RGNZ/MGPB

Machine Learning for Fully Automated Detection of Volcanic Seismic Signals in Real Seismic Records at Sinabung Volcano, North Sumatra

Rudarsko-geološko-naftni zbornik
(The Mining-Geology-Petroleum Engineering Bulletin)
DOI: 10.17794/rgn.2025.4.8

Original scientific paper

Bagas Anwar Arif Nur[1,2]*, Mohammad Hasib[2], Estu Kriswati[2]

[1] *Department of Physics, State University of Jakarta, Jakarta.*

[2] *Research Center for Geological Disaster, National Research and Innovation Agency (BRIN), Bandung.*

Abstract

To mitigate damage from volcanic disasters, experts conduct thorough monitoring mainly through observing volcano earthquake types, as these patterns provide crucial insight into magma movement. Traditionally, the classification of earthquake types has been performed manually, a process that is both time-consuming and subjective. To solve this problem, several studies have developed machine-learning models to classify volcano earthquake types automatically. However, previous research typically trains and evaluates machine learning models on existing datasets. Although these models can classify volcanic earthquake types, they require manual selection of input earthquake signals for classification. To fill this gap, this study aims to develop a machine learning model that integrates with an event detection algorithm to enable fully automated detection in actual seismic recordings at Sinabung Volcano. This study employs the Short-Term Average/Long-Term Average (STA/LTA) method, which calculates the ratio between two time windows for event detection. Two machine learning models, namely a Multi-Layer Perceptron (MLP) based on neural networks and a Random Forest (RF) based on decision trees, are used to classify the events detected by the STA/LTA method. Consequently, this approach enables the machine learning models to operate fully automated. In this study, we first detect events using the STA/LTA method on a daily basis; subsequently, each detected event is classified using a machine learning model developed based on the dataset. RF and MLP successfully predict relative low difference percentage compared to the actual number in earthquake catalogue, with values of VT 5.31%, LF 46.62%, and EXs 30.95%. Automated detection and classification can improve the efficiency of mitigating the risks associated with volcanoes by identifying potential anomalies in advance.

Keywords:
volcanic earthquake, STA/LTA, machine learning, Sinabung Volcano

1. Introduction

Volcanic eruptions are one of the natural disasters that have a significant impact on human life. Volcano monitoring could be investigated in many ways, including seismic activity, surface deformation, gas composition and emission ratio, and heat flux (**Power et al., 2021**). Seismic activity plays an essential role in monitoring volcano because it provides valuable insight into the movement of magma beneath the Earth's surface (**Battaglia et al., 2019; Hasib et al., 2022**). This movement could be used to predict and assess volcanic hazards and determine the level activity of the volcano by analyzing the number of earthquake events that occur (**Battaglia et al., 2019; Sparks et al., 2012**).

The problem with seismic monitoring is that the manual classification of volcanic earthquakes takes time because it is continuously monitored. In contrast to humans, machine learning can solve complicated problems like seismic recognition more quickly and avoid subjective decisions. Several previous studies proposed the usefulness of machine learning in classifying volcanic earthquakes. **Malfante et al. (2018)** developed machine learning models called Support Vector Machine (SVM) and Random Forest (RF) classifiers to classify the six types of volcanic earthquakes. SVM and RF were chosen in this study due to their simplicity and powerful performance, with RF having the best performance, reaching 92% accuracy. Other studies have utilized spectrograms as input to machine learning models, treating them as images to be classified by the Convolutional Neural Networks (CNN) model that has shown to have the ability to learn spatial and temporal features. The results showed accuracy of up to 96% (**Curilem et al., 2018a**) and 94% (**Calderón et al., 2020**). On the other hand, **Curilem et al. (2018b)** measured the similarity of two signals by calculating the correlation coefficient of their spectrograms and using it as input for a K-Nearest Neighbors (KNN)

* Corresponding author: Bagas Anwar Arif Nur
e-mail address: bagasanwararifnur@gmail.com
Received: 30 December 2024. Accepted: 4 March 2025.
Available online: 27 August 2025

model, which resulted in a less than 5% error. Similarly, **Pérez et al. (2022)** classified earthquakes by measuring the morphological similarity of spectrograms using Manhattan Distance (L1-norm), Euclidean Distance (L2-norm), and Distance Correlation, with this method, the system performs faster than other previous studies and performance reaches 97% accuracy.

Most previous research has relied on datasets for model testing, which may not fully reflect the performance of these models on actual seismic records. Therefore, this research aims to develop a machine-learning model that can be used on actual seismic records. The model is trained and evaluated on existing datasets and tested on actual seismic records to better understand how machine learning performs in real-world volcanic monitoring scenarios.

1.1. The Sinabung activity

Sinabung Volcano is part of the Sunda volcanic arc formed by the subduction of the Indo-Australian Plate, located in Karo, North Sumatra (see **Figure 1**). It actively began after the latest caldera-forming of Toba Lake in the Pleistocene-to-Holocene transition (74,000 years ago), divided into two main phases, old and young stages. The old and young stages correspond to the periods before and shortly after the 2010 eruption (**Hasib et al., 2024**). In the old stages, the western part was mainly formed by porphyritic two-pyroxene andesite lava flows. In the early stage, the new edifice was formed with the composition of lava flows and pyroclastic deposits, porphyritic two-pyroxene basaltic andesite to hornblende two-pyroxene andesite (**González et al., 2015**). The Sinabung Volcano was classified as a type-B until 2010, a volcano that has not shown any activity for the past 400 years. Sinabung Volcano is a stratovolcano type with a diameter of less than 5 km and an elevation of 2460 m above sea level (**Nakada et al., 2019; Pallister et al., 2019**). Sinabung Volcano activity in the 21st century first appeared in August 2010 and was followed by six additional eruptions in September, with ash column heights reaching 7 km in altitude (**Wright et al., 2019**).

Since it became active again, the Sinabung Volcano continued to erupt until the following year and reached the "AWAS" level in 2013, the highest level on Indonesia's volcano status. The Sinabung Volcano erupted on September 15 and September 17–18, 2013. A lava dome was formed on the northeast side and collapsed on December 13th, 2013, causing a pyroclastic flow (**Kriswati et al., 2024**). In mid-September 2014, a large eruption occurred with a Volcanic Explosivity Index (VEI) of 3 and ejected ash up to 6 km in altitude (**Pallister et al., 2019**). In 2016, the Center for Volcanology and Geological Hazard Mitigation (CVGHM) reported that during the period of January 4-14, 2016, hot avalanches and pyroclastic flow occurred. In the same period, ash plumes from 40 observed events rose to 3 km in altitude with high seismicity remaining up to April 2016 (**Global Volcanism Program, 2016**). In May 2016, a large eruption occurred with a dense ash plume ejected to 12 km (39,000 ft), resulting in the death of 7 people due to lahars and Pyroclastic Flow. During May-October 2016, ash plumes generally rose to an altitude of 3.3–5.5 km with an alert level of 4 on a scale of 1– 4 (**Global Volcanism Program, 2017; Hasib et al., 2024**). From October 2014 to June 2017, Sinabung continued to erupt and produce pyroclastic flows (**Nugraha et al., 2019; Hasib et al., 2024**). Then, the Sinabung eruption continued until August 2018, with the largest eruption on February 19, 2018 with a column height reaching about 19.3 km in altitude (**Kriswati et al., 2024**).

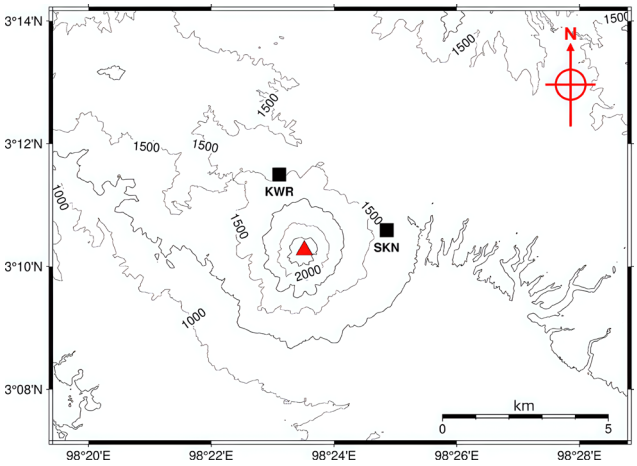

Figure 1. Location of Volcano Sinabung on Sumatra Island. The red triangle indicates the summit's location, while the black rectangles represent the seismic stations Sukanalu (SKN) and Laukawar (KWR).

2. Methods

In this study, the first stage involved development a classification model using a dataset of seismic records. The dataset consists of three classes of volcanic earthquakes: Volcano-Tectonic (VT), Low-Frequency (LF), and Explosion Earthquakes (EXs).

VT earthquakes occur due to cracks forming beneath the volcano, often caused by brittle failure under stress from fluids (**Clarke et al., 2019**). These events can indicate the volcano's stress state (**Traversa & Grasso, 2010**). VT events typically feature clear P and S wave onsets and have a dominant frequency range of 5 – 15 Hz (**Sánchez et al., 2009**). Unlike VT events, LF events lack a distinct P-wave and have dominant frequencies between 0.2 and 5 Hz. They are triggered when fluid moves or fills rock cracks (**Clarke et al., 2019**). On the other hand, EXs is a type of seismicity caused by rapid pressure release in the conduit or beneath the crater (**Hasib et al., 2019**). EXs events had a clear onset with a dominant frequency of 1 – 5 Hz (**Permana et al., 2021**). **Figure 2** is an example of a waveform and associated spectrogram of these events.

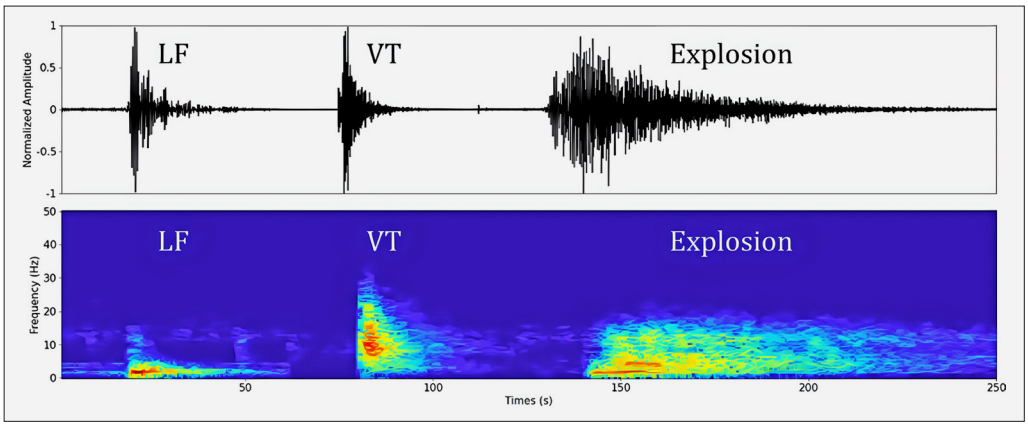

Figure 2. An example of a seismic record at Sinabung Volcano was in 2016. The left to right panel shows examples of the types of Low Frequency (LF), Volcano-Tectonic (VT), and Explosion Earthquake (EXs) events, respectively. The top panel shows the waveforms and the bottom panel shows its spectrogram.

Each class comprises 100 sample data representing the characteristics of the type of earthquake in Sinabung Volcano. The VT and LF events in this study have varying durations, ranging from 55 to 60 seconds, while EXs reach up to 300 seconds. We did not standardize the waveform durations to capture the complete information of each waveform during feature extraction, ensuring that the entire earthquake wave was analyzed. This approach helps prevent potential time domain distortions and minimizes time-frequency representation alterations, particularly for longer-duration waveforms. The label for each data member follows the seismicity criteria on Sinabung Volcano mentioned by **Gunawan et al. (2019)** and the CVGHM catalogue (**CVGHM., 2023**).

After obtaining labelled data, we performed feature extraction, a process that transforms raw data into meaningful information or value called features for analysis (see **Figure 3**). This stage involves statistical calculation and specific properties of the signal from multiple domains, including time, frequency domain, and time-frequency representation (spectrogram). The features suggested in **Table 1** were selected based on their capacity to provide details about wave shape, spectrum content, and time-varying spectrum.

The value of each extracted feature exhibits a different range of values. Certain features will have a smaller range than others, which leads to some classifiers ignoring them and only focusing on features with a large range (**Singh & Singh, 2020**). Normalization was performed to resolve this issue by scaling the feature values to fall within the same range and ensuring that all features contributed evenly during the training phase and were not ignored by classifiers, which improved the model's performance (**Pandey & Jain, 2017; Singh & Singh, 2020; Pei et al., 2023**) (see **Figure S1 in the Supplementary data**). In this stage, we also performed Principal Component Analysis (PCA) to reduce the dimensionality while preserving essential information. PCA offers several benefits. First, it helps uncover hidden patterns that may not be visible in higher dimensional space. Second, it improves the model's performance by removing the influence of noise and irrelevant features. Additionally, PCA facilitates visual analysis by projecting data into lower dimensional space and allowing us to determine the degree of similarity between group members based on their distance from each other.

Table 1. Feature List

Feature Name	Formula	Domain
Mean	$\mu_s = \frac{1}{n}\sum_i s[i]$	Time Domain, Frequency Domain, and Spectrogram
Kurtosis	$\frac{1}{n}\cdot\sum_i\left(\frac{s[i]-\mu_s}{\sigma_s}\right)^4$	
Skewness	$\frac{1}{n}\cdot\sum_i\left(\frac{s[i]-\mu_s}{\sigma_s}\right)^3$	
Shannon Entropy	$-\sum_i p(s_i)\log_2(p(s_i))$	Time Domain
Renyi Entropy	$\frac{1}{1-a}\cdot\log_2\sum_i p(s_i)^a$	
Zero Crossing Rate	$\frac{count(s[k]s[k-1]<0)}{length(s[k])}$	
Energy	$E=\sum_i s[i]^2$	
I of Central Energy	$\bar{\imath}=\frac{1}{E}\cdot\sum_i E_i\cdot i$	
Standard Deviation	$\sigma_s=\sqrt{\frac{1}{(n-1)}\sum_i\left(s[i]-\mu_s\right)^2}$	Frequency Domain
Frequency Dominant	$max_i(fft(s))$	

After the feature extraction process was complete, the model was trained to classify types of volcanic earth-

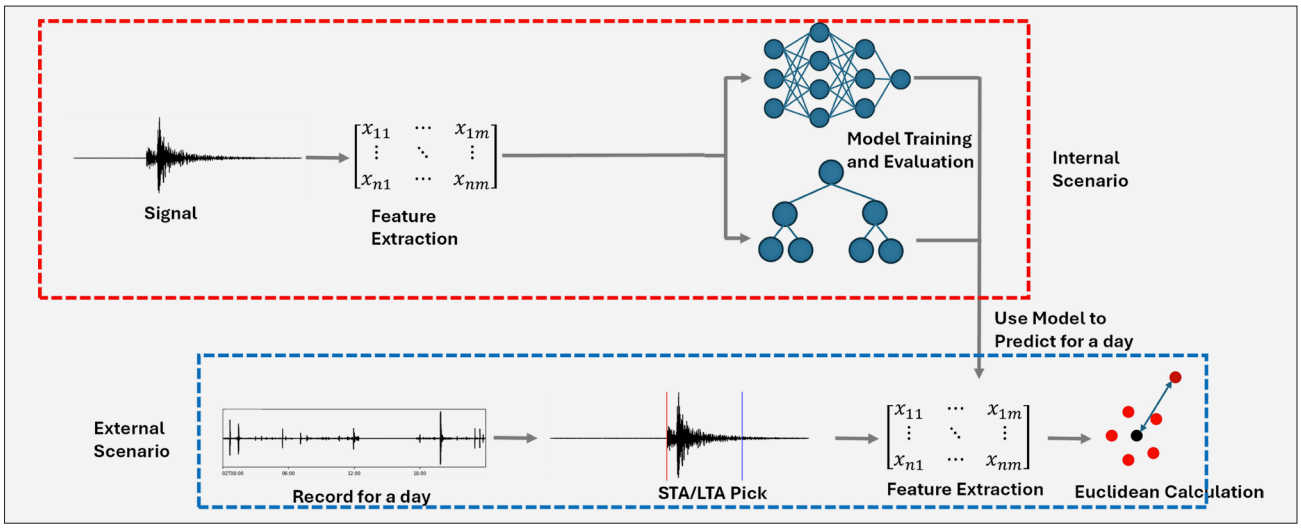

Figure 3. Workflow of the proposed method. The section envelope, a red rectangle, shows the dataset training and evaluation flow. Meanwhile, the section envelope, a blue rectangle, shows the real seismic record testing and evaluation flow.

quakes. The model used is Multi-Layer Perceptron (MLP) and Random Forest (RF), which were successfully used in the case of classifying types of volcanic earthquakes (**Langer et al., 2009; Malfante et al., 2018**). MLP is a model in machine learning that imitates the principles of biological neural networks. MLP has structure generally consists of three parts, namely the input layer, hidden layer, and output layer, which are connected (**Shah et al., 2017**). MLP works in two stages, the feed-forward stage, where the input value is predicted. In the feed-forward stage, the MLP model will calculate the error by comparing the result and prediction. The error value from this calculation will be used to update the weight in the backpropagation phase (**Jahani & Saffariha, 2022**). MLP has good capabilities in the case of non-linear data due to the use of non-linear activation functions in the hidden and output layers (**Jahani & Saffariha, 2022; Langer et al., 2009**).

Meanwhile, Random Forest (RF) is a supervised model in machine learning which can perform classification or regression tasks and can overcome imbalanced data and outliers (**More & Rana, 2017**). Random Forest is an ensemble learning method that uses the concept of bagging or so-called bootstrap aggregating. The idea of Random Forest is to ensemble the number of decision tree. Each tree provides a vote for the predicted class of a given *x,* with the final prediction determined by majority voting among trees in the ensemble (**Kulkarni & Lowe, 2016; Y. Liu et al., 2012**). In general, increasing the number of classifiers used will increase the resulting accuracy in the context of bagging (**Kulkarni & Lowe, 2016**). Both models are reliable for real-word prediction. MLP capable of generalization making it effective to predict unseen data. Meanwhile, Random Forest is robust to noise and well-suited for data with high variability (**Altay & Altay, 2023; Irandzad & Liu, 2024**).

The multi-layer perceptron model was developed using the LBFGS solver with a regularization strength of $\alpha = 10^{-5}$ and a neural network consisting of four hidden layers. Additionally, a Random Forest model was constructed with a maximum depth of 20, a minimum of two samples required for node splitting, and at least one sample per leaf node. The performances of trained models are evaluated on 20% of split data from the dataset and serve as the test set for evaluating model performance in terms of accuracy and precision. For comparison, we also replicate the method proposed by **Pérez et al. (2022)** and calculate the performance of L1 and L2.

Testing with actual seismic records was conducted using the STA/LTA method, with STA set to 10s and LTA 35s window lengths. The threshold of trigger-on and trigger-off for STA/LTA are 3 and 1, respectively. This specific value of window duration and trigger threshold is chosen to make the STA/LTA detection algorithm not overly sensitive and avoid false triggers during detection caused by noise spikes (**Trnkoczy, 2009**). The event detected using STA/LTA has the criteria of 20 seconds before and after the trigger. The cut was not set at the same duration to ensure that the cuts successfully captured the full waveform of the volcanic earthquake events. The result of STA/LTA detection will undergo the same feature extraction, normalization, and dimensionality reduction procedures as those applied to the training set. The normalization and dimension reduction on STA/LTA picking data use the same values that have been fit in the dataset; this is so that the new transformed values follow the criteria and ranges used in the dataset training and evaluation. This ensures that the detected events follow the same feature distribution as the training dataset, maintaining consistency in the analysis.

For the final step in this study, we predicted the event class from STA/LTA detection following the method proposed by **Ibs-von Seht in (2008)**. However, the number of earthquake types that were analysed in this study is limited. To ensure the prediction results belong to one

of predefined class, Euclidean Distance was used as a threshold value for accepting predicted results, with tested values being 0.3, 0.4, and 0.5. Evaluation in actual seismic record conducted by measuring the similarity of the trendline between the results of the model and the earthquake catalogue using Pearson correlation and difference percentage between number predicted and catalogue count. And lastly, we compare the performance of the model against L1 and L2 classifiers.

3. Results and discussion

PCA analysis is carried out to see how the data is distributed (**Ivosev et al., 2008**). **Figure 4** shows that each earthquake type occupies a distinct region in the feature space, indicating high inter-class and low intra class-variation. These characteristics suggest that the selected features effectively capture class variations, which may aid the ability of a model trained with these features set to distinguish each type of earthquake. Furthermore, as explained in the methodology section, we use Euclidean distance to limit and validate the prediction on actual seismic records. In **Figure 4**, the member of each type of earthquake event has a distance to the centroid from 0.3 – 0.5, where this value will be used as a threshold value to validate the prediction.

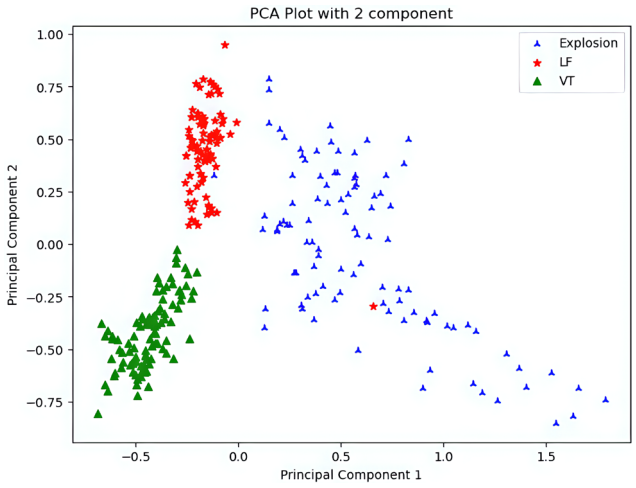

Figure 4. The dataset will be distributed using PCA with 2 principal components. The red star, green triangle, and blue asterisk show the distribution of LF, VT, and EXs, respectively.

Tables 2 and **3** to **6** show the performance and the confusion matrix for all models that have been developed, respectively. For the model developed by RF and MLP, the accuracy and precision values reached 98%. This high performance demonstrates the models' ability to accurately discriminate between different volcanic earthquake. The L1-norm model showed the lowest performance, with an accuracy of only 55% and a precision of 68%. The lower accuracy compared to **Pérez et al. (2022)** may be due to the differences in the number of types of analyzed earthquakes. Our study examines three types of volcano earthquakes (VT, LF, and EXs), while the previous studies only examine two types (VT and LF). This difference in accuracy is likely due to the EXs sharing some characteristics with VT and LF, like low frequency component and long period phase (**Cárdenas-Peña et al., 2013; Miwa et al., 2009; Titos et al., 2020**). (See **Table S3.** and **S4.** in the Supplementary Material).

Testing on 20% of the dataset indicates that the developed model performs robustly. However, testing is necessary on the actual seismic records to develop a fully automated system. According to the scheme explained in the methodology section, testing for real-world data involves detecting volcanic events using STA/LTA and

Table 2. Performance Metric for Classifier

Classifier	Accuracy	Precision
MLP	98.33%	98.42%
RF	98.33%	98.42%
L1	55%	68.65%
L2	73.33%	83.17%

Table 3. Confusion Matrix of MLP

		Predicted Label		
		LF	EXs	VT
True Label	LF	18	0	0
	EXs	1	22	0
	VT	0	0	19

Table 4. Confusion Matrix of RF

		Predicted Label		
		LF	EXs	VT
True Label	LF	18	0	0
	EXs	1	22	0
	VT	0	0	19

Table 5. Confusion Matrix of L1

		Predicted Label		
		LF	EXs	VT
True Label	LF	3	11	4
	EXs	0	22	1
	VT	0	11	8

Table 6. Confusion Matrix of L2

		Predicted Label		
		LF	EXs	VT
True Label	LF	18	0	0
	EXs	12	11	0
	VT	3	1	15

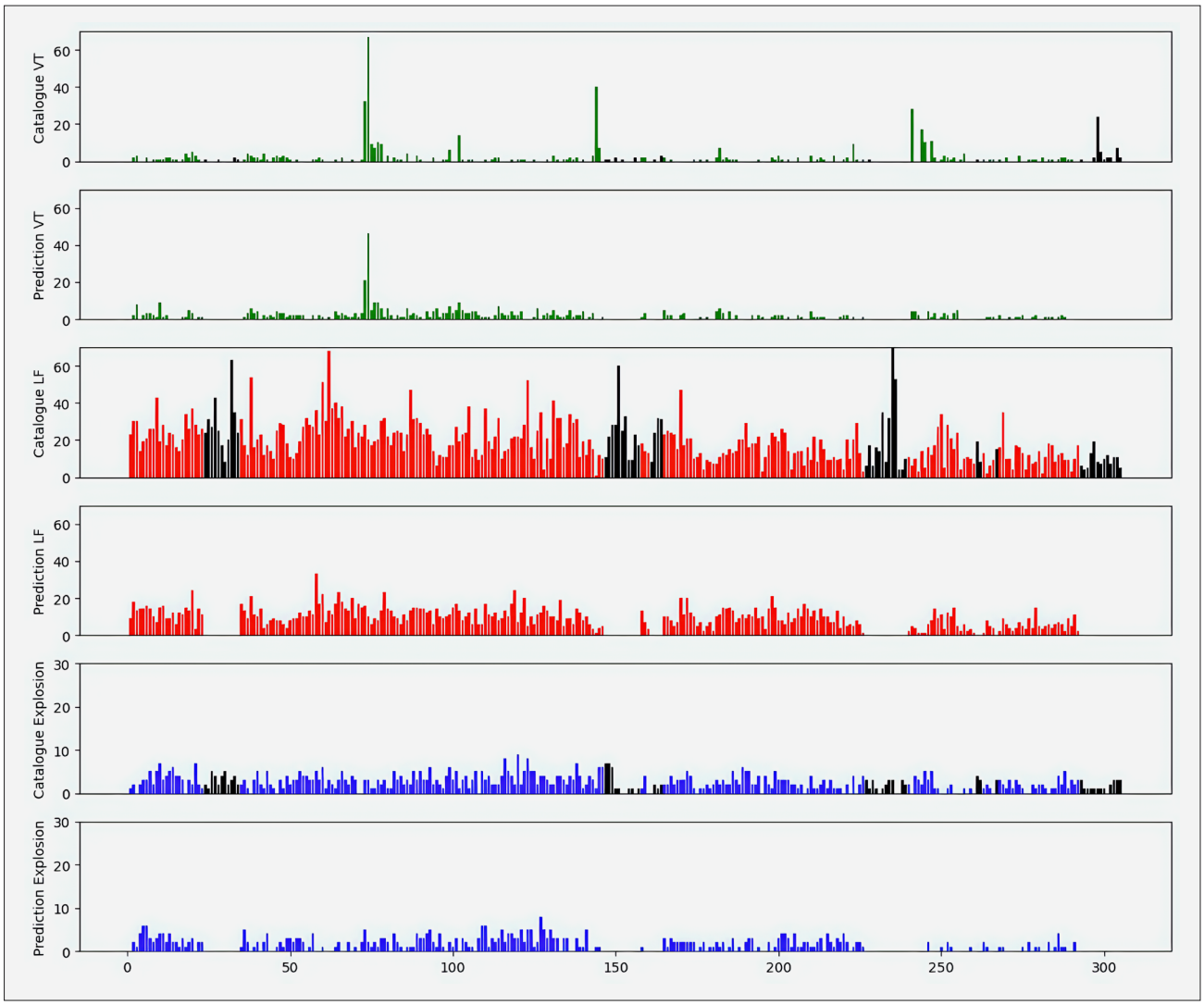

Figure 5. The comparison between prediction results using MLP and earthquake catalogue for each day. The top panel for each colour shows the earthquake catalogue and the bottom panel shows the total number of earthquake predictions. Black does not present record data.

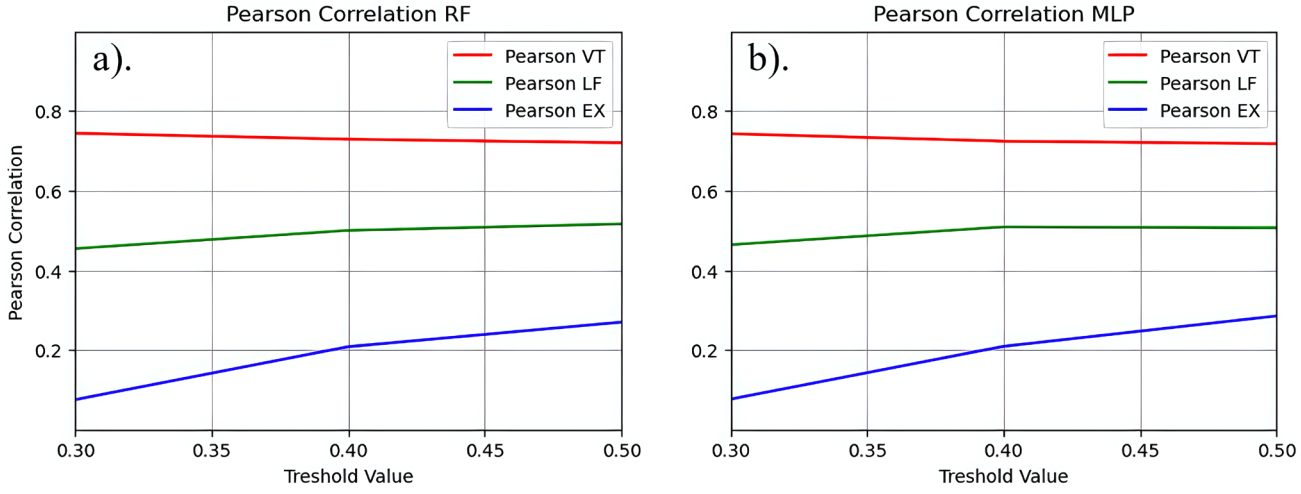

Figure 6. Pearson correlation for each type of volcano earthquake. The left shows the result using a) RF, and right b) using MLP.

classifying with the developed RF and MLP. We also use two additional classifier methods (L1 and L2) for comparison. Euclidean distance is used to measure the distance for each data point to the centroid of the corresponding class, which helps avoid prediction of earthquake type that is not included in our analysis based on smoothness assumption (**Pérez et al., 2022; van Engelen & Hoos, 2020**). **Figure 5** compares the predicted results using a machine learning model (MLP) that has been developed and the earthquake catalogue for daily events in January – October 2016. (For another classifier, please see **Figure S2** to **S4** in the Supplementary data).

Figure 6 illustrates the evaluation of trendline similarity between the predicted results and the earthquake catalogue, using the Pearson correlation coefficient (**Ahlgren et al., 2003; Sun et al., 2022**). The Pearson correlation coefficient has a score ranging from -1 to +1 (**Mana & Sasipraba, 2021**). In this context, we treated the predicted results and the catalogue counts as interdependent variables, meaning both variables should have positive Pearson coefficients indicating that they increase together, reflecting their similarity (**Di Lena & Margara, 2010**).

Evaluation of trendline similarity is conducted by varying the threshold value or Euclidean distance for the model developed by machine learning (RF and MLP) and calculated with Pearson correlation in **Figure 6**. **Figure 6** shows that for both RF and MLP models, the trendline similarity values of VT and LF type tend to be constant or have no significant change for all threshold values. In contrast, the trendline for EXs becomes more similar with an increased threshold. As indicated in **Figure 4**, this outcome is expected since EXs exhibit a broader distribution. Hence, a high threshold is required to predict data points be farther from the centroid.

The calculation of the difference between the predicted number and the catalogue is shown in **Table 7**. A smaller difference percentage indicates a higher similarity between the predicted results and the earthquake catalogue. LF events have characteristics where they often exhibit emergent signals without unclear onset, this characteristic may result in a large difference percentage in LF class as a result using the STA/LTA method. STA/LTA operates by calculating the ratio between the average of the short-term and long-term window, emergent signal lacks a contrast transition between these windows. As a result, the ratio between two window will be low and fail to trigger the detection (**McNutt & Roman, 2015; Saccorotti & Lokmer, 2021**). **Tables 5** and **6** show that L1 and L2 often misclassify other earthquake event types as EXs. Leading to a mismatch between predictions and the catalogue when applied to real seismic records in **Table 7**. On the other hand, the MLP and RF models demonstrate strong capabilities in predicting data from real seismic records, as presented in **Table 7**. For VT and LF events, the predictions from the MLP model align more closely with the values in the catalogue, likely due to its generalization ability. In contrast, the RF model excels in predicting EXs events, likely due to its robustness against noise and outliers, allowing it to be able to predict EXs with significant variability in the dataset.

Table 7. Difference percentage prediction for all classifiers

Classifier	difference Percentage		
	VT	LF	EXs
MLP	5.31%	46.62%	34.75%
RF	22.04%	59.97%	30.95%
L1	17.34%	98.90%	455.03%
L2	8.16%	51.09%	134.146%

Figures 7, **8**, and **9** are examples of the model prediction results derived from STA/LTA picks. **Figure 7** presents the prediction results for an EXs showing that while the waveform characteristics of EXs differ, their spectrum and spectrogram appear similar. This variation in the waveform may be attributed to the nature of EXs, characterized by the partitioning of energy at the source, where part travels as seismic waves through the ground and part as acoustic waves through the air. The acoustic waves can couple back into the ground and be detected by the seismometer, potentially introducing additional complexity in the waveform due to propagation effects such as reflections or refractions, which may lead to alteration in the amplitude of the recorded signal (**McNutt & Roman, 2015**), causing wider data distribution of EXs, as seen in **Figure 4**. On the other hand, **Figures 8** and **9** show that VT and LF events have clearer, more visible similarities. These similarities may have caused the data distribution for these two types, as shown in **Figure 4**, to be more concentrated, with members of each group being closer to one another.

Table 8 displays the computational time for all classifiers tested in the dataset. According to **Table 8**, the L1-norm classifier is the fastest, with a computational time of 3.49 s, whereas the RF classifier is the slowest, requiring 5.75 s, this result is aligned with the previous study. According to **Pérez et al. (2022)**, L1-norm is the fastest because it is performed in linear time by only summing the absolute difference, meaning that its computational complexity is proportional to its input size. However, to build actual seismic record automated classification systems, computational time and performance need to balance each other. Our method offers both efficiency and accuracy.

However, our studies have some limitations that from the detection method where STA/LTA is less sensitive to detect events without a clear onset. Additionally, no parameter tuning was conducted, which may impact the optimality of the model's performance and the limited number of samples in the dataset. Future work should consider alternative event detection methods and incorporate hyperparameter tuning to enhance model performance (**Wu et al., 2019; Yang & Shami, 2020**).

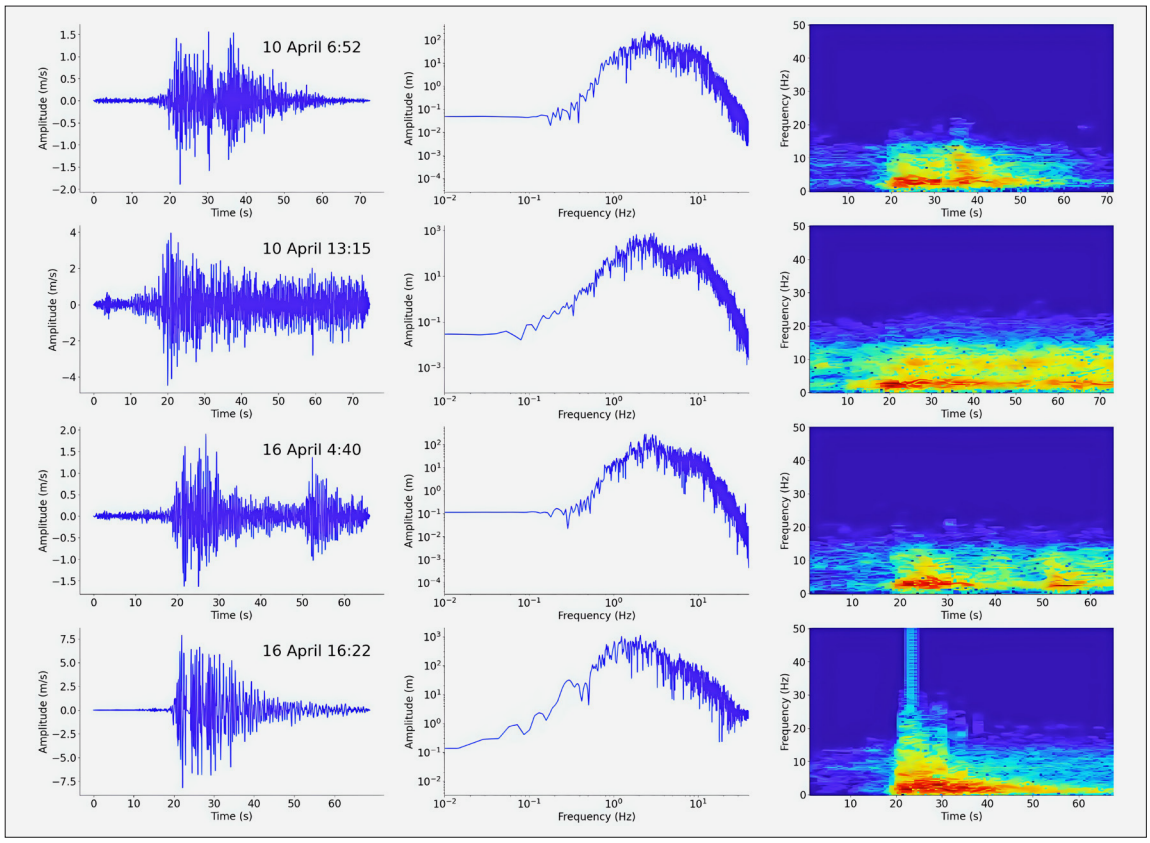

Figure 7. Prediction result from MLP. The left to right panel shows waveforms, frequency spectrum, and corresponding spectrogram, respectively, of EXs.

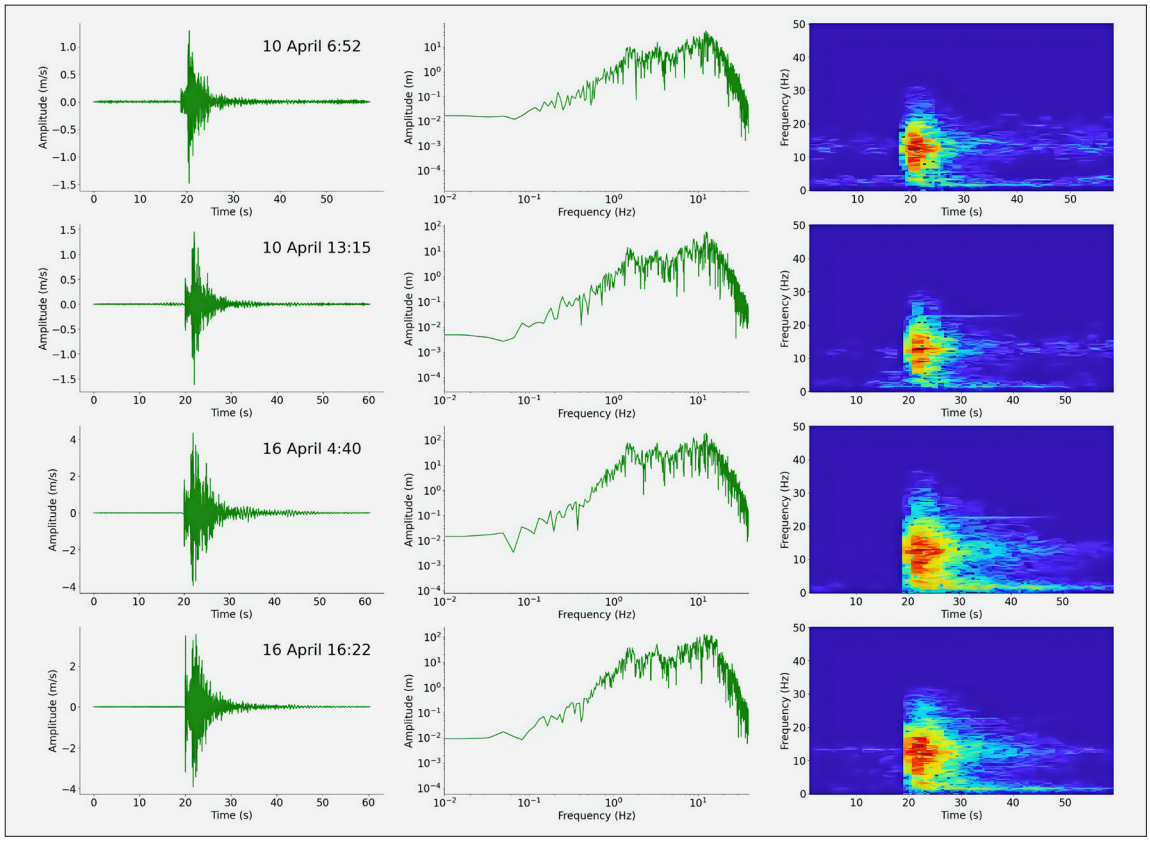

Figure 8. Prediction result from MLP. The left to right panel shows waveforms, frequency spectrum, and corresponding spectrogram, respectively, of VT.

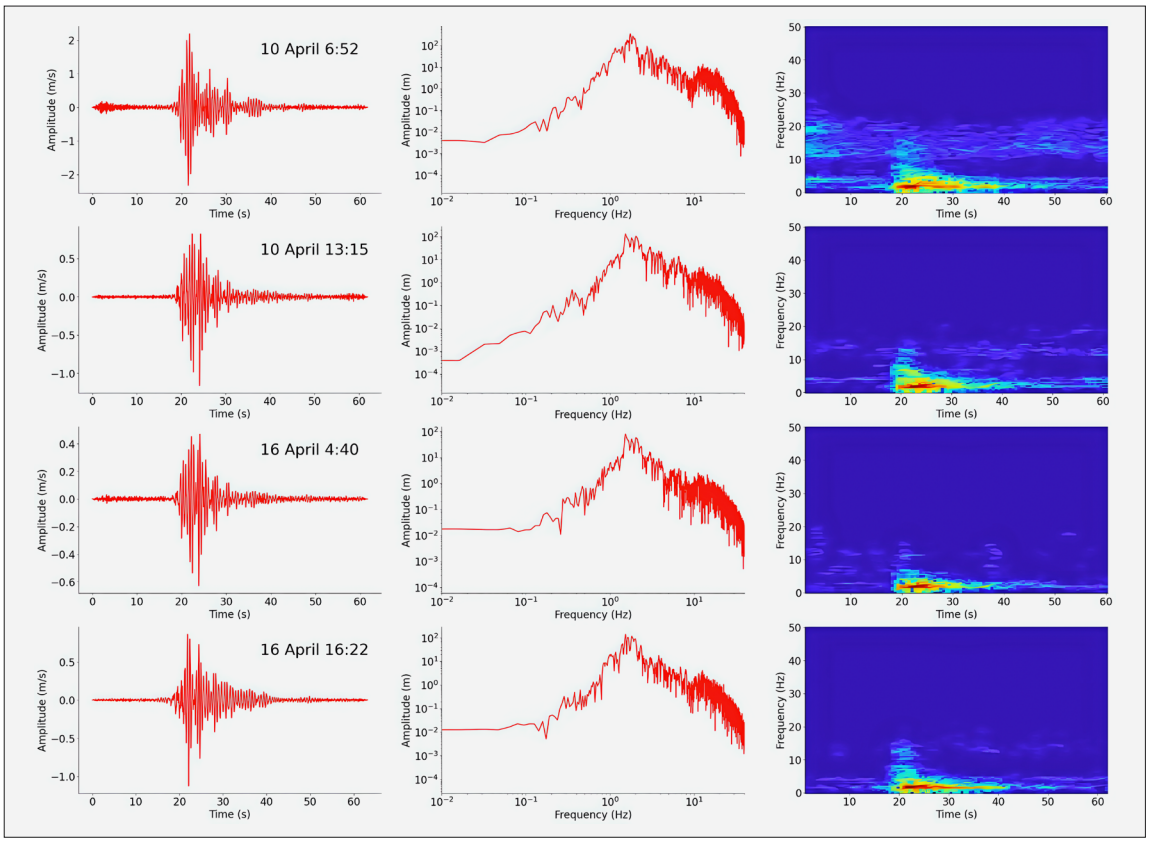

Figure 9. Prediction result from MLP. The left-to-right panel shows waveforms, the frequency spectrum, and the corresponding spectrogram of LF, respectively.

Table 8. Computational time for all classifiers

Classifier	Time (s)
MLP	5.66
RF	5.75
L1	3.49
L2	3.60

5. Conclusions

Understanding volcanic seismic activity is crucial for mitigating volcanic disasters and automated classification system for detecting volcano earthquake events in actual seismic records play a crucial role. In this study, MLP and RF machine learning models were developed to identify the volcanic earthquake types in actual seismic records. The results show that the prediction of machine learning models has a relatively low difference percentage compared to the actual number in earthquake catalogue, with values of VT 5.31%, LF 46.62%, and EXs 30.95%. The MLP has better performance in detecting VT and LF due to its strong generalization capability. Meanwhile, RF is more capable of detecting EXs as its robust to high-variability class. In our result, both models offer time efficiency and high performance, achieving an overall accuracy of 98% on a dataset evaluation with a 5-second duration running time. However, this study has limitations related to the detection methodology, model tuning, and data availability. For future studies, several improvements can be explored, such as employing a detection method more sensitive to all types of volcanic earthquakes, conducting hyperparameter tuning, or using semi-supervised learning to overcome the limitation of data. This study demonstrates that the development of an automatic earthquake classification system can be achieved using machine learning. The implementation of such an automated system offers several benefits, particularly in terms of time efficiency, which is a crucial factor in disaster mitigation related to volcanic activity. Moreover, the ability of machine learning to eliminate subjective assessments provides critical information for decision-making in evaluating volcanic hazard risks.

Acknowledgement

We thank the Centre of Volcanology and Geological Hazard Mitigation (PVMBG) for providing seismic data. This study was funded by the National Research and Innovation Agency (BRIN), under the research assistant scheme 2023-2024 Number/No. 129/II/HK/2023.

6. References

Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special ref-

erence to Pearson's correlation coefficient. Journal of the American Society for Information Science and Technology, 54(6), 550–560. https://doi.org/https://doi.org/10.1002/asi.10242

Altay, O., & Varol Altay, E. (2023). A novel hybrid multilayer perceptron neural network with improved grey wolf optimizer. Neural Computing and Applications, 35(1), 529–556. https://doi.org/10.1007/s00521-022-07775-4

Battaglia, M., Alpala, J., Alpala, R., Angarita, M., Arcos, D., Euillades, L., Euillades, P., Muller, C., Narvaez, L., & others. (2019). Monitoring volcanic deformation. Reference Module in Earth Systems and Environmental Sciences.

Calderón, F., Pérez, N., Benítez, D. S., & Riofrío, D. (2020). Volcanic Seismic Event Classification based on CNN Architectures. 2020 IEEE ANDESCON, 1–6. https://doi.org/10.1109/ANDESCON50619.2020.9272099

Cárdenas-Peña, D., Orozco-Alzate, M., & Castellanos-Dominguez, G. (2013). Selection of time-variant features for earthquake classification at the Nevado-del-Ruiz volcano. Computers & Geosciences, 51, 293–304. https://doi.org/https://doi.org/10.1016/j.cageo.2012.08.012

Clarke, J., Adam, L., Sarout, J., van Wijk, K., Kennedy, B., & Dautriat, J. (2019). The relation between viscosity and acoustic emissions as a laboratory analogue for volcano seismicity. Geology, 47(6), 499–503. https://doi.org/10.1130/G45446.1

Curilem, M., Canário, J. P., Franco, L., & Rios, R. A. (2018a). Using CNN To Classify Spectrograms of Seismic Events From Llaima Volcano (Chile). 2018 International Joint Conference on Neural Networks (IJCNN), 1–8. https://doi.org/10.1109/IJCNN.2018.8489285

Curilem, M., de Mello, R. F., Huenupan, F., San Martin, C., Franco, L., Hernández, E., & Rios, R. A. (2018b). Discriminating seismic events of the Llaima volcano (Chile) based on spectrogram cross-correlations. Journal of Volcanology and Geothermal Research, 367, 63–78. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2018.10.023

CVHGM, (2023). Sinabung Volcano Activity during 2016 Activity Report (in communication).

Di Lena, P., & Margara, L. (2010). Optimal global alignment of signals by maximization of Pearson correlation. Information Processing Letters, 110(16), 679–686. https://doi.org/https://doi.org/10.1016/j.ipl.2010.05.024

Global Volcanism Program, (2016). Report on Sinabung (Indonesia) (Venzke, E., ed.). Bulletin of the Global Volcanism Network, 41:9. Smithsonian Institution. https://doi.org/10.5479/si.GVP.BGVN201609-261080

Global Volcanism Program, (2017). Report on Sinabung (Indonesia) (Venzke, E., ed.). Bulletin of the Global Volcanism Network, 42:2. Smithsonian Institution. https://doi.org/10.5479/si.GVP.BGVN201702-261080

Gunawan, H., Budianto, A., Prambada, O., McCausland, W., Pallister, J., Iguchi, M., & others. (2019). Overview of the eruptions of Sinabung Volcano, 2010 and 2013--present and details of the 2013 phreatomagmatic phase. Journal of Volcanology and Geothermal Research, 382, 103–119.

González, P. J., Singh, K. D., & Tiampo, K. F. (2015). Shallow Hydrothermal Pressurization before the 2010 Eruption of Mount Sinabung Volcano, Indonesia, Observed by use of ALOS Satellite Radar Interferometry. Pure and Applied Geophysics, 172(11), 3229–3245. https://doi.org/10.1007/s00024-014-0915-7

Hasib, M., Nishimura, T., Anggono, T., Syuhada, Febriani, F., Sulaiman, A., Dewi, C. N., & Prasetio, A. D. (2022). Temporal changes of acoustic and seismic wave radiations: Preliminary results of application on Vulcanian eruptions at Sakurajima Volcano, Japan. AIP Conference Proceedings, 2652(1), 30003. https://doi.org/10.1063/5.0106335

Hasib, M., Nishimura, T., & Nakahara, H. (2019). Spectral ratio analyses of explosion earthquakes at Sakurajima Volcano, Japan. Journal of Volcanology and Geothermal Research, 381, 302–311. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2019.05.005

Hasib, M., Saepuloh, A., Aulia, A. N., Muttaqy, F., Anggono, T., Ramdhan, M., Kriswati, E., Syuhada, Febriani, F., Prasetio, A. D., Dewi, C. N., Indrastuti, N., & Ghifari, B. F. (2024). Revealing the shallow magmatic plumbing system of Sinabung Volcano during 2014–2017 eruption events using seismic tomography. Journal of Volcanology and Geothermal Research, 452, 108137. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2024.108137

Ibs-von Seht, M. (2008). Detection and identification of seismic signals recorded at Krakatau volcano (Indonesia) using artificial neural networks. Journal of Volcanology and Geothermal Research, 176(4), 448–456. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2008.04.015

Iranzad, R., & Liu, X. (2024). A review of random forest-based feature selection methods for data science education and applications. International Journal of Data Science and Analytics. https://doi.org/10.1007/s41060-024-00509-w

Ivosev, G., Burton, L., & Bonner, R. (2008). Dimensionality Reduction and Visualization in Principal Component Analysis. Analytical Chemistry, 80(13), 4933–4944. https://doi.org/10.1021/ac800110w

Jahani, A., & Saffariha, M. (2022). Environmental decision support system for Plane trees failure prediction: A comparison of multi-layer perceptron and random forest modeling approaches. Agrosystems, Geosciences & Environment, 5(4), e20316. https://doi.org/https://doi.org/10.1002/agg2.20316

Kriswati, E., Meilano, I., Hasib, M., Saepuloh, A., Kuncoro, H., Dewanto, B. G., & Fuadi, A. (2024). Explosion mechanism and volume estimation of volcanic ash during the eruption of Sinabung Volcano on February 19, 2018: Insight from kinematic GPS and seismic data. Journal of Volcanology and Geothermal Research, 447, 108034. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2024.108034

Kulkarni, A. D., & Lowe, B. (2016). Random forest algorithm for land cover classification.

Langer, H., Falsaperla, S., Masotti, M., Campanini, R., Spampinato, S., & Messina, A. (2009). Synopsis of supervised and unsupervised pattern classification techniques applied to volcanic tremor data at Mt Etna, Italy. Geophysical Journal International, 178(2), 1132–1144. https://doi.org/10.1111/j.1365-246X.2009.04179.x

Liu, Y., Wang, Y., & Zhang, J. (2012). New Machine Learning Algorithm: Random Forest. In B. Liu, M. Ma, & J. Chang (Eds.), Information Computing and Applications (pp. 246–252). Springer Berlin Heidelberg.

Malfante, M., Dalla Mura, M., Mars, J. I., Métaxian, J.-P., Macedo, O., & Inza, A. (2018). Automatic Classification of Volcano Seismic Signatures. Journal of Geophysical Research: Solid Earth, 123(12), 10,610-645,658. https://doi.org/https://doi.org/10.1029/2018JB015470

Mana, S. C., & Sasipraba, T. (2021). Research on Cosine Similarity and Pearson Correlation Based Recommendation Models. Journal of Physics: Conference Series, 1770(1), 12014. https://doi.org/10.1088/1742-6596/1770/1/012014

McNutt, S. R., & Roman, D. C. (2015). Chapter 59 - Volcanic Seismicity. In H. Sigurdsson (Ed.), The Encyclopedia of Volcanoes (Second Edition) (Second Edi, pp. 1011–1034). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-385938-9.00059-6

Miwa, T., Toramaru, A., & Iguchi, M. (2009). Correlations of volcanic ash texture with explosion earthquakes from vulcanian eruptions at Sakurajima volcano, Japan. Journal of Volcanology and Geothermal Research, 184(3), 473–486. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2009.05.012

More, A. S., & Rana, D. P. (2017). Review of random forest classification techniques to resolve data imbalance. 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM), 72–78. https://doi.org/10.1109/ICISIM.2017.8122151

Nakada, S., Zaennudin, A., Yoshimoto, M., Maeno, F., Suzuki, Y., Hokanishi, N., Sasaki, H., Iguchi, M., Ohkura, T., Gunawan, H., & Triastuty, H. (2019). Growth process of the lava dome/flow complex at Sinabung Volcano during 2013–2016. Journal of Volcanology and Geothermal Research, 382, 120–136. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2017.06.012

Nugraha, A. D., Indrastuti, N., Kusnandar, R., Gunawan, H., McCausland, W., Aulia, A. N., & Harlianti, U. (2019). Joint 3-D tomographic imaging of Vp, Vs and Vp/Vs and hypocenter relocation at Sinabung volcano, Indonesia from November to December 2013. Journal of Volcanology and Geothermal Research, 382, 210–223. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2017.09.018

Pallister, J., Wessels, R., Griswold, J., McCausland, W., Kartadinata, N., Gunawan, H., Budianto, A., & Primulyana, S. (2019). Monitoring, forecasting collapse events, and mapping pyroclastic deposits at Sinabung volcano with satellite imagery. Journal of Volcanology and Geothermal Research, 382, 149–163. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2018.05.012

Pandey, A., & Jain, A. (2017). Comparative analysis of KNN algorithm using various normalization techniques. International Journal of Computer Network and Information Security, 10(11), 36.

Pei, X., hong Zhao, Y., Chen, L., Guo, Q., Duan, Z., Pan, Y., & Hou, H. (2023). Robustness of machine learning to color, size change, normalization, and image enhancement on micrograph datasets with large sample differences. Materials & Design, 232, 112086. https://doi.org/https://doi.org/10.1016/j.matdes.2023.112086

Pérez, N., Granda, F. S., Benítez, D., Grijalva, F., & Lara, R. (2022). Toward Real-Time Volcano Seismic Events' Classification: A New Approach Using Mathematical Morphology and Similarity Criteria. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–13. https://doi.org/10.1109/TGRS.2020.3048107

Permana, T., Nishimura, T., Nakahara, H., & Shapiro, N. (2021). Classification of volcanic tremors and earthquakes based on seismic correlation: application at Sakurajima volcano, Japan. Geophysical Journal International, 229(2), 1077–1097. https://doi.org/10.1093/gji/ggab517

Power, J. A., Roman, D. C., Lyons, J. J., Haney, M. M., Rasmussen, D. J., Plank, T., Nicolaysen, K. P., Izbekov, P., Werner, C., & Kaufman, A. M. (2021). Volcanic seismicity beneath Chuginadak Island, Alaska (Cleveland and Tana volcanoes): Implications for magma dynamics and eruption forecasting. Journal of Volcanology and Geothermal Research, 412, 107182. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2021.107182

Saccorotti, G., & Lokmer, I. (2021). Chapter 2 - A review of seismic methods for monitoring and understanding active volcanoes. In P. Papale (Ed.), Forecasting and Planning for Volcanic Hazards, Risks, and Disasters (Vol. 2, pp. 25–73). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-12-818082-2.00002-0

Sánchez, J. J., Núñez-Cornú, F. J., Suárez-Plascencia, C., & Trejo-Gómez, E. (2009). seismicity at ceboruco Volcano, Mexico. Seismological Research Letters, 80(5), 823–830.

Shah, A. S., Shah, M., Fayaz, M., Wahid, F., Khan, H. K., & Shah, A. (2017). Forensic analysis of offline signatures using multilayer perceptron and random forest. International Journal of Database Theory and Application, 10(1), 139–148.

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. Applied Soft Computing, 97, 105524. https://doi.org/https://doi.org/10.1016/j.asoc.2019.105524

Sparks, R. S. J., Biggs, J., & Neuberg, J. W. (2012). Monitoring volcanoes. Science, 335(6074), 1310–1311.

Sun, H., Zhu, L., Guo, L., Luo, Y., Du, D., & Sun, Y. (2022). Understanding the different responses from the similarity between displacement and groundwater level time series in Beijing, China. Natural Hazards, 111(1), 1–18. https://doi.org/10.1007/s11069-021-05041-9

Traversa, P., & Grasso, J.-R. (2010). How is Volcano Seismicity Different from Tectonic Seismicity? Bulletin of the Seismological Society of America, 100(4), 1755–1769. https://doi.org/10.1785/0120090214

Trnkoczy, A. (2009). Understanding and parameter setting of STA/LTA trigger algorithm. In New manual of seismological observatory practice (NMSOP) (pp. 1–20). Deutsches GeoForschungsZentrum GFZ.

van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. Machine Learning, 109(2), 373–440. https://doi.org/10.1007/s10994-019-05855-6

Wright, H. M. N., Pallister, J. S., McCausland, W. A., Griswold, J. P., Andreastuti, S., Budianto, A., Primulyana, S., Gunawan, H., Battaglia, M., Diefenbach, A., Griswold, J., Ewert, J., Kelly, P., Kern, C., LaFevers, M., Lockhart, A.,

Marso, J., Mayberry, G., McCausland, W., … Triastuty, H. (2019). Construction of probabilistic event trees for eruption forecasting at Sinabung volcano, Indonesia 2013–14. Journal of Volcanology and Geothermal Research, 382, 233–252. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2018.02.003

Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., & Deng, S.-H. (2019). Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimizationb. Journal of Electronic Science and Technology, 17(1), 26–40. https://doi.org/https://doi.org/10.11989/JEST.1674-862X.80904120

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing, 415, 295–316. https://doi.org/https://doi.org/10.1016/j.neucom.2020.07.061

SAŽETAK

Strojno učenje za potpuno automatizirano otkrivanje vulkanskih seizmičkih signala u stvarnim seizmičkim zapisima na vulkanu Sinabung, Sjeverna Sumatra

Kako bi ublažili štetu od vulkanskih katastrofa, stručnjaci provode temeljito praćenje uglavnom promatranjem vrsta vulkanskih potresa jer ti obrasci daju ključne uvide u kretanje magme. Klasifikacija tipova potresa tradicionalno se provodila ručno, proces koji je dugotrajan i subjektivan. Kako bi se riješio taj problem, nekoliko je studija razvilo modele strojnoga učenja za automatsku klasifikaciju tipova vulkanskih potresa. Međutim, prethodna istraživanja obično obučavaju i ocjenjuju modele strojnoga učenja na postojećim skupovima podataka. Iako ti modeli mogu klasificirati tipove vulkanskih potresa, oni zahtijevaju ručni odabir ulaznih signala potresa za klasifikaciju. Kako bi se popunila ta praznina, cilj je ove studije razviti model strojnoga učenja koji se integrira s algoritmom za otkrivanje događaja kako bi se omogućilo potpuno automatizirano otkrivanje u stvarnim seizmičkim snimkama na vulkanu Sinabung. Ova studija koristi se metodom kratkoročnoga prosjeka / dugoročnoga prosjeka (STA/LTA), koja izračunava omjer između dvaju vremenskih okvira za otkrivanje događaja. Dva modela strojnoga učenja, Multi-Layer Perceptron (MLP) temeljen na neuronskim mrežama i Random Forest (RF) temeljen na stablima odlučivanja, koriste se za klasifikaciju događaja otkrivenih metodom STA/LTA. Posljedično, ovaj pristup omogućuje potpuno automatizirani rad modela strojnoga učenja. U ovoj studiji prvo se otkrivaju događaji pomoću STA/LTA metode na dnevnoj bazi, nakon toga se svaki otkriveni događaj klasificira pomoću modela strojnoga učenja razvijenoga na temelju skupa podataka. RF i MLP uspješno predviđaju relativno nizak postotak razlike u odnosu na stvarni broj u katalogu potresa, s vrijednostima od VT 5,31 %, LF 46,62 % i EXs 30,95 %. Automatizirano otkrivanje i klasifikacija mogu poboljšati učinkovitost ublažavanja rizika povezanih s vulkanima utvrđivanjem potencijalnih anomalija unaprijed.

Ključne riječi:
vulkanski potres, STA/LTA, strojno učenje, vulkan Sinabung

Author's contribution

Bagas Anwar Arif Nur (B. Sc, Graduate Student) provided the methodology research, developing the model, and analyzed the results. **Mohammad Hasib** (M. Sc, Researcher) provided the information regarding volcano seismicity, and analyzed the results. **Estu Kriswati** (M. Sc, Researcher) provided the seismic records.
All authors have read and agreed to the published version of the manuscript.