: Random Forest Model Evaluation

Dataset	R^2	RMSE	MAE
Training Set	0.7382	0.5454	0.4116
Test Set	0.0975	1.0613	0.8756

The feature importance analysis shows that longitude contributes the most to magnitude prediction at 43.7%, followed by depth (29.6%), latitude (20.6%), and hour (6.0%). Figure 3 illustrates the results of the prediction analysis.

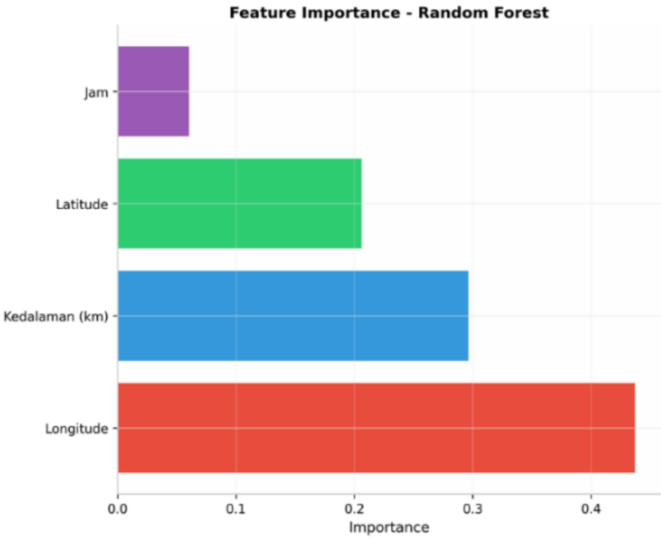

Figure 6: Feature Importance – Random Forest

The first section (Feature Importance – Random Forest) shows the contribution of each variable to magnitude prediction. Longitude is the most important, followed by depth, latitude, and time of occurrence. The dominance of longitude and latitude suggests that spatial factors play the largest role in determining earthquake magnitude in Indonesia [12]. This is consistent with Indonesia's tectonic location at the intersection of several large global plates, making geographic distribution an important indicator of seismic patterns. Depth also contributes significantly, as shallow earthquakes generally have different energy release characteristics than deep earthquakes. In contrast, the time of occurrence has a relatively small influence, suggesting that earthquake magnitude is not influenced by diurnal patterns but rather by subsurface geological dynamics [1].

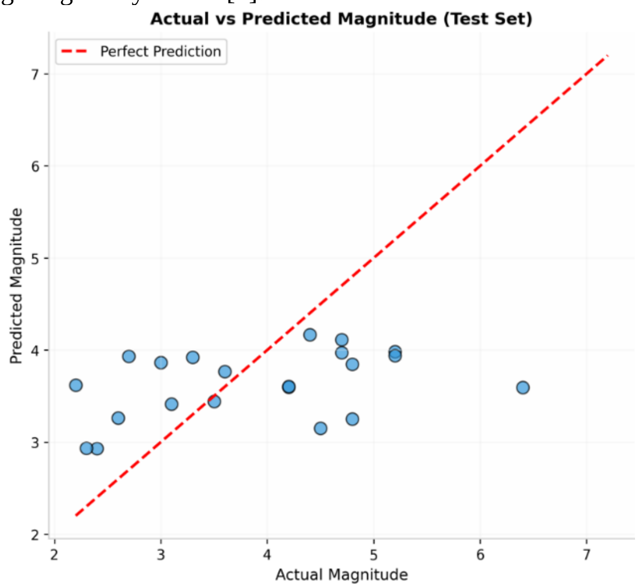

Figure 7: Actual vs. Predicted Magnitude (Test Dataset)

The second visualization (Actual vs. Predicted Magnitude) shows a comparison between the actual magnitude values and the model's predictions based on the test data. The red diagonal represents a perfect prediction, where the actual values match the predicted values. Most data points lie very close to this line, indicating that the model has good generalization ability in the low to medium magnitude range. However, deviations are observed at some higher magnitude values, suggesting that the model tends to slightly underestimate large earthquakes. This phenomenon is common in machine learning when the amount of data in the extreme categories is relatively small, allowing the model to learn more from the majority pattern (small to medium earthquakes). Overall, however, the model shows stable prediction consistency [19].

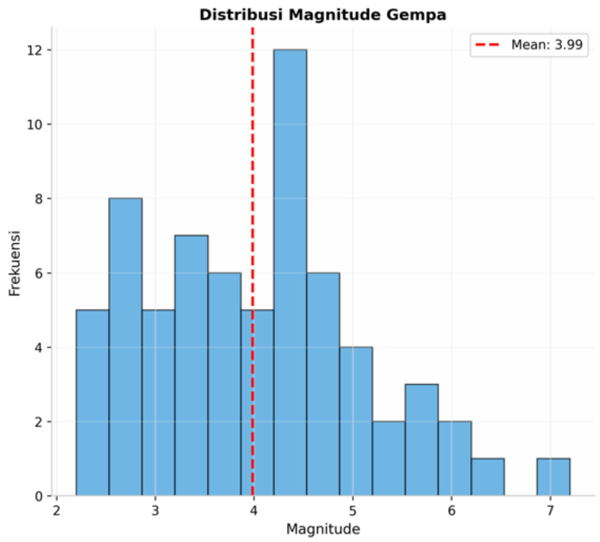

Figure 8: Distribution of earthquake magnitude

The histogram shows that most events have a magnitude between 3.0 and 4.5 on the Richter scale, with a mean of about 3.99. This distribution tends to be right-skewed, indicating that strong earthquakes are relatively rare compared to small and medium-sized earthquakes. This pattern is consistent with the general characteristics of seismic activity in Indonesia, where small earthquakes are more common as part of tectonic energy release. This uneven distribution also affects model performance, as unbalanced data can influence the model's sensitivity to extreme events.[20]

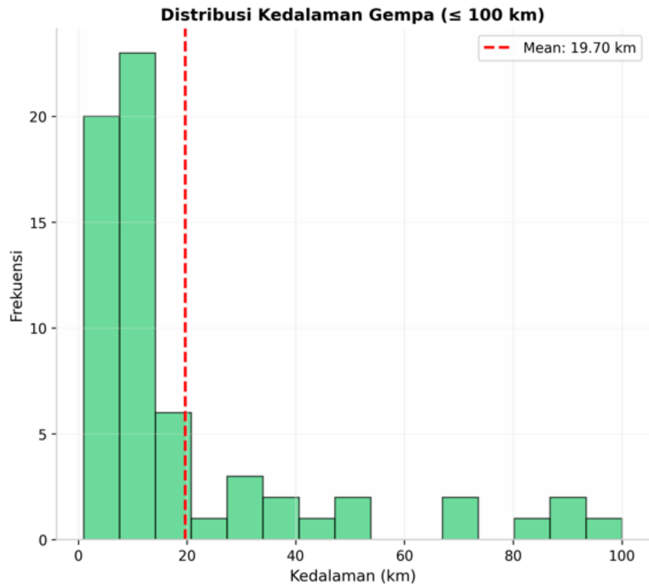

Figure 9: Earthquake Depth Distribution

The earthquake depth distribution shows that most events occurred at depths of less than 20 km, with an average depth of approximately 19.70 km. This suggests a predominance of shallow earthquakes, which, from a geophysical perspective, are more likely to be felt at the surface and affect infrastructure.

The occurrence of several earthquakes at greater depths (up to 100 km) suggests variations in the earthquake focus mechanisms, including possible activity in subduction zones or deep plates. This distribution confirms the results of the feature importance analysis, in which depth is a significant variable in the prediction model [21].
The cluster analysis reveals four distinct earthquake regions, each with its own characteristics. Cluster 0 (Aceh-Sumatra) has an average magnitude of 3.61 at a depth of 16.6 km and is typical of shallow earthquakes of medium intensity. Cluster 1 (Papua-Maluku) had the highest magnitude (4.53) and extended across eastern Indonesia. Cluster 2 (Java-Bali-NTT) was the largest cluster with 30 earthquakes (44.8%), reflecting the high seismic activity in this densely populated region. Magnitude prediction using Random Forest performed well in the training dataset ($R^2 = 0.7382$) but deteriorated in the test dataset ($R^2 = 0.0975$). This suggests overfitting due to the relatively small dataset size (67 samples). Longitude analysis yielded the strongest predictor variable, consistent with the geographical fact that eastern Indonesia (with its higher longitude) experiences more intense earthquake activity due to tectonic plate interactions [22]. Limitations of this study include: (1) the small dataset size, (2) the lack of geological features such as fault types and subsurface structures, and (3) the failure to consider long-term temporal factors. For future research, it is recommended to use a larger dataset and to include geophysical features to improve prediction accuracy [23].

IV. Conslusion

1. Based on the research findings, it can be concluded that machine learning provides initial insights into earthquake distribution in Indonesia. The K-means clustering method successfully identified four clusters of seismic regions representing the spatial distribution of earthquake events. The random forest model also showed that spatial parameters, particularly longitude and depth, have a significant influence on the estimation of earthquake magnitude.
2. However, the model's performance with the test data indicates limited predictive power, likely due to the relatively small dataset. Therefore, it is recommended that future studies utilize a larger seismic dataset and incorporate additional geophysical parameters to improve the model's generalizability and enhance earthquake hazard analysis in Indonesia.

V. References

[1] B. Zhou *et al.*, “Annotated data for semantic role labeling of crisis events in Indonesian Tweets,” *Data Brief*, vol. 61, p. 111688, Aug. 2025, doi: 10.1016/j.compenvurbsys.2022.101824.

[2] M. Akhoondzadeh, “Earthquake prediction using satellite data: Advances and ahead challenges,” *Advances in Space Research*, vol. 74, no. 8, pp. 3539–3555, Oct. 2024, doi: 10.1016/j.asr.2024.06.054.

[3] A. A. Akyol, O. Arikan, and F. Arikan, “A Machine Learning-Based Detection of Earthquake Precursors Using Ionospheric Data,” *Radio Sci.*, vol. 55, no. 11, p. e2019RS006931, Nov. 2020, doi: 10.1029/2019RS006931.

[4] C. Qingping Wang *et al.*, “Earthquake Magnitude Estimation Based on Machine Learning: Application to Earthquake Early Warning System,” *J. Phys. Conf. Ser.*, vol. 1951, no. 1, p. 012057, Jun. 2021, doi: 10.1088/1742-6596/1951/1/012057.

[5] S. Ommi and M. Hashemi, “Machine learning technique in the north zagros earthquake prediction,” *Applied Computing and Geosciences*, vol. 22, no. 1, p. 100163, Jun. 2024, doi: 10.1016/j.acags.2024.100163.

[6] A. Devaux, C. Proust-Lima, and R. Genuer, “Random Forests for Time-Fixed and Time-Dependent Predictors: The DynForest R Package,” *R J.*, vol. 17, no. 1, pp. 24–46, Aug. 2025, doi: 10.32614/rj-2025-002.

[7] A. Leema, P. Balakrishnan, G. G. Kiruba, G. Rajarajan, S. Goel, and P. Aggarwal, “SeismoQuakeGNN: a hybrid framework for spatio-temporal earthquake prediction with transformer-enhanced models,” *Front. Artif. Intell.*, vol. 8, 2025, doi: 10.3389/frai.2025.1690476.

[8] A. Wu, I. Khan, and Y. W. Kwon, “A transformer-based real-time earthquake detection framework in heterogeneous environments,” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-024-82087-2.

[9] Y. Hu, W. Wang, L. Li, and F. Wang, “Applying Machine Learning to Earthquake Engineering: A Scientometric Analysis of World Research,” *Buildings 2024, Vol. 14, Page 1393*, vol. 14, no. 5, p. 1393, May 2024, doi: 10.3390/buildings14051393.

[10] P. Kolivand *et al.*, “A systematic review of Earthquake Early Warning (EEW) systems based on Artificial Intelligence,” *Earth Science Informatics 2024 17:2*, vol. 17, no. 2, pp. 957–984, Feb. 2024, doi: 10.1007/s12145-024-01253-2.

[11] S. Li, Q. Liu, and X. Sun, “HTCNN-Attn: a fine-grained hierarchical multi-label deep learning model for disaster emergency information intelligent extraction from social media,” *PeerJ Comput. Sci.*, vol. 11, p. e2992, Jul. 2025, doi: 10.7717/peerj-cs.2992.

[12] H. Kebede, Z. Demissie, H. Tadesse, and A. Eshetu, “Spatial interpolation techniques comparison and evaluation: The case of ground-based gravity and elevation datasets of the central Main Ethiopian rift,” *Heliyon*, vol. 10, no. 12, Jun. 2024, doi: 10.1016/j.heliyon.2024.e32806.

[13] D. Strnad, D. Mongus, Š. Horvat, and E. Šegina, “A multi-task deep learning approach for landslide displacement prediction with applications in early warning systems,” *Sci. Rep.*, vol. 16, no. 1, Dec. 2026, doi: 10.1038/s41598-025-29084-1.

[14] H. Setya Hadi and E. Gustina, “WEB-BASED INFORMATION SYSTEM FOR THE RECAPITULATION OF THE ELECTION OF THE CHAIRMAN OF RW 011 PADANG SARAI PERMAI HOUSING,” *Jurnal Manajemen Teknologi Informatika*, vol. 2, no. 1, pp. 219–230, Apr. 2024, doi: 10.70038/jentik.v2i1.87.

[15] C. M. Rosca and A. Stancu, “Earthquake Prediction and Alert System Using IoT Infrastructure and Cloud-Based Environmental Data Analysis,” *Applied Sciences 2024, Vol. 14, Page 10169*, vol. 14, no. 22, p. 10169, Nov. 2024, doi: 10.3390/app142210169.

[16] Y. Zhang, M. Kang, X. Li, Y. Qiu, and Z. Li, “Optimizing Smartphone App Usage Prediction: A Click-Through Rate Ranking Approach,” *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 24, pp. 6281–6290, Aug. 2024, doi: 10.1145/3637528.3671567.

[17] R. Jena *et al.*, “Explainable Artificial Intelligence (XAI) Model for Earthquake Spatial Probability Assessment in Arabian Peninsula,” *Remote Sensing 2023, Vol. 15, Page 2248*, vol. 15, no. 9, p. 2248, Apr. 2023, doi: 10.3390/rs15092248.

[18] H. S. Hadi *et al.*, “Internet Of Thing(Iot) : Prinsip Dan Implementasinya”.

[19] N. Farahmandi and A. Patel, “Earthquake Prediction: Irregular Time-series Forecasting with Deep Learning,” Jun. 2025, doi: 10.18260/1-2--56311.

[20] T. Han and G. Xu, “The Relationship Between Learning Environment Perception, Achievement Goals, and the Undergraduate Deep Learning Approach: A Longitudinal Mediation Model,” *J. Intell.*, vol. 13, no. 2, Feb. 2025, doi: 10.3390/jintelligence13020019.

[21] B. Hou, Y. Zhou, S. Li, Y. Wei, and J. Song, “Real-time earthquake magnitude estimation via a deep learning network based on waveform and text mixed modal,” *Earth, Planets and Space*, vol. 76, no. 1, Dec. 2024, doi: 10.1186/s40623-024-02005-8.

[22] Q. He *et al.*, “Evaluation of landslide susceptibility of mountain highway based on RF and SVM models,” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-08774-w.

[23] M. Azizah and H. S. Hadi, “Perancangan Sistem Antrian Dengan Speech Recognition Berbasis Web Pada Puskesmas Gasan Gadang Kab. Padang Pariaman,” *Jurnal Manajemen Teknologi Informatika*, vol. 2, no. 3, pp. 154–160, Dec. 2024, doi: 10.70038/jentik.v2i3.96.