ISSN 2723-6471

A Comparative Study of Machine Learning Approaches to Megathrust Earthquake Prediction in Subduction Zones

Wella[1,*], Ririn Ikana Desanti[2], Suryasari[3]

[1,2,3]*Universitas Multimedia Nusantara, Jl. Scientia Boulevard Gading Serpong, Tangerang, 15810, Indonesia*

(Received: March 07, 2025; Revised: May 19, 2025; Accepted: July 31, 2025; Available online: September 01, 2025)

Abstract

Megathrust earthquakes are one of the most severe threats to countries situated along tectonic subduction zones, particularly Indonesia, where the movement of converging plates frequently triggers large-scale seismic events and tsunamis. Although recent developments in seismology have introduced various predictive tools, many of these models still face challenges, especially due to limitations in hydrogeological data quality. This study aims to investigate how three different machine learning algorithms perform in predicting megathrust earthquake events. The algorithms tested are Support Vector Machine, Random Forest, and Artificial Neural Network, applied to a dataset dominated by earthquake records from the Indonesian and Pacific regions. Each model was evaluated based on accuracy, precision, recall, and F1 score to provide a comprehensive performance analysis. The results show that Random Forest produced the highest accuracy, reaching 96%, followed closely by Support Vector Machine with 95%, while Artificial Neural Network achieved 83%. In terms of the F1 score, Random Forest led with a score of 0.95, indicating balanced performance in classification. However, recall, which is critical in disaster preparedness because it measures the model's ability to detect high-risk events, Artificial Neural Network reached 92% for tsunami-related classifications. This suggests that while Random Forest is the most accurate overall, Artificial Neural Network could be more appropriate for early warning systems where the cost of missing a true event is much higher than issuing a false alarm. The contribution of this research is the direct comparison of multiple machine learning methods using real earthquake data, focusing not only on accuracy but also on practical disaster management considerations such as recall. This study also presents a novel perspective by analyzing the trade-off between model accuracy and disaster risk, emphasizing the need for probabilistic forecasts that can support timely public decision-making during seismic crises.

Keywords: Earthquake Prediction, Machine Learning, Random Forest, Neural Networks, Subduction Zones

1. Introduction

Given its position along the Indian Ocean, particularly in regions such as Aceh, Indonesia is especially susceptible to tsunami hazards [1]. This high level of vulnerability is primarily due to the frequent occurrence of large-magnitude earthquakes generated by megathrust faults. A notable example is the catastrophic earthquake on December 26, 2004, which triggered a massive tsunami, resulting in significant loss of life and extensive infrastructural damage [2], [3]. Situated on the Pacific Ring of Fire, Indonesia faces persistent seismic threats, making the development of robust and reliable earthquake prediction models a critical component of disaster risk reduction. In recent years, Artificial Intelligence (AI) and Machine Learning (ML) have emerged as valuable tools in enhancing the accuracy of seismic forecasting. Despite their growing application in seismological research, several challenges remain [4], [5]. A major issue lies in the insufficient accuracy of hydrogeological data used during model training. Many current models fail to account for key underlying geophysical factors—such as fault behavior, variations in subsurface hydrology, and stress accumulation patterns—focusing instead on seismological and meteorological indicators, which are often secondary outcomes [6], [7]. As a result, these limitations reduce the reliability and interpretability of earthquake predictions, highlighting the need for more comprehensive approaches that incorporate both physical and geological dynamics [8].

For earthquake prediction, many machine learning techniques have been extensively studied. Each method has different advantages and disadvantages. Since they can analyze very large seismic data, models like Support Vector Machines, Random Forests, and Artificial Neural Networks are some of the most widely used methodologies [9], [10], [11]. SVM

***Corresponding author: Wella (wella@umn.ac.id)**
DOI: https://doi.org/10.47738/jads.v6i4.904

is also very good at classifying earthquake patterns and indicating seismic risk zones, but its efficacy is highly dependent on feature selection and kernel optimization [12], [13]. The Random Forest algorithm, despite challenges in capturing extended temporal dependencies, exhibits robust performance in estimating earthquake magnitudes, largely because of its resilience to noisy and incomplete datasets [14]. Likewise, Artificial Neural Networks (ANNs), especially deep learning architecture such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), have proven effective at detecting complex seismic patterns. However, their utility is frequently constrained by the necessity for large volumes of labeled training data and the high computational resources involved [15], [16].

It is difficult to reach a consensus on the most effective megathrust prediction model because the performance of different algorithms is inconsistent [17], [18]. Random Forest produces more accurate magnitude estimates, according to some studies [14]. On the other hand, other studies have found that conventional techniques in nonlinear earthquake dynamics modeling are inferior to deep learning-based ANN models [15], [19]. Due to its robustness in classification tasks, SVM remains a strong choice. However, its effectiveness compared to RF and ANN is still debated [8]. To address this challenge, this research focuses on assessing the effectiveness of Support Vector Machines (SVM), Random Forest (RF), and ANN for predicting megathrust earthquakes. The study seeks to determine the most dependable forecasting model by examining critical performance indicators such as accuracy, precision, recall, and computational cost. The results are expected to improve disaster preparedness and early warning systems, which will help refine mitigation strategies for megathrusts in Indonesia and other seismically active regions around the world.

2. Literature Review

2.1. Previous Research

The table 1 is previous studies presented to support the justification for conducting this research. Recent studies have investigated the use of ML approaches for earthquake forecasting, with ANN and RF outperforming other techniques in accuracy [20]. In one case, seismic data from the Northern Zagros region were analyzed using SVM and ANN, while a hybrid approach combining RF and Multi-Layer Perceptron (MLP) was applied to estimate earthquake magnitudes, yielding encouraging results with a Mean Absolute Error (MAE) of 0.0738 [21], [22]. Attempts to enhance predictive capabilities by integrating various methods revealed certain limitations inherent in hybrid ML models [23].

Table 1. Comparison from Related Research

Ref	Background	Method	Result
[20]	This research focuses on the study of seismicity changes and the potential for large earthquakes in a seismic zone.	This study used a seismic catalog from the Northern Zagros region, which is a seismically active zone with many large cities at high risk of earthquakes.	This research focuses on the study of seismicity changes and the potential for large earthquakes in a seismic zone.
[21]	Western Nepal is highly prone to earthquake-induced landslides due to its active seismic zones and fragile terrain. Assessing landslide susceptibility is vital for hazard mitigation.	The study used 16 causative factors to train and test ANN and Random Forest (RF) models. Model performance was evaluated using AUC metrics.	The Random Forest model achieved higher accuracy (AUC 0.933) compared to the ANN (AUC 0.889), indicating better performance for landslide susceptibility mapping.
[22]	This study explores the use of soft computing techniques, focusing on ANN and the Adaptive Neuro-Fuzzy Inference System (ANFIS), to predict earthquake magnitudes.	Historical seismic data were utilized to train both ANN and ANFIS models for the purpose of predicting earthquake magnitudes. Their predictive capabilities were then assessed and compared to evaluate which model performs more effectively.	Both models showed good predictive ability, but the ANFIS model provided higher accuracy compared to ANN, indicating superiority in handling uncertainty and non-linearity in seismic data.
[23]	This research is centered on building an earthquake forecasting model by leveraging hybrid machine learning	Researchers combined several machine learning algorithms in a hybrid model, including Random	The hybrid model showed improved prediction accuracy compared to a single model, with

Ref	Background	Method	Result
	approaches to enhance both its accuracy and reliability.	Forest, Support Vector Machine, and Neural Networks. The model was trained and tested using seismic datasets to evaluate its performance.	better ability to capture complex patterns in seismic data.
[24]	SVM are commonly applied in classification problems; however, their effectiveness is strongly influenced by the appropriate tuning of hyperparameters.	The authors explore two optimization techniques—Grid Search and Genetic Algorithms (GA)—to fine-tune SVM parameters. They apply these methods to various datasets to evaluate their effectiveness.	The research shows that Genetic Algorithms efficiently explore the hyperparameter space, frequently surpassing Grid Search in both classification accuracy and computational performance.
[25]	Predicting software defects is essential for maintaining software quality. Random Forest (RF) classifiers are commonly used, but their performance can be influenced by hyperparameter settings and data imbalance.	The authors compare different hyperparameter tuning methods for RF classifiers, integrating Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance and employing Genetic Algorithms for feature selection.	The combination of SMOTE and Genetic Algorithm-based feature selection enhances the RF classifier's ability to predict software defects, leading to improved performance metrics.
[26]	Accurate classification of mammogram images is critical for early breast cancer detection. ANN require optimal hyperparameter settings to perform effectively in this domain.	This research presents a meta-heuristic method employing Artificial Bee Colony (ABC) optimization to optimize the hyperparameters of ANN for classifying mammograms.	The ANN model, optimized with the ABC algorithm, achieves superior classification accuracy compared to models tuned with traditional methods.

Additionally, a comparative analysis of Grid Search and Genetic Algorithm (GA) for SVM hyperparameter tuning underscored the importance of optimization, with GA demonstrating improvements in classification accuracy [24]. In software defect prediction, combining SMOTE with GA-based feature selection further improved Random Forest classifier performance, highlighting the value of mitigating class imbalance and ensuring relevant feature selection [25]. Another study employed an ABC optimization meta-heuristic to fine-tune ANN hyperparameters for mammogram classification, resulting in enhanced accuracy [26]. Collectively, these investigations showcase progress in machine learning applications and stress the necessity of careful model selection and hyperparameter tuning across diverse domains including earthquake prediction and healthcare diagnostics.

3. Methodology

The CRISP-ML(Q) framework on figure 1, a systematic methodology for machine learning projects that prioritizes iterative procedures and quality assurance, is depicted in this graphic [27]. The framework is an extension of the ideas of Cross-Industry Standard Process for Data Mining (CRISP-DM) and is specifically designed to help develop and deploy effective machine learning applications. The process consists of three main stages: understanding the business and data, model development, and model operations. Each stage is essential to building a reliable predictive model.

Business and Data Understanding. The first phase involves clarifying the machine learning use case, aligning with organizational objectives, and assessing the data at hand. In the context of earthquake prediction studies, this step is vital for pinpointing the main factors affecting seismic events and choosing suitable machine learning algorithms such as ANNs, SVMs, and RF. By thoroughly examining the seismic and hydrogeological features within the dataset, researchers can guarantee that the models are trained on data that is both relevant and meaningful.

Model Development. This stage encompasses the technical processes involved in building the predictive model after gaining a comprehensive understanding of the data. Key activities include data preprocessing, feature engineering, model training, and evaluation. In earthquake prediction research, this phase is particularly important due to the large volumes of seismic data that must be cleaned of noise and filtered for the most informative features. Machine learning algorithms such as ANNs, SVMs, and RF are trained and validated during this process. Their effectiveness is assessed using metrics including F1 score, accuracy, precision, and recall. Given its iterative nature, this phase allows for

 ISSN 2723-6471

ongoing refinement, enabling the model to more accurately detect seismic patterns and enhance its prediction performance.

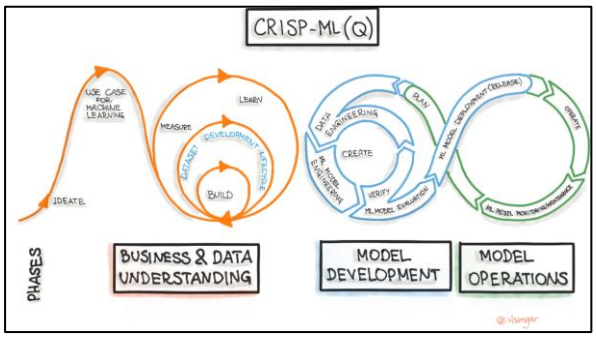

Figure 1. CRISP-ML Life Cycle Process [27]

Model Operations. The final phase entails the implementation, deployment, and monitoring of the machine learning model as it is used in the real world. This process is critical to integrating the trained model into the early warning system, which will allow for real-time assessment of seismic data and prompt issuance of alarms. Repeated monitoring of the model's performance ensures that it remains accurate, allowing for improvements as new earthquake data is acquired. This continuous improvement builds confidence in the model, making it an essential tool for disaster preparedness and risk mitigation.

4. Results and Discussion

4.1. Business & Data Understanding

The dataset applied to this research was retrieved from: https://www.kaggle.com/code/ravivarmaodugu/earthquakes-data-analytics/input and it combines two datasets: earthquake_1995-2023 dan earthquake_data.csv [28]. Figure 2 presents several key parameters used to describe earthquake events. The first parameter is the Community Determined Intensity, or CDI, which represents the perceived intensity of an earthquake as reported by people in the affected community. This value ranges from 1 to 12, reflecting increasing levels of perceived shaking. The second parameter is the Modified Mercalli Intensity, or MMI, which assesses earthquake intensity based on observable physical impacts on individuals, buildings, and the surrounding environment. Like CDI, MMI also ranges from 1 to 12.

	title	magnitude	date_time	cdi	mmi	alert	tsunami	sig	net	nst	dmin	gap	magType	depth	latitude	longitude	location	continent
0	M 6.5 - 42 km W of Sola, Vanuatu	6.5	16-08-2023 12:47	7	4	green	0	657	us	114	7.177000	25.0	mww	192.955	-13.8814	167.1580	Sola, Vanuatu	NaN
1	M 6.5 - 43 km S of Intipucá, El Salvador	6.5	19-07-2023 00:22	8	6	yellow	0	775	us	92	0.679000	40.0	mww	69.727	12.8140	-88.1265	Intipucá, El Salvador	NaN
2	M 6.6 - 25 km ESE of Loncopué, Argentina	6.6	17-07-2023 03:05	7	5	green	0	899	us	70	1.634000	28.0	mww	171.371	-38.1911	-70.3731	Loncopué, Argentina	South America
3	M 7.2 - 98 km S of Sand Point, Alaska	7.2	16-07-2023 06:48	6	6	green	1	860	us	173	0.907000	36.0	mww	32.571	54.3844	-160.6990	Sand Point, Alaska	NaN
4	M 7.3 - Alaska Peninsula	7.3	16-07-2023 06:48	0	5	NaN	1	820	at	79	0.879451	172.8	Mi	21.000	54.4900	-160.7960	Alaska Peninsula	NaN

Figure 2. The Dataset [28]

Another important variable is the Significance Index, abbreviated as Sig. This index provides a quantitative measure of the earthquake's importance by incorporating factors such as magnitude, depth, and geographic location. A higher significance index indicates a greater potential impact of the earthquake. The Number of Stations, referred to as Nst, indicates how many seismic stations recorded the event. A larger number of stations generally results in more accurate calculations of the earthquake's location and magnitude.

The Minimum Distance, or Dmin, measures the shortest distance between the earthquake's epicenter and the nearest seismic monitoring station. This distance is recorded in degrees. The Azimuthal Gap, known as Gap, describes the

 ISSN 2723-6471

largest angular gap between stations around the earthquake epicenter, also measured in degrees. Smaller gap values indicate better coverage and higher precision in determining the earthquake's location.

The Magnitude Type, labeled as MagType, refers to the specific method or scale used to calculate the earthquake's magnitude. Finally, Depth describes how far below the Earth's surface the earthquake occurred, measured in kilometers. Earthquakes are categorized by depth into three groups: shallow earthquakes occur at depths less than 70 kilometers, intermediate-depth earthquakes range from 70 to 300 kilometers, and deep earthquakes take place at depths exceeding 300 kilometers. The countries where earthquakes have happened, as noted in the dataset as depict in figure 3.

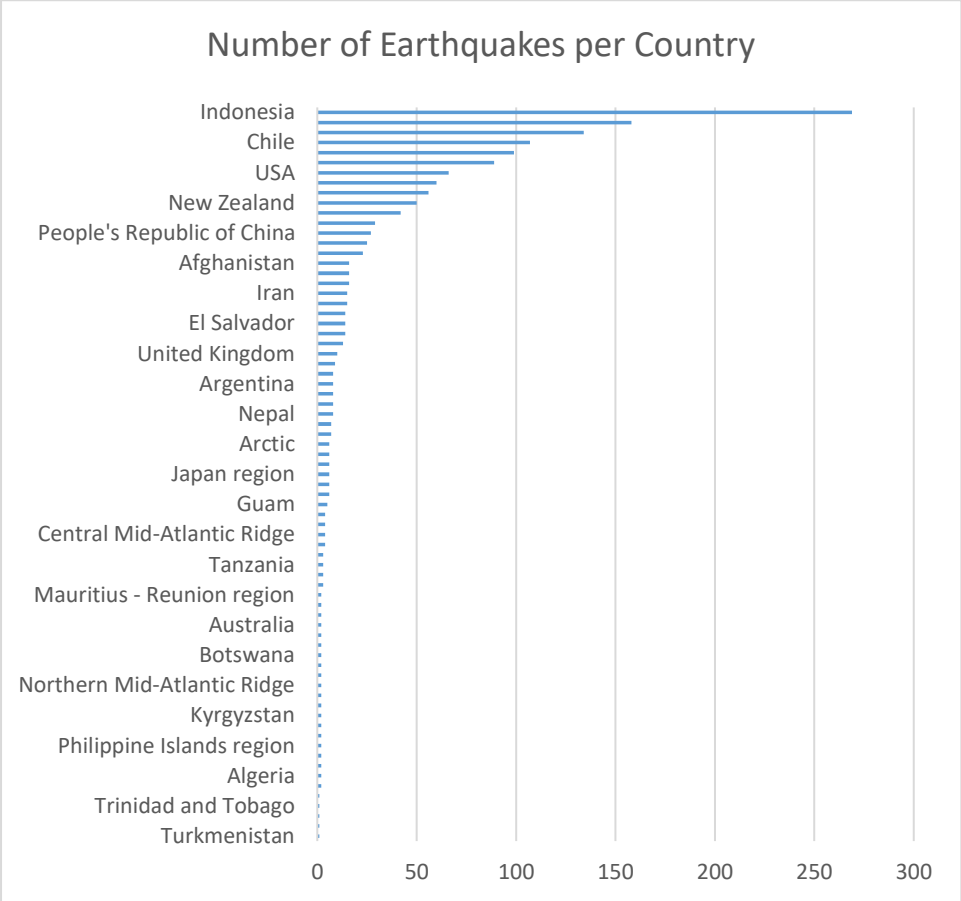

Figure 3. Detailed dataset of country [28]

With a total dataset of 1,776 earthquakes in 1995-2023, Indonesia is the country with the highest frequency of earthquake occurrences. Figure 4 is a world map that illustrates the geographical spread of earthquake-affected areas across the globe. The blue markers indicate the epicenters, which are predominantly clustered along the circum-Pacific belt, the Himalayan region, and parts of the Mediterranean-Asian seismic belt. This visualization confirms the global reach of seismic activity but also emphasizes the disproportionate clustering in tectonically active zones. The Pacific Ring of Fire, in particular, is visibly the most affected, consistent with the prior country-level analysis. The spatial distribution provides critical insights into regional vulnerability and is essential for global disaster risk reduction frameworks.

Figure 5 is a top 5 earthquake-prone countries. The bar chart presents the top five countries with the highest number of recorded earthquakes. Indonesia leads with 269 events, followed by Papua New Guinea (158), Japan (134), Chile (107), and Vanuatu (99). This distribution highlights the concentration of seismic activity within the Pacific Ring of Fire, where tectonic plate boundaries are highly active. These countries are situated along subduction zones that are known to generate frequent and potentially damaging seismic events. The chart supports the need for region-specific mitigation strategies and continuous seismic monitoring in these high-risk areas.

ISSN 2723-6471

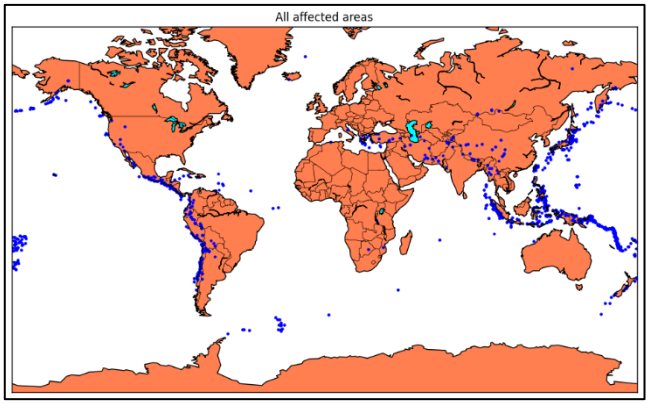

Figure 4. Distribution of earthquakes that occurred in 1995-2023

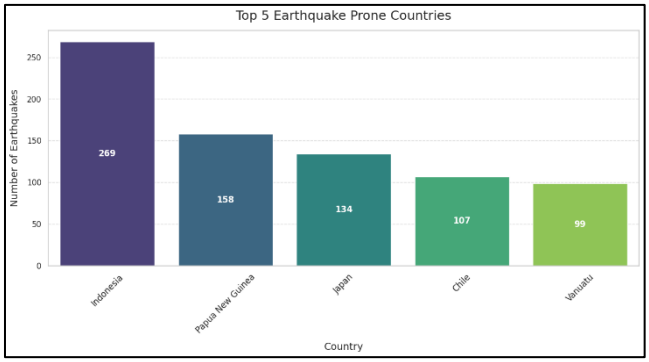

Figure 5. Top 5 Earthquake-Prone Countries

Magnitude distributions refer to the pattern of distribution or frequency of earthquake occurrences based on their size (magnitude) which can be seen in Figure 6. Based on figure 6, the frequency of earthquakes based on their magnitude per event. The majority of recorded earthquakes fall within the 6.0 to 6.3 magnitude range, with a peak at 6.0 (296 events). As magnitude increases, the frequency of events decreases substantially, with only 50 events recorded at magnitude 7.5. This inverse relationship between magnitude and frequency is consistent with the Gutenberg-Richter law, which postulates that larger earthquakes occur less frequently. Understanding this distribution is important for probabilistic seismic hazard assessment and infrastructure design codes.

It can be seen in figure 7, the time-series line graph illustrates the annual number of earthquakes from 1998 to 2023. A general upward trend is observed from the early 2000s, peaking in 2010 and 2013 with over 100 events each. However, a noticeable decline is apparent after 2016, reaching its lowest point in 2023. The temporal variation may be influenced by changes in monitoring capabilities, tectonic cycles, or reporting thresholds. Longitudinal analysis of this nature is valuable for evaluating temporal seismicity patterns and forecasting potential future activity.

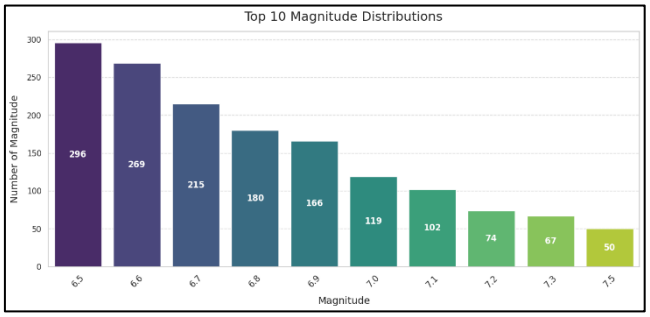

Figure 6. Magnitude distributions

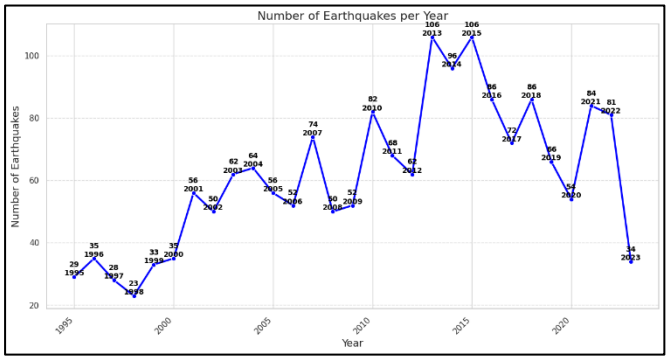

Figure 7. Quantity of earthquakes from 1995-2023

Community Determined Intensity (CDI) refers to earthquake intensity determined by observation of the community, as seen in figure 8. It is a measure of earthquake effect obtained from direct experience and observation of individuals residing in a particular locality. The CDI distribution chart reflects the community-perceived shaking intensities reported by the population. A large number of events are associated with a CDI of 0 (609 reports), indicating either unperceived events or remote epicenters. Nevertheless, considerable numbers are also seen at mid-level intensities, particularly at CDI levels 5 (220), 6 (156), and 8 (179). This reflects the variability in felt intensity despite similar magnitudes, which is influenced by depth, location, and local site conditions. These data are crucial for calibrating intensity prediction models and public preparedness planning.

Figure 9 depicts the MMI scale. The MMI distribution illustrates the physical shaking intensity as estimated from seismic data and impact reports. The most frequent values are MMI 6 and 7, with 486 and 480 counts respectively, indicating moderate to strong shaking. Lower intensities such as MMI 2 and 3 are rare, and very high intensities (MMI

 ISSN 2723-6471

9–10) are uncommon but still present. The distribution underscores the relatively high severity of ground motion experienced during these events. This metric is vital for impact-based forecasting and post-event damage assessment.

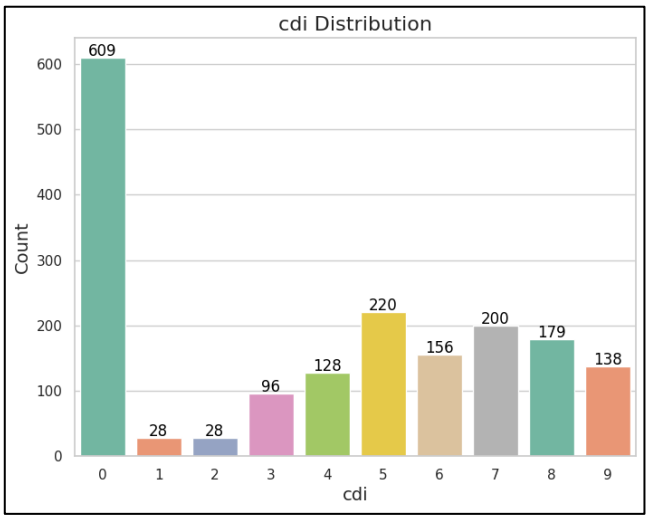

Figure 8. CDI (Community Determined Intensity)

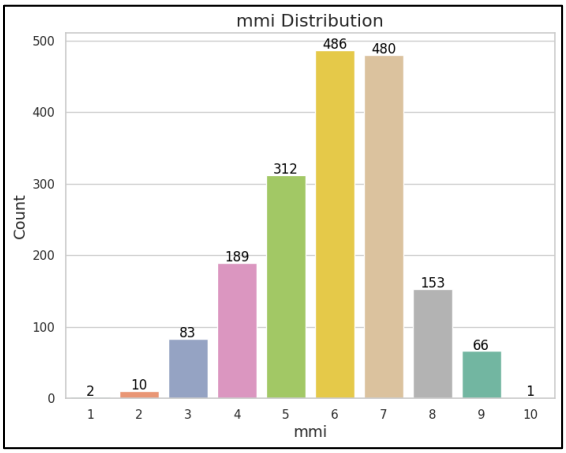

Figure 9. Modified Mercalli Intensity (MMI)

The tsunami distribution bar chart (figure 10) indicates that among the total earthquake events, 629 were associated with a tsunami, while 1,153 were not. This binary classification highlights that while the majority of earthquakes do not generate tsunamis, a substantial portion still poses significant hydrodynamic hazards. The occurrence of tsunamis is often linked to undersea megathrust events, reinforcing the importance of early warning systems in coastal regions. Quantifying this risk is critical for integrated coastal hazard management.

Based on figure 10, it can be analyzed that 1,153 earthquake events occurred in non-oceanic areas, making tsunami occurrences unlikely. Meanwhile, 629 earthquake events occurred in oceanic areas, where tsunamis are possible (0 = No, 1 = Yes). Figure 11 is monthly distribution of earthquake events is relatively uniform across the year, with slight peaks observed in October (174 events) and November (209 events). This suggests a weak or non-existent seasonality in global earthquake occurrences, as expected from tectonic processes which are not driven by climatic or seasonal variations. However, subtle clustering in certain months may warrant further investigation into possible triggering factors. These insights are useful for temporal risk communication and readiness efforts.

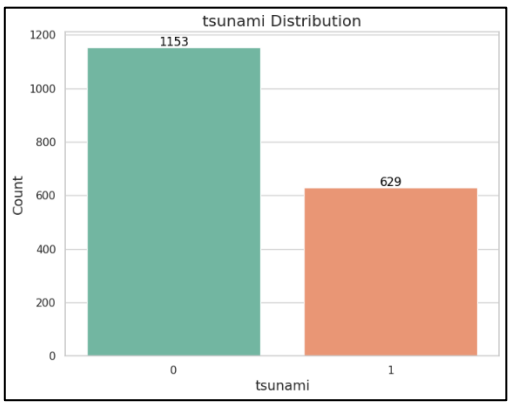

Figure 10. Tsunami Distribution

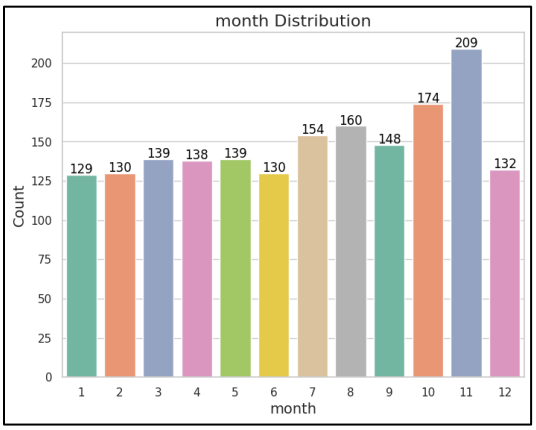

Figure 11. Month Distribution

4.2. Model Development

Through the correlation heatmap on figure 12, it is possible to see several strong correlations between variables within the earthquake dataset. One of the strongest relationships is that between magnitude and earthquake importance (sig), which demonstrates that higher-magnitude earthquakes have a higher level of significance. Moreover, there is a positive correlation between earthquake magnitude and tsunami occurrence, suggesting that earthquakes of larger magnitudes

 ISSN 2723-6471

have a higher likelihood of generating tsunamis. However, there is a negative relationship between magnitude and earthquake depth, meaning that stronger earthquakes occur at more shallow depths.

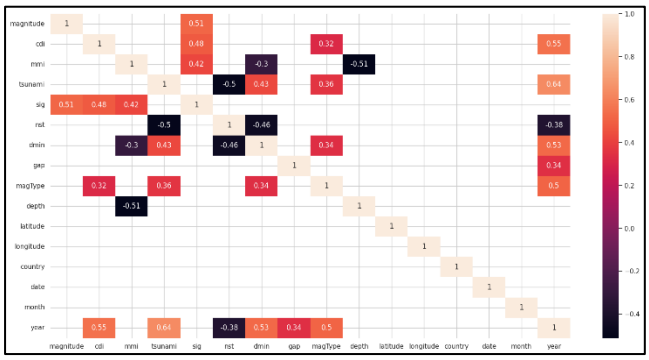

Figure 12. Correlation heatmap

Along with magnitude-based correlations, the heatmap also shows temporal trends. The year variable is positively correlated with earthquake intensity values (CDI and MMI), meaning that earthquakes have had a larger impact than before in recent years. This could be because geological changes or improvements in earthquake recording and reporting systems have occurred (see figure 13). AI techniques have demonstrated great effectiveness in uncovering intricate patterns and complex connections within historical seismic datasets [29]. For each model, a Random Search Optimization with 20 iterations and k-fold=5 is done to get the best hyperparameter configuration.

Since the dataset contains a mixture of both tsunami and non-tsunami occurrences, a linear kernel is unsuitable for this classification task. When implemented, the linear kernel yielded a maximum attainable overall accuracy of only 60%, even for Class 1 (tsunami) predictions (see figure 14). Therefore, a non-linear kernel was necessary to capture the complex, non-linear relationships within the data. The primary objective of an SVM is to identify the optimal hyperplane that maximally separates data points from different classes within a high-dimensional feature space, enabling effective classification of new, unseen data [20].

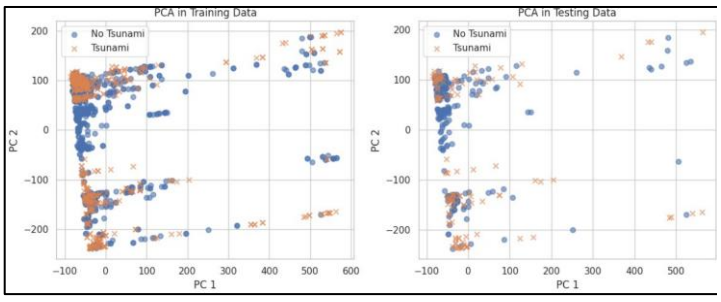

Figure 13. Tsunami Classification using SVM

```
svm - Cross-validation score: 0.8701754385964913
Test score: 0.9467787114845938
Best model for svm: Pipeline(steps=[('smote', SMOTE(random_state=11)), ('sc
aler', MinMaxScaler()),
                ['classifier',
                 SVC(C=7250.01, gamma=90.01, probability=True,
                     random_state=11)]])
Best parameters for svm: {'classifier__kernel': 'rbf', 'classifier__gamma':
90.01, 'classifier__C': 7250.01}

                    precision    recall  f1-score   support

                 0       0.94      0.98      0.96       231
                 1       0.96      0.89      0.92       126

          accuracy                           0.95       357
         macro avg       0.95      0.93      0.94       357
      weighted avg       0.95      0.95      0.95       357
```

Figure 14. SVM Result

The hyperparameter range for the Random Search referred to the previous research [24], and the details can be seen in table 2. The optimized hyperparameters obtained from the tuning process are as follows: the gamma parameter was set to 90.01, while the regularization parameter C was determined to be 7250.01. According to this result-high recall for tsunami prediction means that the model can capture true occurrences of tsunamis on one end, whereas on the other end, it would rather make evacuation preparations than stay undefined. The analysis highlights two key points regarding the model's performance. First, for Class 0 (non-tsunami events), the balance between precision and recall suggests that some negative instances were incorrectly classified as positive (i.e., as tsunamis). This pattern indicates a conservative model behavior, prioritizing the minimization of false negatives. In other words, the model tends to avoid missing actual tsunami events, even if it means accepting a higher rate of false positives for non-tsunami cases. Second, for Class 1 (tsunami events), the precision exceeding recall implies that the model is highly accurate when it does predict a tsunami, but it occasionally fails to detect some actual tsunami events. This suggests that while the model's

 ISSN 2723-6471

positive predictions are generally reliable, it may overlook certain tsunami incidents, particularly in ambiguous cases where the signal is weak or similar to shipwreck scenarios.

Table 2. SVM Random Search Hyperparameter Range

Parameters	Kernel	Min	Max	Type	Steps	Scale
C	Linear	0.001	10,000	Real	10	logarithmic or logarithmic legacy
gamma	Linear, RBF, sigmoid	0.001	10,000	Real	10	logarithmic or logarithmic legacy
degree	Polynomial	1	5	Integer	1	Linear (1,2,3,4,5)

Class 0 has better classification performance concerning Class 1 because oversampling by SMOTE could have taken place. However, both classes do procure quite a good accuracy and F1 score, meaning the model performance is balanced. This performance of SVM model is superb by getting very high accuracy even in training (87%) and testing (94%) data. Overfitting is not occurring, because the test accuracy is not much different from the cross-validation score; meaning, it generalizes well to previously unseen data.

The hyperparameter range for the Random Search procedure was determined based on the previous study conducted by Suryadi et al. [25]. The outcomes of the Random Search optimization are presented in figure 15, which shows that the optimal configuration includes an n_estimator value of 70, min_sample_split set to 3, min_sample_leaf set to 1, max_depth left undefined (none), and bootstrap set to false. The same figure also summarizes the model's training performance, which demonstrates results consistent with the SVM findings, particularly in terms of the model's behavior for Class 0 and Class 1. For Class 0 (non-tsunami events), the precision is lower than the recall. This indicates that the model adopts a precautionary approach, favoring the reduction of false negatives and minimizing the likelihood of overlooking actual tsunami events.

For Class 1 (tsunami events), the precision is notably higher than the recall. This suggests that the model is accurate when it predicts a tsunami occurrence; however, it tends to detect fewer tsunami events overall, indicating a conservative detection threshold that prioritizes correctness over completeness. Random Forest model is doing quite exceptionally high, with accuracy-89.24% being for training data and 95% for test data. No overfitting could be said, as accuracy obtained in a test set is directly comparable to the cross-validation score, indicating that it can generalize quite well. The result of both ROC Curve and ROC AUC Score (figure 16) being 1 indicates a perfect performance in predicting megathrusts. The score means that the model can distinguish between all classes perfectly, without making any false negative or false positive predictions.

```
rf - Cross-validation score: 0.9024561403508772
Test score: 0.9551820728291317
Best model for rf: Pipeline(steps=[('smote', SMOTE(random_state=11)), ('sca
ler', MinMaxScaler()),
                ['classifier',
                 RandomForestClassifier(bootstrap=False, min_samples_split=
3,
                                        n_estimators=70, random_state=1
1)]])
Best parameters for rf: {'classifier__n_estimators': 70, 'classifier__min_s
amples_split': 3, 'classifier__min_samples_leaf': 1, 'classifier__max_dept
h': None, 'classifier__bootstrap': False}

              precision    recall  f1-score   support

           0       0.96      0.97      0.97       231
           1       0.94      0.93      0.94       126

    accuracy                           0.96       357
   macro avg       0.95      0.95      0.95       357
weighted avg       0.96      0.96      0.96       357
```

Figure 15. RF Result

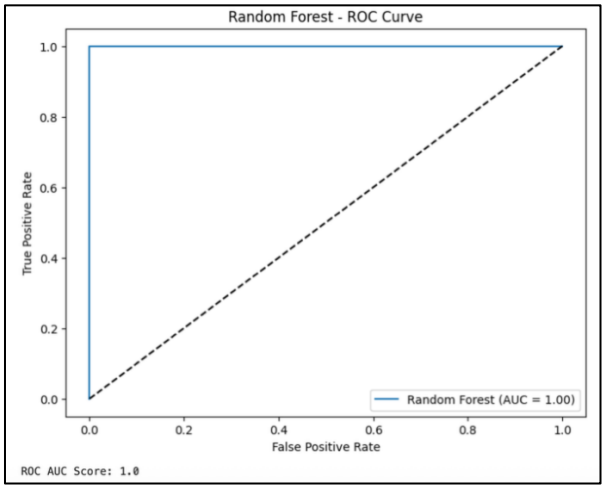

Figure 16. RF ROC AUC Result

Figure 17 shows the outcomes obtained from both Random Search Optimization and the model training process. The hyperparameter range used in the Random Search Optimization for ANN refers to those used in the study by Mamindla & Ramadevi [26], as described in table 3. The hyperparameter tuning process resulted in the following configuration

 ISSN 2723-6471

for the model. The learning rate was set to 0.0001, indicating a cautious update during the optimization process. The model architecture employs three hidden layers with respective units of 64, 32, and 16, forming a progressively narrowing structure that facilitates feature abstraction. No dropout was applied, as the dropout rate was set to 0.0, suggesting that regularization through node deactivation was not utilized in this case. The Rectified Linear Unit (ReLU) function was selected as the activation function to introduce non-linearity into the model. Training was conducted over 200 epochs with a batch size of 128, balancing computational efficiency with gradient stability during learning.

```
Best Parameters: {'classifier__optimizer__learning_rate': 0.0001, 'classifi
er__model__hidden_layers': [64, 32, 16], 'classifier__model__dropout': 0.0,
'classifier__model__activation_function': 'relu', 'classifier__fit__epoch
s': 200, 'classifier__fit__batch_size': 128}
              precision    recall  f1-score   support

           0       0.89      0.85      0.87       231
           1       0.74      0.81      0.78       126

    accuracy                           0.83       357
   macro avg       0.82      0.83      0.82       357
weighted avg       0.84      0.83      0.84       357
```

Figure 17. ANN Result

Table 3. ANN Random Search Hyperparameter Range

Parameters	Range
Number of Hidden Layers	1 to 3
Number of Hidden Nodes	1 to 10
Number of Training Cycles	10 to 1,000
Learning Rate	0.0001 to 0.1
Learning Algorithm	RMDprop, SDG, Adam
Adam, SDG, RMDprop	Linear, Tangent
Learning Rate Decay	Linear, Exponential
Error Function	Mean Square Error, Log Loss
Epoch Limit	Maximum number of learning iterations
Mini Batch Size	10, 20, 30
Patience	2, 5, 10

Compared to SVM and RF, the ANN model shows higher precision than recall in Class 0, i.e., it is more conservative when it is predicting the non-occurrence of a tsunami. Thus, it is more likely to misclassify non-tsunami events as tsunami occurrences. Conversely, in Class 1, the model demonstrates recall lower than precision, i.e., the high recall ensures only a few true tsunami events are left out. This shows that the ANN model is more capable of identifying true tsunami events. A key limitation of this study lies in the absence of hyperparameter tuning and architectural adjustments for the ANN model. This decision was made intentionally to ensure a fair, baseline-level comparison among the three models SVM, Random Forest, and ANN using default or minimal configurations. As a result, the relatively lower accuracy of the ANN model (83%) may be partially attributed to the lack of optimization, which could have otherwise enhanced its performance. Furthermore, the conclusion that ANN is more suitable for tsunami prediction due to its higher recall in Class 1 is based solely on a single confusion matrix. No cross-validation, statistical testing, or confidence interval analysis was conducted to verify the stability or generalizability of this result. Future studies should incorporate systematic hyperparameter tuning, deeper network exploration, and robust evaluation techniques to better assess the full potential of ANN in this context.

Table 4 compares the accuracy of three machine learning models: SVM, RF, and ANN. The findings reveal that Random Forest attained the highest accuracy of 96%, with SVM close behind at 95%, whereas ANN showed the lowest accuracy at 83%. This suggests that, within the context of the data set used, Random Forest and SVM provide superior performance in classifying earthquake-related data compared to ANN.

 ISSN 2723-6471

Table 4. Comparison of Accuracy

	SVM	RF	ANN
Accuracy	0.95	0.96	0.83

Meanwhile of the three models, Random Forest achieves the highest F1-score, outperforming the other models (RF > SVM > ANN). The lower effectiveness of the neural network model stems from suboptimal parameter initialization, which demands more extensive experimentation than conventional machine learning methods.

But according to the principle that "for tsunami prediction, a high recall is more desirable because it indicates how well the model can capture actual tsunami events (TP). It is better to take evacuation procedures than to not act at all," the ANN model is more desirable. This is because, in Class 1 classification, the ANN model has recall > precision, whereas the other models have precision > recall. Each forecast published does risk inducing unnecessary fear or a false sense of security. This implies that there must be openness as an underlying principle. Each probabilistic prediction must be accompanied by its own uncertainties to enable the general public to be adequately informed in their decisions [30].

4.3. Model Operations

Prediction of the Megathrust using Best Algorithm (Random Forest). Megathrust earthquakes occur in subduction zones [31]. From a geological perspective, a subduction zone is an area where a denser tectonic plate sinks beneath a lighter one. Subduction zones are typically characterized by deep ocean trenches, high seismicity, and extensive crustal deformation. Subduction zones are also often the location of great earthquakes and tsunamis caused by the unloading of energy along the contact boundary between the two plates. Subduction zones exist along the Cascadia Subduction Zone off North America, the Sunda Subduction Zone between Sumatra and Java, and the Japan Subduction Zone, comprising the Nankai Trough and the Japan Trench. Another notable example is the South American Subduction Zone, where the Nazca Plate is descending beneath the South American Plate. Megathrust earthquakes are prevalent in these regions, generating strong seismic activity with the potential to generate tsunamis that can cause widespread effects.

Facts from SMS Tsunami Warning about the largest magnitude earthquake [32]. Several of the most devastating earthquakes in recorded history have occurred along subduction zones, where one tectonic plate is forced beneath another. One such event is the Valdivia earthquake in Chile, which struck on May 22, 1960, with a magnitude of 9.5. This megathrust earthquake occurred along the boundary where the Nazca Plate subducts beneath the South American Plate. The disaster resulted in the deaths of approximately 1,655 people, injured around 3,000 others, and displaced nearly two million individuals. The economic losses in Chile were estimated at 550 million US dollars. The earthquake also triggered a massive tsunami that caused significant damage in distant locations, including Hawaii, Japan, and the Philippines, leading to further casualties and destruction. The rupture zone of the earthquake extended over 1,000 kilometers, making it the largest ever recorded. Moreover, on May 24, 1960, just two days after the quake, the Puyehue volcano in Chile erupted, releasing ash and steam that reached altitudes of up to 6 kilometers and continued for several weeks.

Another significant event was the Prince William Sound earthquake in Alaska, which occurred on March 28, 1964, with a magnitude of 9.2. This megathrust earthquake took place in the subduction zone where the Pacific Plate descends beneath the North American Plate. Although the overall destruction was less severe than in the 1960 Chile event, the earthquake generated a tsunami that led to 128 fatalities and caused an estimated 311 million US dollars in financial losses. The areas most affected included Alaska and parts of Canada, but the tsunami also reached Hawaii, resulting in additional damage. The city of Anchorage, located about 120 kilometers northwest of the epicenter, experienced the most intense impacts. Notably, seismic shaking from the event lasted for nearly three minutes, making it one of the longest-duration earthquakes ever recorded.

A more recent catastrophe was the Sumatra earthquake in Indonesia on December 26, 2004, which had a magnitude of 9.1. This event occurred along the subduction zone where the Indo-Australian Plate is forced beneath the Eurasian Plate. The earthquake triggered the Boxing Day Tsunami, which resulted in catastrophic human and material losses. Approximately 227,900 people were reported dead or missing, and about 1.7 million individuals were displaced across

14 countries in South Asia and East Africa. The earthquake's epicenter was located about 250 kilometers southeast of Banda Aceh, Indonesia, at a depth of 30 kilometers. In the aftermath, on December 28, 2004, a mud volcano near Baratang in the Andaman Islands erupted, an event likely linked to the massive tectonic activity generated by the earthquake.

All three events are classified as megathrust earthquakes because they happened in subduction zones, where one tectonic plate slides beneath another. This means they all exhibit the thrust faulting mechanism typical of megathrust earthquakes. The model correctly predicted the three as megathrust, as shown in figure 18 and figure 19, reflecting its ability to correctly predict megathrust in real-world scenario.

To test its reliability in predicting true negative events, the data of three random earthquakes that occurred in non-subduction zones were used, as shown in figure 19. The model correctly classified all of them as non-megathrust earthquakes, demonstrating its ability in predicting true negative accurately.

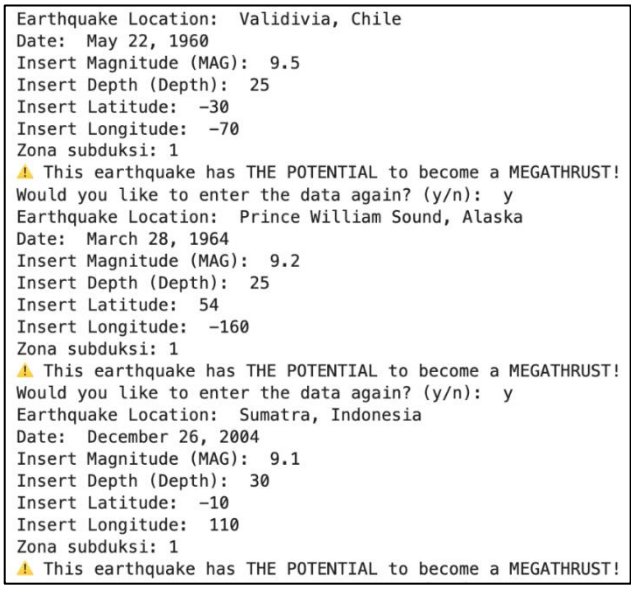

Figure 18. True Positive Result of Random Forest Algorithm

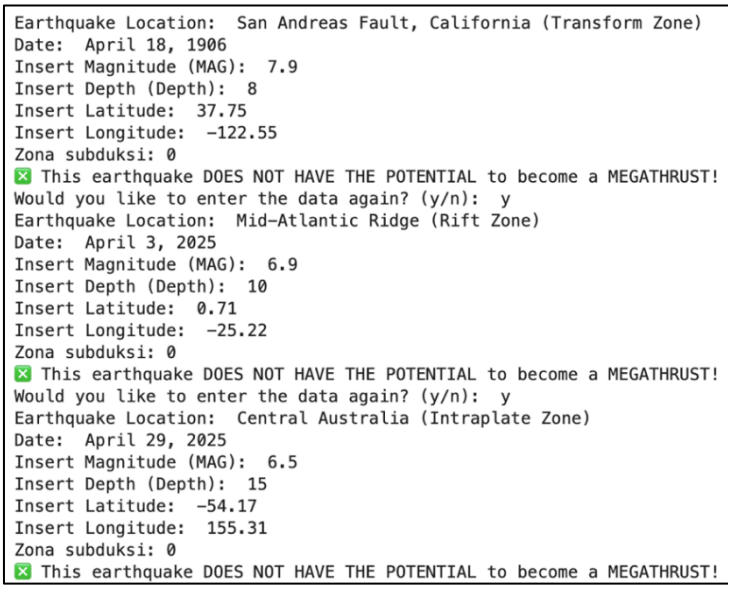

Figure 19. True Negative Result of Random Forest Algorithm

4.4. Discussion

This study compared the predictive capabilities of SVM, RF, and ANN models for megathrust earthquake forecasting. The results showed that RF achieved the highest accuracy at 96%, followed closely by SVM at 95%, while ANN had the lowest accuracy at 83%. These outcomes are consistent with previous research conducted in the North Zagros region, which highlighted that careful feature selection enhances model accuracy. However, unlike the current study, that research did not perform a direct comparison of multiple algorithms to identify the most effective model [20]. One important limitation of the current study is the lack of analysis on feature importance generated by the RF model. Since RF inherently provides model interpretability, discussing the most influential features could have offered deeper insights into the key factors driving earthquake predictions.

Conversely, an ensemble learning approach combining RF and MLP was applied, showing that a combined model could improve accuracy compared to individual approaches, as used in this study. However, while lower in accuracy, ANN was found to be more desirable in tsunami prediction due to its higher recall, an aspect not emphasized in the ensemble learning study [21].

Soft computing techniques employing the ANFIS were also implemented, demonstrating that combining ANN with fuzzy-based techniques could improve accuracy. This contrasts with the present study, where ANN performed the worst. A hybrid ML approach was also proposed, which was found to yield more accurate predictions than single methods, differing from this research, which focused on individual comparisons between SVM, RF, and ANN [22], [23].

Optimizing hyperparameters is essential for boosting model performance, as evidenced by a study where tuning SVM parameters with Grid Search and Genetic Algorithms led to improved classification accuracy. Another study found that RF performance could be improved through hyperparameter tuning using SMOTE and feature selection based on Genetic Algorithms. However, hyperparameter tuning was not a primary focus in this study, suggesting that ANN could perform better with more intensive optimization, as observed in research utilizing meta-heuristic techniques to enhance ANN performance [24], [25], [26].

Although the current results affirm that RF outperforms other models in terms of accuracy, and ANN remains relevant in recall-sensitive scenarios such as tsunami prediction, a key limitation of this study lies in the dataset composition. The earthquake data used in this research are predominantly concentrated in the Indonesian and Pacific regions, which may introduce geographic bias and limit the model's generalizability to other tectonic contexts. Since no stratification or bias correction was applied, future work should consider incorporating geographically diverse data and applying correction techniques to ensure broader applicability of the models. Additionally, the binary classification scheme used in this study—distinguishing only between tsunami and non-tsunami events—as a proxy for identifying megathrust earthquakes, oversimplifies the geological complexity of such phenomena. This approach may overlook important variations in fault mechanisms, rupture characteristics, and subduction dynamics that do not always correlate directly with tsunami occurrence. Exploring hybrid methods and more extensive parameter optimization may yield even more reliable and generalizable earthquake prediction models.

5. Conclusion

Out of the three algorithms tested: SVM, RF, and ANN, the Random Forest model stood out as the most effective at predicting megathrust earthquakes. It successfully identified several major historic quakes, like those in Valdivia, Chile; Prince William Sound, Alaska; and Sumatra, Indonesia, as having megathrust potential. On the other hand, earthquakes from other fault types, such as transform faults like the San Andreas Fault, rift faults like the Mid-Atlantic Ridge, and intraplate faults in Central Australia, were correctly recognized as unlikely to be megathrusts. The model also revealed a clear pattern: megathrust earthquakes tend to happen in subduction zones, have magnitudes above 7.5, and occur at depths shallower than 60 km under the ocean floor. Additionally, it can tell whether an earthquake is within a subduction zone, which helps improve its accuracy in spotting potential megathrust events.

6. Declarations

6.1. Author Contributions

Conceptualization: W., R.I.D., S.; Methodology: S.; Software: W.; Validation: W., S., and R.I.D.; Formal Analysis: W., S., and R.I.D.; Investigation: W.; Resources: S.; Data Curation: S.; Writing Original Draft Preparation: W., S., and R.I.D.; Writing Review and Editing: S., W., and R.I.D.; Visualization: W.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

 ISSN 2723-6471

References

[1] S. Sufri and J. A. Lassa, "Integration of disaster risk reduction and climate change adaptation in Aceh: Progress and challenges after 20 years of Indian Ocean tsunamis," *Int. J. Disaster Risk Reduct.*, vol. 113, no. 1, pp. 1–12, Oct. 2024, doi: 10.1016/j.ijdrr.2024.104894.

[2] P. A. Kohl, A. P. O'Rourke, D. L. Schmidman, W. A. Dopkin, and M. L. Birnbaum, "The Sumatra-Andaman earthquake and tsunami of 2004: The hazards, events, and damage," *Sci. Tsunami Hazards*, vol. 20, no. 6, pp. 355–363, Jun. 2012, doi: 10.1017/S1049023X00002880.

[3] M. Munirwansyah, R. P. Munirwan, V. Listia, I. Irhami, and R. P. Jaya, "Sumatra-fault earthquake source variation for analysis of liquefaction in Aceh, Northern Indonesia," *Open Civ. Eng. J.*, vol. 17, no. 1, pp. 1–12, Nov. 2023, doi: 10.2174/0118741495270939230921154841.

[4] A. Chitkeshwar, "The role of machine learning in earthquake seismology: A review," *Arch. Comput. Methods Eng.*, vol. 1, no. 1, pp. 1–12, Mar. 2024, doi: 10.1007/s11831-024-10099-2.

[5] R. Zhu, X. Wang, Y. Li, Y. Chen, M. Liu, Z. Wu, and L. Zhang, "Anomaly detection using machine learning in hydrochemical data from hot springs: Implications for earthquake prediction," *Wiley Online Library*, vol. 60, no. 6, pp. 1–12, Jun. 2024, doi: 10.1029/2023WR034748.

[6] J. Li, H. Wang, Y. Zhang, Q. Liu, M. Chen, Y. Xu, and X. Zhao, "Application of artificial intelligence technology in the study of anthropogenic earthquakes: A review," *Artif. Intell. Rev.*, vol. 58, no. 5, pp. 1–12, Mar. 2025, doi: 10.1007/s10462-025-11157-2.

[7] Z. Ying, W. Congcong, S. Didier, and Z. Chengxiang, "Integrating artificial intelligence and geophysical insights for earthquake forecasting: A cross-disciplinary review," *arXiv*, vol. 1, no. 1, pp. 1–12, Feb. 2025.

[8] O. N. Akarsu, "A bibliometric review of earthquake and machine learning research," *Civ. Eng. Beyond Limits*, vol. 5, no. 1, pp. 1–10, Apr. 2024, doi: 10.36937/cebel.2024.1908.

[9] N. Alidadi and S. Pezeshk, "State of the art: Application of machine learning in ground motion modeling," *Eng. Appl. Artif. Intell.*, vol. 149, no. 1, pp. 1–12, Jun. 2025, doi: 10.1016/j.engappai.2025.110534.

[10] M. S. Chowdhury, "Comparison of accuracy and reliability of random forest, support vector machine, artificial neural network and maximum likelihood method in land use/cover classification of urban setting," *Environ. Challenges*, vol. 14, no. 1, pp. 1–12, Jan. 2024, doi: 10.1016/j.envc.2023.100800.

[11] R. Jena, I. F. Arman, K. C. Das, B. P. Singh, and S. K. Singh, "Integrated model for earthquake risk assessment using neural network and analytic hierarchy process: Aceh province, Indonesia," *Geosci. Front.*, vol. 11, no. 2, pp. 613–634, Mar. 2020, doi: 10.1016/j.gsf.2019.07.006.

[12] L. Tang, M. Zhang, and L. Wen, "Support vector machine classification of seismic events in the Tianshan Orogenic Belt," *J. Geophys. Res. Solid Earth*, vol. 125, no. 1, pp. 1–12, Jan. 2020, doi: 10.1029/2019JB018132.

[13] N. S. M. Ridzwan and S. H. Md. Yusoff, "Machine learning for earthquake prediction: A review (2017–2021)," *Earth Sci. Inform.*, vol. 16, no. 2, pp. 1133–1149, Jun. 2023, doi: 10.1007/s12145-023-00991-z.

[14] S. S. Shukla, J. Dhanya, P. Kumar, and V. Dutt, "An ensemble random forest model for seismic energy forecast," *Nat. Hazards Earth Syst. Sci. Discuss.*, vol. 1, no. 1, pp. 1–12, Oct. 2024, doi: 10.5194/nhess-2024-129.

[15] M. S. Alvarez-Alvarado, R. Pérez-Londoño, F. J. Moncada, and L. G. Triviño, "Cyber-physical power systems: A comprehensive review about drivers, standards, and future perspectives," *SSRN Electron. J.*, vol. 1, no. 1, pp. 1–12, 2024, doi: 10.2139/ssrn.4687769.

[16] O. M. Saad, A. G. Hafez, and M. S. Soliman, "Deep learning approach for earthquake parameters classification in earthquake early warning system," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 7, pp. 1293–1297, Jul. 2021, doi: 10.1109/LGRS.2020.2998580.

[17] A. Mignan and M. Broccardo, "Neural network applications in earthquake prediction (1994–2019): Meta-analytic and statistical insights on their limitations," *Seismol. Res.* Lett., vol. 91, no. 4, pp. 2330–2342, Jul. 2020, doi: 10.1785/0220200021.

[18] D. Blank and J. Morgan, "Can deep learning predict complete ruptures in numerical megathrust faults?," *Geophys. Res. Lett.*, vol. 48, no. 18, p. e2021GL092607, Sep. 2021, doi: 10.1029/2021GL092607.

 ISSN 2723-6471

[19] B. Feng and G. C. Fox, "Spatiotemporal pattern mining for nowcasting extreme earthquakes in Southern California," in *Proc. 2021 IEEE 17th Int. Conf. eScience (eScience)*, vol. 2021, no. Sep., pp. 99–107, 2021, doi: 10.1109/eScience51609.2021.00020.

[20] S. Ommi and M. Hashemi, "Machine learning technique in the North Zagros earthquake prediction," *Appl. Comput. Geosci.*, vol. 22, no. 1, p. 100163, Jun. 2024, doi: 10.1016/j.acags.2024.100163.

[21] M. Kamal, B. Zhang, J. Cao, X. Zhang, and J. Chen, "Comparative study of artificial neural network and random forest model for susceptibility assessment of landslides induced by earthquake in the Western," *Sustainability*, vol. 14, no. 21, pp. 1–12, 2022, doi: 10.3390/su142113739.

[22] A. Pandit and S. Panda, "Prediction of earthquake magnitude using soft computing techniques: ANN and ANFIS," *Arab. J. Geosci.*, vol. 14, no. 13, pp. 1–12, Jul. 2021, doi: 10.1007/s12517-021-07594-2.

[23] M. A. Salam, L. Ibrahim, and D. S. Abdelminaam, "Earthquake prediction using hybrid machine learning techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 5, pp. 1–12, 2021, doi: 10.14569/IJACSA.2021.0120578.

[24] I. Syarif, A. Prugel-Bennett, and G. Wills, "SVM parameter optimization using grid search and genetic algorithm to improve classification performance," *TELKOMNIKA (Telecommun. Comput. Electron. Control)*, vol. 14, no. 4, pp. 1502–1509, Dec. 2016, doi: 10.12928/telkomnika.v14i4.3956.

[25] M. K. Suryadi, R. Herteno, S. W. Saputro, M. R. Faisal, and R. A. Nugroho, "Comparative study of various hyperparameter tuning on random forest classification with SMOTE and feature selection using genetic algorithm in software defect prediction," *J. Electron. Electromed. Eng. Med. Inform.*, vol. 6, no. 2, pp. 137–147, Mar. 2024, doi: 10.35882/jeeemi.v6i2.375.

[26] A. K. Mamindla and Y. Ramadevi, "ANN-ABC meta-heuristic hyperparameter tuning for mammogram classification," *J. Theor. Appl. Inf. Technol.*, vol. 101, no. 1, pp. 1–12, 2023.

[27] L. Visengeriyeva, A. Kammer, I. Bär, A. Kniesz, and M. Plöd, "CRISP-ML(Q). The ML lifecycle process," ML-Ops.org, vol. 1, no. 1, pp. 1–12, Mar. 2025.

[28] R. V. Odug, "Earthquakes | Data Analytics," Kaggle, Apr. 2025.

[29] J. Pwavodi, A. U. Ibrahim, P. C. Pwavodi, F. Al-Turjman, and A. Mohand-Said, "The role of artificial intelligence and IoT in prediction of earthquakes: Review," *Artif. Intell. Geosci.*, vol. 5, no. 1, p. 100075, Dec. 2024, doi: 10.1016/j.aiig.2024.100075.

[30] D. Baktibayev, B. Baigozha, I. Akhmetov, R. Mussabayev, A. Krassovitskiy, and A. Toleu, "Literature review on aftershock and earthquake prediction models aided by NLP summarization and ontology extraction techniques," *Procedia Comput. Sci.*, vol. 238, no. 1, pp. 579–586, 2024, doi: 10.1016/j.procs.2024.06.064.

[31] S. M. Plescia and G. P. Hayes, "Geometric controls on megathrust earthquakes," *Geophys. J. Int.*, vol. 222, no. 2, pp. 1270–1282, Aug. 2020, doi: 10.1093/gji/ggaa254.

[32] Highest Magnitude and Biggest Earthquakes," *SMS-Tsunami-Warning.com*. [Online]. Available: https://www.sms-tsunami-warning.com/pages/highest-magnitude-earthquake. [Accessed: Jan. 12, 2025].