ISSN: 2723-9535

Available online at www.HighTechJournal.org

HighTech and Innovation Journal

Vol. 6, No. 1, March, 2025

Hybrid Time Series Methods and Machine Learning for Seismic Analysis and Volcano Eruption Predict

Fridy Mandita [1, 2], Ahmad Ashari [3*], Moh. Edi Wibowo [3], Wiwit Suryanto [4]

[1] *Department of Engineering, Universitas 17 Agustus 1945 Surabaya, Surabaya 60118, Indonesia.*

[2] *Doctoral Programme of Computer Science, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.*

[3] *Department of Computer Science and Electronics, Universitas Gadjah Mada,, Yogyakarta 55281, Indonesia.*

[4] *Department of Physics, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.*

Received 02 September 2024; Revised 15 January 2025; Accepted 03 February 2025; Published 01 March 2025

Abstract

Volcanic eruption refers to a natural catastrophe on Earth that poses imminent danger to communities surrounding volcanoes. Therefore, ongoing monitoring of volcanic processes is crucial for effective analysis and observation of volcanic activities preceding an eruption. In response to this, the study presents a novel hybrid time series approach, integrated with machine learning techniques, to enhance the identification and classification of seismic events associated with volcanic eruptions. In this case, time series techniques, including STA/LTA, template matching, and autocorrelation, were implemented to facilitate the detection and classification process. The challenges, however, lie in addressing noise and ensuring accuracy in the analysis of seismic signals. To resolve this, a new hybrid time series method was proposed to improve signal analysis accuracy by integrating multiple time series techniques. In practice, the dataset was collected from Mount Merapi in Indonesia between 2019 and 2021, consisting of a compilation of seismic data categorized by event type, thus enhancing classification accuracy. On top of that, prior to implementing machine learning techniques for signal classification, the hybrid method was employed to efficiently remove noise, ensuring that genuine seismic events were clearly distinguished from spurious signals. Notably, the experimental learning rate was set at 0.01. The results demonstrated that the proposed hybrid method outperformed stand-alone time series techniques, achieving an accuracy of 0.93 to 0.95. This signifies the effectiveness of precise seismic event recognition and categorization, greatly enhancing the volcano monitoring system. Furthermore, the findings offer substantial improvements in the forecasting and risk mitigation associated with volcanic eruptions, hence, advancing reliable seismic analysis methodologies. Ultimately, the method enhances hybrid methods and machine learning for seismic event analysis and volcano monitoring.

Keywords: Seismic Events; Hybrid Time Series; Machine Learning; Volcano Eruption.

1. Introduction

Indonesia is situated in the convergence of three tectonic plate boundaries and occupies a geographically unique position referred to as the Ring of Fire (ROF), characterized by intense tectonic activity, leading to numerous active volcanoes in Indonesia, including [1] approximately 130 active volcanoes, from Sabang to Merauke [2]. This geographical position significantly increases the potential for spontaneous volcanic eruptions, such as the eruptions recorded in one of the active and hazardous volcanoes in Indonesia, Mount Merapi [3, 4].

* Corresponding author: ashari@ugm.ac.id

http://dx.doi.org/10.28991/HIJ-2025-06-01-08

As illustrated in Figure 1, Mount Merapi volcano, located in Central Java, is marked by densely populated slopes, where many residents live as close as 28 km (17 miles) north of Yogyakarta, the city near Mount Merapi with a population of 2.4 million. Considering the dense population, in addition to 73 recorded eruptions over the past 500 years, thorough investigation into the risks posed by Merapi is critical [5]. In fact, between 1672 and 2010, over 80 eruptions were reported. The rest intervals range from 1 to 18 years, averaging 4 years. The 2010 eruption caused extensive damage to community-owned properties on the slopes of the mount [6]. The latest eruption was reported on October 4, 2021, when magma supply triggered a shallow volcanic earthquake 8 km below the Earth's surface in October 2019 [7].

Considering Mount Merapi's high level of activity, monitoring the mountain's activities is essential for effective analysis. Specifically, monitoring the volcano activity requires proper data on the eruption phase of the mountain. This is critical for assessing an unsettled volcano's activity level and predicting the probability and timing of a future eruption. Moreover, tracking volcanoes is an essential component of scientific approaches to minimize hazards to human society [8]. Another point to consider: the effectiveness of event detection depends on the data quality, including the precision, completeness, consistency, and frequency of past events. Moreover, maintaining substantial datasets of eruptions and consistently prepared monitoring data is required to effectively utilize statistical analysis [9].

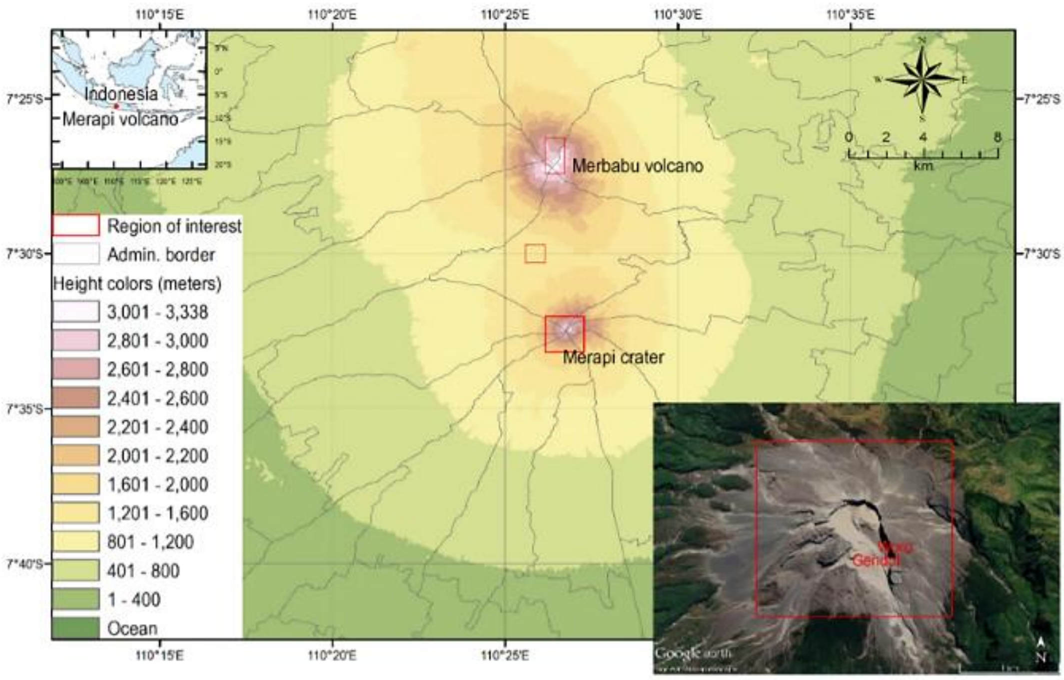

Figure 1. The Map of Merapi Mountain, Indonesia

While the majority of detection works have been automated, the categorization task remains largely reliant on manual intervention. Manual classification tasks exist, although in a limited number. In addition, a number of factors potentially cause labeling to be less reliable. As the classification depends on the operator's subjective assessment, different individuals potentially arrive at other criteria when multiple individuals perform the task. One technique for analyzing seismic data waveforms involves a time series algorithm. The technique is categorized into three primary types: 1) Short-Term Average/Long-Term Average (STA/LTA), 2) Template Matching, and 3) Autocorrelation/Cross-correlation [10]. Although time series algorithm techniques pose different ways of working from one another, the approach shares a single way of processing data from seismic signals.

The time and effort required to collect information regarding the sequence and locations of events are significantly minimized through efficient detection methods, particularly in regions with moderate seismic activity at local or regional scales [11]. Moreover, seismic signals are identified by analyzing time series data or matching patterns in seismic waveforms to determine the correlation with volcanic seismic events [12]. Typically, the event detection task is framed as a problem of classifying, wherein cutout seismic waveforms are categorized into two primary categories: earthquakes and noise [13]. In seismic investigations, signal classification classifies waveform data by attributes, including cepstrum, spectrum, and temporal waveforms. The automatic classification of seismic signals, however, remains a significant challenge, as a substantial portion of the process is conducted manually. As outlined in Falcin et al. (2021) [8], automatic recognition requires models linked to volcanic activity, heavily relying on waveform and spectrum analysis. While the process is frequently executed semi-automatically or automatically during the detection phase, the classification phase remains primarily manual. Therefore, the accuracy of manual classification varies depending on the user and is time-consuming [14-16].

Notably, the STA/LTA technique has been employed by Vaezi & van der Baan (2015) [17], utilizing statistical criteria to compare the Short-Term Average/Long-Term Average (STA/LTA) with the Power Spectral Density (PSD). This is performed by calculating the ratio of the mean energies of the data calculated consecutively over two subsequent moving time frames—a short window followed by a long window. The findings indicate that the Power Spectral Density (PSD) method outperforms the Short-Term Average/Long-Term Average (STA/LTA) method in automatically detecting seismic signals. When compared to STA/LTA, despite exhibiting superior performance in analyzing weak signals, PSD works have yet to be tested for real-time data. Moreover, this technique involves constraints and variables that require meticulous calibration. Additionally, these algorithms are sensitive towards sudden spikes in amplitude, thus capable of identifying noise as microearthquake events that possess energy equal to or surpassing actual microearthquake events [18].

Another experiment utilizing STA/LTA to detect seismic events has been conducted by Pantobe et al. (2024) [19]. The experiment, combined with CNN-SE-T for real-time detection of seismic occurrences for timely alerts and responses, is a complex endeavor that requires precise identification of P-wave arrivals. The challenge, however, relies on the complex detection of eruptive precursory signals, posing difficulty in predicting dangerous sudden phreatic or hydrothermal non-magmatic eruptions within a timely framework. The findings demonstrate that the model is capable of optimizing a velocity model in the shallow dome, despite the limited ability to automate the magnitude determination due to the low SNR and frequency resonance. Additionally, the STA/LTA presents notable disadvantages when the signals are exceedingly weak. In this instance, selecting an appropriate threshold value is particularly challenging, frequently leading to incorrect judgment.

The next method is template matching techniques that represents one of the time series methods to analyze seismic events. The method operates by detecting events based on the similarity of waveforms by calculating the cross-correlation coefficient (CC) value of a seismic signal. According to the study by Ma et al. (2020) [20] on icequake detection, the technique is effective in identifying larger events with low SNR, whereas, limited to simple waveforms.

Likewise, Yang et al. (2021) [21] performed a study by detecting the depth of microearthquake sources, in which various machines were employed to analyze micro-earthquakes utilizing a benchmarked dataset from an underground cavern collapse in South Louisiana, comprising a total of 444 datasets and 444 micro-earthquakes. The findings suggest that the CNN model demonstrates an ability to discern essential elements from the input signal and elucidate the outputs at the hidden layers. Further analysis indicates that feature selection and transformation are crucial for the performance of feature-based classifiers. Additionally, the degree of rectilinearity demonstrates a notably stronger correlation with the source depth. Regardless, template matching emerges as the most effective technique to identify recurrent, more minor earthquakes in the different traditional detection techniques that rely on the similarity of the whole waveforms. In this context, the detection accuracy is contingent upon the number of templates and the difficulty in differentiating between periodic signals and noise.

The last technique of time series is autocorrelation. Ikeda & Takagi (2019) [22] have proposed a study that involved autocorrelation to identify temporal variations in seismic velocity and scattering characteristics, particularly by autocorrelation analysis of ambient seismic noise. The study emphasizes the efficacy of seismic interferometry and autocorrelation analysis in identifying temporal alterations in subsurface structures as a result of seismic activities. Moreover, the approach provides valuable insights into the subsurface dynamics led by earthquakes. Additionally, processing requirements obstruct real-time monitoring, further influencing the outcomes due to dependency on predetermined frequency bands. Therefore, supplementary tools and enhanced methodologies.

Further study guided by Nurtas et al. (2024) [23] provides an analysis of earthquake time series forecasting utilizing Integrated Moving Average with Exogenous Variables (SARIMAX) models. The objective is to evaluate the efficacy of the SARIMAX model in predicting earthquakes by incorporating pertinent exogenous variables, including historical seismic activity, geological attributes, and geodetic data. In practice, the parameters of the SARIMAX model were determined through an analysis of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the time series. Afterward, the model was trained on historical data to predict future values of the series. The study reported valuable insights into the integration of time-series analysis with external geological elements to enhance predictive modeling. The forecast accuracy, however, was inadequate—40% reliability, reflecting a limited level of resilience. The performance was significantly affected by the quality and comprehensiveness of prior earthquake data. Additionally, the resampling of data to daily intervals introduced certain biases.

Apart from the studies, one of the primary challenges in applying machine learning in seismology is automating the process of volcano-monitoring data, which remains predominantly performed manually. Recently, machine learning (ML) has been implemented in seismology, with various applications for identifying invisible signals and patterns and extracting information features related to the field of seismology [24]. On the other hand, implementing ML in seismology has not extended to suppressing volcanic eruptions, rather essentially processing seismic data to convey information related to volcanic activity. Despite the progress of ML applications, challenges persist in seismology implementation.

The application of ML in seismology has experienced significant advancements, as demonstrated by a number of related studies. Seismological studies utilizing machine learning have been implemented in diverse areas, such as volcano monitoring, earthquake forecasting, volcanic eruption predicting, and identification and categorization of seismic vibrations. In this study, seismic signal data were specifically utilized and subjected to ML algorithm processing. Furthermore, Wiszniowski et al. (2021) [11] conducted a study on auto-discovery to initiate the event by monitoring volcano activity, requiring a vast volume of data evaluated in real-time. The findings indicate that the modified SLRNN is more effective at detecting seismic occurrences than the deep learning techniques. Furthermore, the SLRNN method requires fewer data for training samples. Nevertheless, this experiment is not applicable in real-time use.

To add to this, waveform autocorrelation and template matching methods have been widely utilized to classify seismic signals for detection. In this case, the classification based on waveform similarity achieved high recognition accuracy. These methods, however, generally require extensive databases to improve accuracy [25]. This strategy frequently decreases the magnitude of completeness by approximately one, leading to a tenfold increase in discovered occurrences. Despite the effectiveness, the methods are typically restricted to the analysis of short-term data, such as data preceding significant earthquakes, due to the expensive computational requirements described by Ross et al. (2021) [26].

These features gather seismic-related data from seismic signals such as the time waveform, spectrum, and cepstrum [14]. In addition, a variety of machine learning methods have been employed in seismology, including Neural Networks (NN), Multilayer Perceptron (MLP), Support Vector Machine (SVM), Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) [13, 27]. The methods frequently decrease the magnitude of completeness by approximately one, leading to a tenfold increase in discovered occurrences. While effective, the methods are typically limited to the analysis of data over short periods, particularly in the lead-up to significant earthquakes, due to the expensive computational requirements.

A notable advancement has been established by Centeno et al. (2024) [28], in which the study successfully implemented a CNN and U-Net to segment and measure volcanic plumes in photographs. Moreover, the study utilized boosting-based machine learning ensembles to categorize the seismic events associated with ash plumes, demonstrating the efficacy of the methods in managing data produced during seismic and eruptive emergencies. In alignment with this, Ozkaya et al. (2024) [29] introduced the use of KNN and SVM for earthquake detection and classification from seismic signal data. The study utilized a public dataset consisting of three categories: (1) noise, (2) P-waves, and (3) S-waves to define earthquakes. Seven features of the vector were implemented as inputs to the classifier seismic signal by employing KNN and SVM algorithms. The findings reached an accuracy of 90%; nonetheless, further tests with a larger dataset of seismic signals are required.

While ML technologies have markedly progressed in a multitude of domains, distinct challenges persist. The imbalance in natural datasets, for example, leads to mis-assessment or misinterpretation in numerous instances. The efficacy, precision, and adaptability of machine learning are the primary factors of the wide application within earthquake seismology. Despite the progress, numerous issues persist that machine learning effectively addresses. Moreover, the application continues to enhance and extend the understanding of earthquake seismology [13].

To further elaborate, the objective of the study is to create a forecasting model for volcanic eruptions utilizing seismic signal data, noise identification and categorization, time-series analysis, and machine-learning techniques. To ensure the comprehensiveness of the data, unprocessed seismic data were collected from Mount Merapi in a defined time frame from 2019 to 2021. Additionally, the daily frequency of occurrences was recorded. Correlating to this, the newly presented methods exhibited higher accuracy in identifying and classifying volcanic eruptions while working with a short dataset, compared to other methods. Evidently, the hybrid time-series model outperformed other models in accurately recognizing and detecting seismic events associated with volcanic eruptions.

2. Time Series Algorithm

2.1. STA/LTA Algorithm

The STA/LTA algorithm represents a significant technique that effectively reduces high-energy transients in ambient vibration recordings. A notable characteristic of this method lies in the ability to be triggered or identified based on the STA/LTA ratio [30]. The Short-Term Amplitude (STA) / Long-Term Amplitude (LTA) method is a commonly utilized methodology for earthquake detection, as depicted in Figure 2. This information stems from the findings of human experts regarding earthquake detection. According to these experts, significant fluctuations in amplitude serve as visual indicators of potential earthquakes. As a further point, the STA/LTA method relies on two fundamental parameters: the short-term and long-term window duration. The standard parameter selections consist of a brief time frame of three seconds for the short-term analysis and an extended time frame of 30 seconds for the long-term study. Furthermore, the third parameter option is introduced to modify the overlap between the short-term window and the tail end of the long-term window [10].

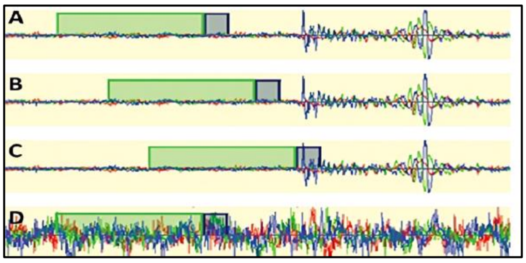

Figure 2. STA/LTA Algorithm

The STA/LTA algorithm is defined by the following equations:

$$STA\ (x_i) = \frac{1}{ns} \sum_{j=i-ns}^{i} x_i^2$$
$$LTA\ (x_i) = \frac{1}{nl} \sum_{j=i-nl}^{i} x_i^2 \tag{1}$$
$$r_i = \frac{STA_i}{LTA_i}$$

where: x_i: current sample of time series data; ns : length of the short window; nl: length of the long window; r_i: ratio of short amplitude and long amplitude

STA/LTA operates independently of historical data, thus being valuable for newly established stations where existing data are unavailable. The STA/LTA method is advantageous for the minimal prerequisites, the linear time complexity of O(n), and the ability to detect signals with distinctive properties. The effectiveness, however, depends on a substantial signal-to-noise ratio (SNR).

2.2. Template Matching

Template matching is a prevalent method in signal processing that quantifies the similarity of signals through cross-correlation. Conventional template-matching methods offer notable benefits, particularly when dealing with low signal-to-noise ratio (SNR) data [20]. This phenomenon is rooted in the physical nature of earthquakes, where seismic waves generated from a common source exhibit a consistent waveform when traveling through Earth. In this case, an initial event catalogue listing is essential and is regarded as a template. In addition, cross-correlation was performed between the new window and the template, while a threshold was applied to establish the minimal level of similarity required for identifying a match. This is commonly regarded as cross-correlation that has been adjusted to be amplitude invariant [10]. More specifically, the template matching is defined as follows:

$$r_i = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\ (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2\ (y_i - \bar{y})^2}} \tag{2}$$

where: $\bar{x}$: first signal; $\bar{y}$: second signal; r: correlation coefficient; i: current sample; n: total number of samples.

Template matching involves iterating through each template and comparing the template with the chosen waveform to identify a match. This is analogous to the template matching format employed in other disciplines. One clear example is in computer vision, where digital image model matching is utilized to locate cases of image models within a larger image. In these scenarios, a similarity threshold is employed to selectively exclude matches, particularly in digital image processing (computer vision). The temporal complexity of template matching is O(kn), where k represents the number of templates and n represents the number of windows in the signal being processed.

This technique exhibits higher sensitivity compared to STA/LTA and is capable of detecting noise in the signal. In addition, the technique contributes to identifying novel signals by correlating the signals with the equivalent signals in the template list. Furthermore, the detection requirements are more stringent compared to STA/LTA; therefore, a specific set of template catalogs is required. Furthermore, increasing the number of template catalogue lists improves the precision of model matching and the computational complexity.

2.3. Autocorrelation / Cross-Correlation

Autocorrelation compares a signal with a delayed version, a phenomenon referred to as auto-covariance in specific scientific disciplines. This technique is employed in signal processing to detect repetitive patterns in a signal, in addition, examines the complete waveform as a continuous signal separated into separate windows of a predetermined length. The autocorrelation is computed by correlating the seismic signature with the copy. Moreover, seismic interferometry creates images of underground structures by comparing seismic signals recorded by several receivers [31] using the above mentioned concept.

Furthermore, each window within the signal is correlated with every other window, reflecting the relative shift of the signal. In this case, various redundancy strategies are implementable to eliminate potential noise, such as setting a minimum correlation threshold or requiring a minimum number of matches [10]. As a further enhancement, normalized autocorrelation was utilized to ensure resilience against amplitude variations. In fact, autocorrelation is a statistical concept describing a correlation between a variable and the past value. To clarify, the equation is defined as follows:

$$r_i = \frac{\sum_{t=k+1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})} \tag{3}$$

where: $\bar{y}$: signal length; k: signal delay length.

The time complexity of autocorrelation is quadratic, denoted as O(n2), where n is the length of the wave. Specifically, autocorrelation is highly effective in identifying earthquake signals by accurately and precisely detecting recurring signals. Unlike template matching, autocorrelation does not require a catalogue list. Autocorrelation, however, is commonly regarded as technical resources for detecting earthquakes at shorter intervals.

3. Related Works

In recent years, seismology research has been conducted to detect and classify volcanic eruptions utilizing seismic signal data by [8, 15, 16] comprehensively examining cutting-edge machine learning methods for analyzing volcanic seismic data. The area of study is divided into two stages: detecting and categorizing seismic signal data for volcanic eruptions. Despite the positive outcomes, identifying and categorizing volcanic seismic events that co-occur in continuous data remains particularly challenging and thus requires extensive effort. A related study by Coombs et al. (2018) [32] focuses on identifying and classifying volcanic eruptions by employing near and real-time data. The finding reports that the capacity in detecting is approximately 60 explosives. Despite the success, enhancing the precision of the alerts is required. In 2020, a large number of machine-learning models were developed to analyze raw seismic velocity data. The models divide the sliding time windows into categories to automatically extract data for volcanic eruption recognition and real-time forecasting [33]. While being exclusively applicable for short-term alerts and offering a minimum four-hour warning for events, the models are unsuitable for predicting long-term increased eruptions.

According to Manley et al. (2020) [34], machine-learning approaches effectively categorize seismic time series into eruptive and non-eruptive behavior patterns. The overall state of a volcano is classified through single-station seismic data by assembling a model. However, analyzing a more extensive and varied dataset is necessary to ascertain whether these crucial characteristics are present across the overall volcanoes. Moreover, Saad & Chen (2021) [35] reported a study on applying a machine-learning algorithm to automatically recognize and classify the noise of events and earthquake signals. By identifying the arrival time of the P-wave using a number of recording data from different observation stations, the study achieved improved performance. Nevertheless, the model is limited to processing the data at a sampling rate of 100 Hz. On the other hand, Wiszniowski et al. (2021) [11] introduced improvements to machine learning models for interpreting seismic signal data by incorporating the polarization analyzer feature. This enhancement was explicitly applied to regions characterized by moderate seismic activity. In this case, precise seismic phase detection and identification are required for detecting seismic events and calculating parameters. Despite this, a significant drawback persists in the implementation of event detection methods that rely on manual data processing. The SLRNN demonstrates the ability to detect low-strength events impractical for manual analysis. This is primarily led by the larger ratio of seismic noise to the signal at most stations.

As a further point, Mandita et al. (2024) [36] combined STA/LTA and machine learning for detecting and classifying seismic signals. In the study, three distinct ML algorithms—Classis, Vanilla, and BiLSTM—were employed for detection and classification, combined with STA/LTA. The findings suggest that a combination of STA/LTA and ML provides an accuracy of around 0.70 and 0.80 in terms of detecting and classifying seismic events. Despite the ability of the model to detect and classify, challenges persist in terms of implementing datasets from different mountains and adding information related to the eruption status of volcanoes.

Furthermore, a number of ML algorithms have been employed to detect volcanic eruptions. Each approach demonstrates varying levels of precision when applied in seismology, such as in the domains of detecting and categorizing seismic signals and in the annotation and evaluation of both annotated and unannotated data, as reported by Mustafa et al. [37]. The accuracy is nearly identical when utilized in detecting volcanic eruptions. To expand on this, Sandhya et al. (2023) [38] predicted the magnitude of earthquakes utilizing data from the Horn of Africa. The study employs LSTM and BiLSTM models to predict earthquakes' magnitude. The objective is to perform a multivariate time-series regression to predict earthquakes with magnitudes of three or higher for the next three months. In practice, the outcomes and outputs acquired from long-term memory were compared.

A number of automated defect detection methods have been developed to enhance productivity and minimize time usage, among which deep-learning-based systems have demonstrated high efficiency [39]. The proposed technique was implemented in two phases: training and prediction. During the training phase, a Convolutional Neural Network (CNN) model was trained, incorporating actual data extracted from seven annotated seismic volumes. Each data point in these volumes was labeled, indicating the probability of faults. During the prediction stage, the trained network worked to compute the probability of faults at each location within the new seismic picture volumes. Despite this, further study is required to determine the effectivity of the trained CNN model when applied in different input samples.

Along those lines, the study provides a comprehensive assessment of machine learning (ML) applications in wide areas of earthquake seismology [13]. Specifically, ML is employed to create earthquake catalogs, analyze seismic activity, predict ground motion, and utilize geodetic data. Machine learning technologies have advanced significantly in these sectors. However, distinct problems require solutions. For instance, disparities in natural datasets pose a challenge in numerous scenarios, potentially leading to inaccurate assessments or misrepresentations. A number of unresolved issues in earthquake seismology have been addressed efficiently using machine learning (ML). Moreover, the implementation of ML broadens and enhances understanding in this area.

Adding to the above idea, the study offers a thorough evaluation by identifying and classifying earthquakes by utilizing KNN and SVM algorithms through the provided seismic signal data provided by Ozkaya et al. (2024) [29]. In this case, the dataset—noise, P-waves, and S-waves—was employed to characterize earthquakes. In addition, seven vector features were implemented as inputs for classifying the seismic signals using the KNN and SVM algorithms. The study successfully achieved 90% accuracy. To advance the finding, further testing is required with a more extensive dataset of seismic signals.

Despite the notable advancements, a number of limitations persist. In addition, a significant number of models have been tested on limited or specific datasets, thereby impeding an accurate assessment on generalizability across different volcanic and seismic regions. Furthermore, persistent noise in seismic data continues to obstruct, hindering accurate classification, particularly for low-strength events. While certain models excel at providing short-term alerts, further development is required to enhance long-term prediction capabilities. Moreover, a number of methods rely on manual processing, diminishing the efficiency for real-time applications. To resolve this, developing standardized datasets, generating noise-robust models, and expanding research into long-term forecasting methods are essential. Additionally, as machine learning is integrated into seismology, further innovation is indispensable to enhance understanding and mitigate seismic hazards.

4. Material and Methods

4.1. Data Analysis

The seismic signal data utilized in the study were pre-processed to eliminate poor signal quality from the seismic signal database. To identify any discrepancies within the seismic signal database, a thorough re-evaluation is essential, conducted with the assistance of experts in seismology. A number of observation stations, however, are limited to effectively observe phenomena related to volcanic activity. Another point to consider is inspecting an event to assure the appropriateness to be visually labeled. This task is more accessible when signal quality outweighs the surrounding noise level. Seismic signal categorization involves manually dividing the information into smaller segments of varying durations to identify underlying patterns. Following the extraction, each segment is categorized into a specific class based on the characteristics of the underlying physical event (reference class). In this study, the seismic signal data contained information related to volcanic activity, called occurrences.

On top of that, feature extraction describes the process of obtaining information from a dataset. In fact, analyzing seismic signal data in pattern-recognition systems is a crucial stage. The primary objective is to offer significant characteristics for the discerning procedure for seismic signals. During the feature extraction process, signal parameters are computed from the raw data by incorporating valuable information to distinguish between different classes of seismic signals. In this study, volcanic activity data were collected from observational locations near the volcano. Moreover, the primary data were collected from the Centre for Research and Development of Geological Disaster Technology (BPPTKG) at the Mount Merapi observation station in Yogyakarta. Specifically, the data were collected from multiple observation points over a specific time frame.

Located in the provinces of Yogyakarta, Central Java, Indonesia, Mount Merapi has been selected as the subject of the study due to the unique nature of each eruption, wherein seismic signal data varies across different eruptions. Accordingly, the study is focused on analyzing the seismic event data set of Mount Merapi in Indonesia. On the other hand, Mount Merapi has exhibited considerable volcanic activities in recent decades, as indicated by a number of eruptions throughout the current decade. Following this, seismic waveform data were collected from observation stations within a specified period.

In addition, seismic signal data related to the activity were collected from observation stations surrounding Mount Merapi throughout a specific timeframe, under particular frequencies of 0.5 Hertz to 50 Hz. To illustrate more clearly, Figure 3 displays the data obtained from Mount Merapi's activities in the form of connected photos, serving as the subjects of this study.

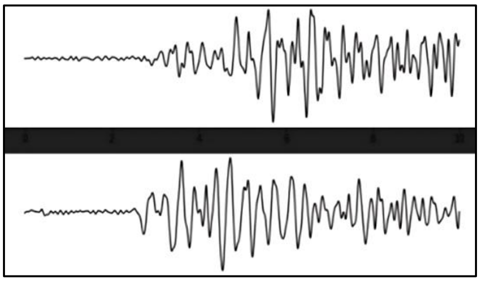

Figure 3. Data Seismic Events

These events are divided into a large number of classes based on wave patterns and spectral characteristics, as presented in Tables 1 and 2. Information was collected from Mount Merapi, between 2019 to 2021. The data used in the tests were primers obtained from a monitoring station near Mount Merapi. The seismic event data was examined prior to categorization according to data type with the assistance of domain experts. Data were collected from Mount Merapi, and subsequently categorized into eight and four distinct classifications of the seismic signals. A total of approximately 5000 to 10000 seismic event data were successfully gathered, encompassing diverse indications. Following this, the seismic event data were classified into eight classes, as listed in Table 1.

Table 1. The Classification of Eight Class-Seismic Signal Type

No	Seismic Signal Type
1.	AP
2.	DG
3.	Low Frequency
4.	Multiple Phase
5.	Rockfalls
6.	Tremor
7.	VT-A
8.	VT-B

Table 2 presents the seismic event data classified into four classes.

Table 2. The Classification of Four Class- Seismic Signal Type

No	Seismic Signal Type
1.	DG
2.	MP
3.	Rockfalls
4.	VT-B

4.2. Data Preprocessing

Preprocessing the seismic signal data is essential in eliminating low-quality signals from the seismic signal database and ensuring the accuracy and reliability of the research. Another point to consider is, identifying discrepancies in a seismic signal database requires professionals with a specialized background in seismology. Additionally, understanding that certain observation stations work effectively only when observing volcanic activity is essential. Moreover, ensuring precision requires a visual inspection to assure the accuracy of the event labeling. This procedure is more feasible when the signal quality surpasses the noise level of the surrounding seismic signal. Seismic signal data pattern categorization entails partitioning the dataset into smaller segments to subsequently be categorized into specific classes based on the underlying physical event. Regarding seismic signal data, the assessment of volcanic activity relies on the analysis of waves and spectrum, with each segment being categorized according to the corresponding reference class.

This study focuses on collecting seismic data by categorizing the events as presented in Table 1. The seismic data were sampled at frequencies of 0.5 Hz to 100 Hz, incorporating measurements from the time waveform, spectrum, and cepstrum. Afterwards, the data were filtered using a Butterworth bandpass filter under the frequency of 1–25 Hz. Notably, a sliding window method is commonly applied in seismic data processing to analyze seismic signals in specific time segments while reducing and differentiating noise in seismic signals. In addition, a Butterworth filter is a frequency filter employed in signal processing, including seismic signals. This filter is designed to provide a smooth frequency response in the pass-band (allowed frequency band) and roll-off (reduction of amplitude outside the band). On the other hand, a sliding window is a flexible tool that provides a dynamic analysis on seismic signals, enhancing the precision of detection and data processing. In this case, the window size and stride are typically adjusted based on the nature of the data and the purpose of the analysis. Furthermore, the features listed in Table 1 were obtained from multiple mountain observation stations that continuously monitored volcanic activity. For a more detailed illustration, Figure 4 depicts the pre-processing and feature extraction processes.

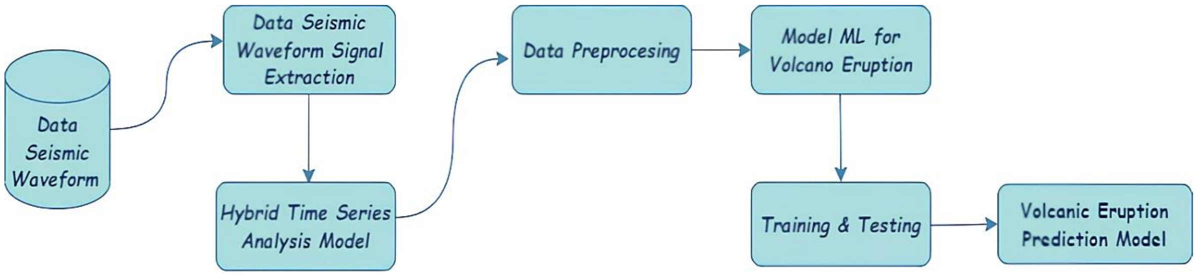

Figure 4. Extraction Features and Data Preprocessing

Figure 4 displays the seismic signal feature extraction and data preprocessing, commencing with seismic waveform data. In the initial phase, the data was fed into a bandpass filter utilizing Butterworth and sliding window techniques. This step is essential in extracting relevant information and distinguishing between noise and genuine or spurious events during the seismic signal detection stage. In this case, a hybrid time series analysis model was utilized to classify the seismic waves according to the labeled seismic event classes. Afterward, the pre-processed data was fed into further data preprocessing prior to applying machine learning techniques for further processing.

During the preprocessing stage of developing a model for predicting volcanic eruptions, the data set was systematically divided into training and testing data. The results of training and testing were subsequently utilized to form a hybrid time series and ML model to detect and classify seismic signals, in addition to predicting seismic signals, as well as the status of the anticipated type of eruption.

4.3. Volcano Activity Level

The Indonesian government, through the National Disaster Management Agency (BNPB), has developed volcano status levels as a critical component of mitigation plans. The classification is based on the severity and potential high-risk impacts of volcanic activity [40]. Moreover, the level of volcanic activity is categorized into the following:

- Normal (Level 1), assigned to volcanoes with inactive magma. Moreover, normal status indicates a volcano with essential volcanic activity.
- Waspada (Level 2), where volcanic activity exhibits an observable increase, detected by abnormal visual or seismic observations, changes in magma activity, hydrothermal increases, and tectonic events.
- Siaga (Level 3), where an eruption potentially occurs; nevertheless, the outcome is indeterminate. In this level, observational data indicate an increase in seismic and volcanic monitoring, in addition to visual and non-visual changes in the volcanic crater activity.
- Awas (Level 4), indicating that a volcanic eruption is imminent or actively occurring. At this level, the alert status alarms a disastrous condition.

4.4. Hybrid Time Series Method

Mandita et al. (2024) [36] have successfully developed an STA/LTA model combined with ML, including Classic, Vanilla, and BiLSTM, which effectively detected and classified a seismic signal. While predicting the status of a volcanic eruption, the accuracy level remained approximately 70-80. In contrast, the model built in this study integrates the time series method with two distinct models: the Proposed Method Hybrid STA/LTA & Template Matching and the Proposed Method Hybrid STA/LTA & Autocorrelation collaborated with the ML algorithm for detecting and classifying seismic signals, as well as predicting the status of volcanic eruptions. While both studies utilize the time series method to detect and classify seismic signals, the previous experiment employed a one-time series model, whereas this study combined several time series models. The novelty of this research involves a hybrid time series model that merges a combination of several time series methods. Additionally, the built model, combined with the ML algorithm, is proficient in predicting volcanic eruption status.

Correlating to this, a hybrid time-series approach for volcanic prediction has been proposed. This methodology integrates algorithms, machine learning, and time-series analysis. A hybrid time series combines a large number of components or models to capture different patterns or characteristics of the data. In addition, two types of hybrid time series built in this experiment comprise STA/LTA and template matching, as well as STA/LTA and autocorrelation. Specifically, the hybrid time series method is defined as follows:

$$r_i = \frac{STA_i}{LTA_i} \frac{\sum_{i=1}^{n}(x_i - \bar{x})\,(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2\,(y_i - \bar{y})^2}} \tag{4}$$

$$r_i = \frac{STA_i}{LTA_i} \frac{\sum_{t=k+1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})} \tag{5}$$

where: r_i: ratio of STA/LTA & template matching; and the ratio of STA/LTA & autocorrelation; $\bar{x}$: first signal; $\bar{y}$: second signal; n: total number of samples; k: signal delay length.

Equations 4 and 5 present a new method for categorizing and identifying volcanic eruptions using time-series techniques and machine learning. The proposed method integrates a time-series algorithm to classify and identify signal seismic occurrences. In this case, two approaches were proposed: short-term average/long-term average (STA/LTA) and template matching, as well as a combination of STA/LTA with autocorrelation combined with ML algorithms—Classic, Vanilla, and BiLSTM. Subsequently, the ML algorithms were compared to determine the model that exhibited the highest accuracy in analyzing a seismic signal. To clearly illustrate, the proposed method for volcanic eruptions is presented in Figure 5, depicting the newly proposed method for volcanic eruptions, commencing with detecting seismic event data to detect actual or false seismic events from the volcano, using a hybrid time series algorithm to process the seismic data during the classification stage.

Equally important, two types of hybrid time series algorithms were implemented, including STA/LTA with template matching and STA/LTA with autocorrelation. The objective is to detect and classify seismic signal data. A predictive model using hybrid time series and ML collaboration was built to predict volcanic eruptions. Additionally, the study employed a number of machine learning models—Classic, Vanilla, and BiLSTM—to analyze seismic signal data. In this case, the ML algorithms were compared to determine the model that achieved the highest accuracy during the seismic signal analysis process. The subsequent phase involved training and testing the model to forecast volcanic eruptions. Upon completing the overall required operations, the highest-performing ML model for volcanic prediction was performed and validated. Lastly, the final stage is expected to produce a validated model that accurately predicts daily seismic events with higher precision.

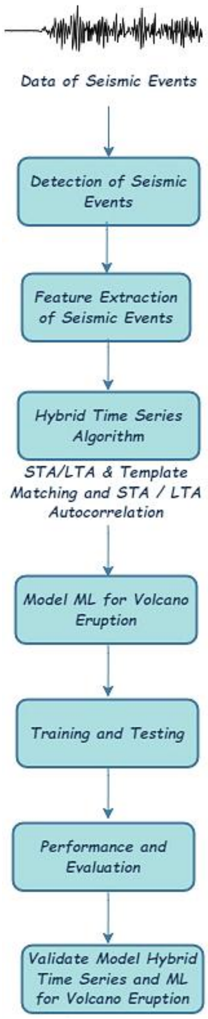

Figure 5. The Proposed Method

5. Results and Discussions

This chapter discusses the findings of the study in detecting true or false seismic events. The seismic event data were analyzed based on detection and classification to identify actual or false events. As previously outlined, the seismic events were divided into eight classes and four classes of seismic event types. At this point, data identification and feature extraction were performed. The data were analyzed to determine the possibility of generating the desired machine-learning model. Moreover, the data were purified to build the highest-performing machine-learning model for detecting and classifying seismic occurrences.

Furthermore, the inputs for the ML model were divided into eight and four inputs, respectively. In particular, the eight inputs are defined as follows:

$$x_t = \begin{matrix} AG \\ DG \\ LF \\ MP \\ RF \\ TR \\ VT\text{-}A \\ VT\text{-}B \end{matrix} \quad (6)$$

The definitions of the four inputs are as follows:

$$x_t = \begin{matrix} DG \\ MP \\ RF \\ VT\text{-}B \end{matrix} \quad (7)$$

Data analysis was conducted on the data classification outcomes involving seismology experts. During the iterative data testing, performing data validation is essential to obtain the ideal outcomes from the constructed model. The process involved training and testing the model, focusing on detecting and classifying the seismic occurrences. Moreover, three models were employed in detecting and classifying the seismic data.

5.1. Support Vector Machine (SVM)

The built model incorporates SVM with the time series method, specifically the STA/LTA method, to detect and classify seismic signals. The model was employed to analyze seismic signals and predict volcanic eruptions. Moreover, two classes of seismic events—four classes and eight classes of seismic events—were employed, as presented in Tables 1 and 2, the results of which are as follows:

Figure 6 represents the results of the time series method. In the process, SVM was utilized in detecting and classifying seismic events, as well as predicting volcanic eruption status. In this study, three Support Vector Machine (SVM) models—linear, polynomial, and RBF—were employed to identify and categorize seismic occurrences and distinguish between genuine and spurious events for forecasting volcano eruptions. As a result, the model exhibited the lowest linear accuracy of 0.88 for the input of four classes, compared to the polynomial accuracy of 0.9. Additionally, the polynomial models outperformed the RBF model, achieving an accuracy of 0.88, similar to the accuracy exhibited by the linear and RBF models.

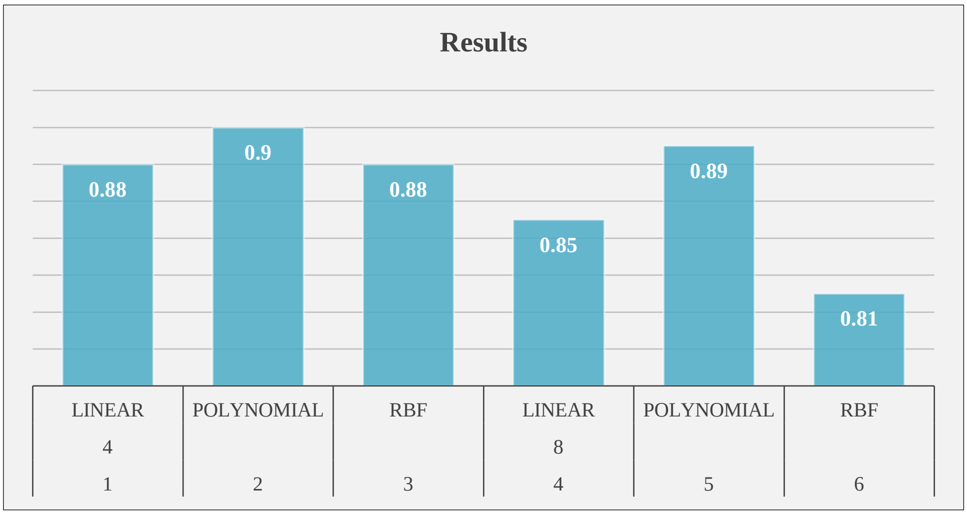

Figure 6. The results of the SVM Model with four classes and eight classes input

Parallel to this, the employment of eight classes as input generated an accuracy of approximately 0.81 to 0.89. The linear model achieved an accuracy of 0.85, surpassing the RBF model's accuracy of 0.81 in classifying and detecting seismic events. The accuracy of the linear model, however, was lower than that of the polynomial model (0.89). Both classes with four or eight inputs signify that the polynomial model generates superior outcomes, with an accuracy of 0.9 for the four input classes and 0.89 for the eight input classes. The model, therefore, was employed to detect and classify seismic events and predict volcanic eruptions.

5.2. K Nearest Neighbors (KNN)

In practice, this study employed the KNN to classify and detect signal seismic events using four and eight classes, respectively. The results of which are presented as follows:

Expanding further, Figure 7 illustrates the experimental results for eight classes using five K. For accuracy, K = 1 achieved the highest-performing value at 0.84, outperforming other results. For other additional K, results range from 0.7 to 0.73, with K = 3 achieving a value of 0.73, K = 5 achieving a value of 0.74 and K = 9 achieving a value of 0.71, with poor accuracy observed at K = 7, with a value of 0.7. In terms of precision, the highest-performing results were exhibited at K = 1 with a value of 0.85. While other K achieved values between 0.66 to 0.75, where K = 3 achieved a value of 0.75, with K = 5 achieving a value of 0.71, K = 7 achieved a value of 0.66 and K = 9 achieved a value of 0.67. This suggests that K = 7 performs the lowest results among all. For recall, the highest-performing results are exhibited at K = 1 with a value of 0.84. For other K, the values range between 0.7 to 0.74, with the lowest result at K = 7, with a value of 0.7 compared to other K values. For the F1 score, the highest-performing result was obtained with several 0.84 at K = 1. For other K results, K = 3 at 0.73, whereas for K = 5, the value was 0.72. For other K, K = 7 achieving a value of 0.68, and finally K = 9, with a value of 0.69. The results demonstrate that K = 7 with a value of 0.66 exhibits the lowest performance among all.

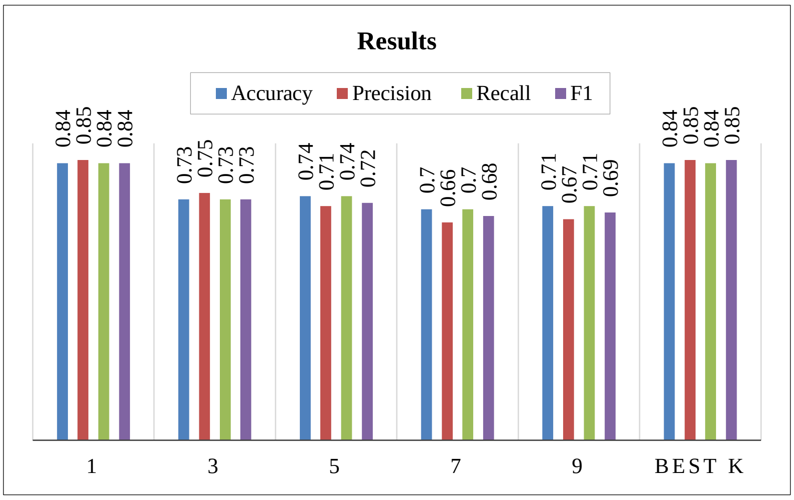

Figure 7. The results with eight classes input

Furthermore, Figure 8 presents the results for the four classes, evaluated across different five K experiments. Among these, a single value emerged as the highest-performing K experiment, in which K = 1 achieved the highest-performing results, compared to other values. For the accuracy, K = 1 achieved 0.87, while the results of other K ranged from 0.77 to 0.78 for K with values 5, 7, and 9, and K = 3 achieved the lowest output. For precision, K = 1 achieved the highest-performing value compared to other K values. K = 5, K = 7, and K = 9 achieved similar value of 0.76. For accuracy, K = 3 achieved the lowest value compared to other K values, with a value of 0.75. In recall, K = 1 with a value of 0.87 emerged as the highest-performing value. K = 5 and K = 9 achieved a value of 0.77, slightly lower than the results at K = 7 with a value of 0.78. Meanwhile, the output with the lowest value at K = 3 achieved a value of 0.75. Regarding the F1 score value, K = 1 achieved the highest-performing result with a value of 0.87. Additionally, K=5, K=7 and K=9 achieved similar value of 0.76. In contrast, K=3, with a value of 0.75 achieved the lowest result.

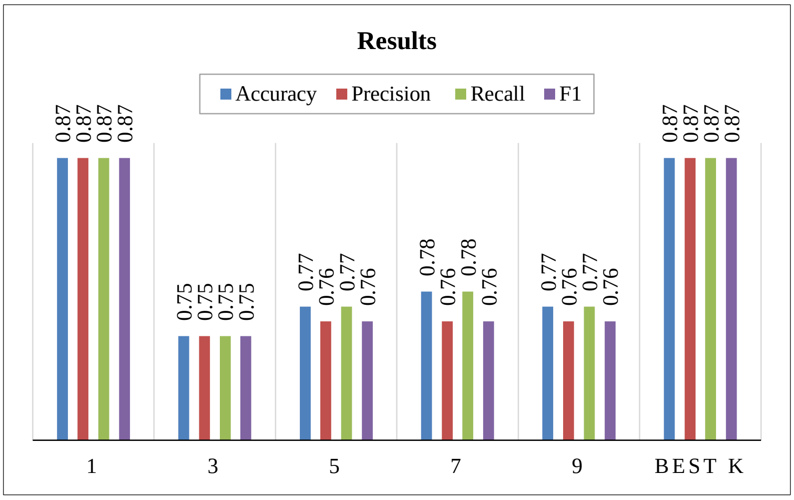

Figure 8. The results with four classes input

5.3. Long Short-Term Memory (LSTM)

The LSTM employed to classify and detect seismic events, the results of which are presented in Tables 3 and 4.

Table 3. The accuracy results for eight classes

No.	Methods	Accuracy
1.	Classic LSTM	0.84
2.	Vanilla LSTM	0.85
3.	BiLSTM	0.87
4.	Proposed Method Hybrid STA/LTA & Template Matching and Classic	0.88
5.	Proposed Method Hybrid STA/LTA & Autocorrelation and Classic	0.89
6.	Proposed Method Hybrid STA/LTA & Template Matching and Vanilla	0.90
7.	Proposed Method Hybrid STA/LTA & Autocorrelation and Vanilla	0.91
8.	Proposed Method Hybrid STA/LTA & Template Matching and BiLSTM	0.93
9.	Proposed Method Hybrid STA/LTA & Autocorrelation and BiLSTM	0.93

Table 3 presents the accuracy results for eight classes of seismic event classification. Classic LSTM achieved an accuracy of 0.84, while Vanilla LSTM achieved a slightly different accuracy of 0.85, compared to BiLSTM, with an accuracy of 0.87. This reflects a difference of 0.02 points between Vanilla LSTM and BiLSTM. In practice, the proposed method involved three methods: Classic, Vanilla, and BiLSTM. The hybrid STA/LTA and Template Matching method with Classic LSTM, achieved an accuracy of 0.88, slightly different from the hybrid STA/LTA and Autocorrelation method with an accuracy of 0.89. Combined with Vanilla LSTM, the hybrid method achieved an accuracy of 0.89, while the STA/LTA & Template Matching and STA/LTA & Autocorrelation models achieved an accuracy of 0.90. Additionally, the hybrid methods with BiLSTM achieved an accuracy of 0.93 in terms of the proposed hybrid STA/LTA and Template Matching method, and the proposed hybrid STA/LTA and Autocorrelation method achieved similar accuracy. Moreover, the proposed method achieved similar accuracy in terms of the STA/LTA & Template Matching and STA/LTA & Autocorrelation with BiLSTM – with a value of 0.93. Notably, the method outperformed the other methods in terms of accuracy.

The results of the proposed method for the four classes are listed in Table 4.

Table 4. The accuracy results for four classes

No.	Methods	Accuracy
1.	Classic LSTM	0.85
2.	Vanilla LSTM	0.86
3.	BiLSTM	0.88
4.	Proposed Method Hybrid STA/LTA & Template Matching and Classic	0.89
5.	Proposed Method Hybrid STA/LTA & Autocorrelation and Classic	0.90
6.	Proposed Method Hybrid STA/LTA & Template Matching and Vanilla	0.91
7.	Proposed Method Hybrid STA/LTA & Autocorrelation and Vanilla	0.93
8.	Proposed Method Hybrid STA/LTA & Template Matching and BiLSTM	0.95
9.	Proposed Method Hybrid STA/LTA & Autocorrelation and BiLSTM	0.95

Table 4 lists the accuracy results for the four seismic event classification classes, presenting that Classic LSTM achieved an accuracy of 0.85, slightly lower than the Vanilla LSTM model's accuracy of 0.86. This is particularly different from Classic LSTM, with an accuracy of 0.01. Likewise, the results of the Vanilla LSTM were lower than BiLSTM, with an accuracy of 0.88, representing an enhancement of 0.02 over the Vanilla LSTM. Additionally, the proposed method for STA/LTA & Template Matching and Classic LSTM achieved an accuracy of 0.89, slightly different from the STA/LTA & Autocorrelation and Classic LSTM methods with an accuracy of 0.90. Similarly, the proposed method with Vanilla LSTM for STA/LTA & Template Matching achieved an accuracy level of 0.91, with a slight difference for the STA/LTA & Autocorrelation, with an accuracy of 0.93. Moreover, the last proposed method, hybrid STA/LTA & Template Matching and hybrid STA/LTA & Autocorrelation, achieved a similar level of accuracy. In addition, the proposed methods with hybrid STA/LTA & Template Matching and hybrid STA/LTA & Autocorrelation achieved an accuracy of 0.95. Notably, the results of the proposed hybrid time series and BiLSTM method outperform the other methods.

5.4. Comparison of Others Model with the Proposed Model

The accuracies of the comparison and proposed model are presented in Table 5.

Table 5. The comparison results of the utilized model and the proposed model for eight classes

No.	Methods	Accuracy
1.	SVM	0.90
2.	KNN	0.87
3.	Classic LSTM	0.84
4.	Vanilla LSTM	0.85
5.	BiLSTM	0.87
6.	Proposed Method Hybrid STA/LTA & Template Matching and Classic	0.88
7.	Proposed Method Hybrid STA/LTA & Autocorrelation and Classic	0.89
8.	Proposed Method Hybrid STA/LTA & Template Matching and Vanilla	0.90
9.	Proposed Method Hybrid STA/LTA & Autocorrelation and Vanilla	0.91
10.	Proposed Method Hybrid STA/LTA & Template Matching and BiLSTM	0.93
11.	Proposed Method Hybrid STA/LTA & Autocorrelation and BiLSTM	0.93

Table 5 represents the comparative accuracy between the proposed approach and the other methods. In detail, in terms of detecting and classifying seismic event signals, SVM and KNN achieved an accuracy of 0.90 and 0.87, respectively. In this task, SVM outperforms KNN. The Classic LSTM model achieved a performance score of 0.84, whereas the Vanilla LSTM model achieved a slightly higher score of 0.85. Comparatively, the BiLSTM model surpassed the Vanilla LSTM and BiLSTM models with a performance score of 0.87, demonstrating a 0.02-point disparity. However, when comparing the accuracy of the SVM with the proposed technique, the SVM achieved lower accuracy than that of the proposed method. The proposed method demonstrates higher accuracy, achieving a precision of 0.93, surpassing the SVM and KNN. Compared with the SVM, KNN, Classic LSTM, Vanilla LSTM, BiLSTM, and hybrid method with Classic LSTM or Vanilla LSTM, the proposed method—the hybrid STA/LTA & Template Matching and STA/LTA & Autocorrelation with BiLSTM—achieved a superior accuracy of 0.93. This represents an accuracy improvement from 0.06 to 0.09 in the proposed model.

Table 6 illustrates the comparison accuracy between the proposed method with SVM and KNN, achieving an accuracy of 0.90 and 0.85, respectively. This signifies that SVM outperforms KNN for detecting and classifying seismic event signals. While the Classic LSTM model achieved an accuracy of 0.85, the Vanilla LSTM model achieved a slightly lower accuracy of 0.86, resulting in a difference of 0.01 compared to the Classic LSTM model. Moreover, the Vanilla LSTM achieved poorer results compared to the BiLSTM, which achieved an accuracy of 0.88. Meanwhile, when compared to the proposed method, the accuracy of SVM was lower than that of the proposed method. The proposed method achieved an accuracy of 0.95, surpassing the SVM and KNN. Furthermore, when evaluated against the SVM, KNN, Classic LSTM, Vanilla LSTM, BiLSTM, and hybrid method with Classic LSTM or Vanilla LSTM—hybrid STA/LTA & Template Matching and STA/LTA & Autocorrelation with BiLSTM—the proposed method achieved the highest accuracy of 0.95. This represents an improvement in the proposed model's accuracy from 0.05 to 0.1.

Table 6. The comparison results of the utilized model and the proposed model for four classes

No.	Methods	Accuracy
1.	SVM	0.90
2.	KNN	0.85
3.	Classic LSTM	0.85
4.	Vanilla LSTM	0.86
5.	BiLSTM	0.88
6.	Proposed Method Hybrid STA/LTA & Template Matching and Classic	0.89
7.	Proposed Method Hybrid STA/LTA & Autocorrelation and Classic	0.90
8.	Proposed Method Hybrid STA/LTA & Template Matching and Vanilla	0.91
9.	Proposed Method Hybrid STA/LTA & Autocorrelation and Vanilla	0,93
10.	Proposed Method Hybrid STA/LTA & Template Matching	0.95
11.	Proposed Method Hybrid STA/LTA & Autocorrelation	0.95

Table 7 presents an overview of the accuracy, precision, recall, and F1 score for eight seismic event classes, summarizing the performance of a number of machine learning models and methods employed for classification tasks. These metrics offer a comprehensive assessment of the model's effectiveness.

Table 7. The Accuracy, Precision, Recall, and F1 Score Model LSTM for Eight Classes of Seismic Events

No.	Methods	Accuracy	Precision	Recall	F1 Score
1.	Classic LSTM	0.84	0.85	0.83	0.839
2.	Vanilla LSTM	0.85	0.86	0.84	0.849
3.	BiLSTM	0.87	0.88	0.86	0.869
4.	Proposed Method Hybrid STA/LTA & Template Matching and Classic	0.88	0.89	0.87	0.879
5.	Proposed Method Hybrid STA/LTA & Autocorrelation and Classic	0.89	0.90	0.88	0.889
6.	Proposed Method Hybrid STA/LTA & Template Matching and Vanilla	0.90	0.91	0.89	0.899
7.	Proposed Method Hybrid STA/LTA & Autocorrelation and Vanilla	0.91	0.92	0.90	0.909
8.	Proposed Method Hybrid STA/LTA & Template Matching and BiLSTM	0.93	0.94	0.92	0.929
9.	Proposed Method Hybrid STA/LTA & Autocorrelation and BiLSTM	0.93	0.94	0.92	0.929

The Classic LSTM model achieved an accuracy of 84%, with precision, recall, and F1 scores of 85%, 83%, and 0.839, respectively, reflecting a balanced performance in identifying relevant instances with reliable predictions. The Vanilla LSTM model achieved a slight improvement of 85% accuracy, 86% precision, 84% recall, and an F1 score of 0.849. The BiLSTM model further enhanced these metrics, with an accuracy of 87%, precision of 88%, recall of 86%, and an F1 score of 0.869, therefore, the incorporation of the bidirectional layers enabled the model to capture contextual dependencies with greater efficacy.

Moreover, significant advancements are presented through the Proposed Hybrid Methods, which combine Hybrid STA/LTA & Template Matching or Autocorrelation techniques with LSTM-based models. These hybrid approaches consistently improved the accuracy, precision, recall, and F1 scores. Notably, the Hybrid STA/LTA & Template Matching with Classic LSTM achieved 88% accuracy, 89% precision, 87% recall, and an F1 score of 0.879. In addition, the Autocorrelation with Classic LSTM hybrid improved the accuracy to 89% and F1 score to 0.889.

The upward trend persisted as Vanilla LSTM served as the baseline in these hybrid methods. The Template Matching and Vanilla LSTM hybrid achieved 90% accuracy, while the Autocorrelation and Vanilla LSTM hybrid achieved 91% accuracy and an F1 score of 0.909. The highest performing results were obtained with hybrid methods incorporating BiLSTM, where both Template Matching and Autocorrelation achieved 93% accuracy, with precision of 94%, recall of 92%, and an F1 score of 0.929.

These findings highlight the significant performance enhancements exhibited by the proposed hybrid methods. The improvements in precision and recall demonstrate the ability to reliably identify relevant instances while minimizing errors, establishing the methods to be highly suitable for applications requiring robust predictive accuracy.

Table 8 presents the accuracy, precision, recall, and F1 score for eight classes of seismic events, demonstrating the efficacy of various machine learning models and hybrid techniques in classification tasks. These metrics—Accuracy, Precision, Recall, and F1 Score—provide a comprehensive evaluation of model reliability, efficacy in detecting relevant events, and the balance between precision and recall.

Table 8. The Accuracy, Precision, Recall, and F1 Score Model LSTM for Four Classes Seismic Events

No.	Methods	Accuracy	Precision	Recall	F1 Score
1.	Classic LSTM	0.85	0.86	0.84	0.849
2.	Vanilla LSTM	0.86	0.87	0.85	0.859
3.	BiLSTM	0.88	0.89	0.87	0.879
4.	Proposed Method Hybrid STA/LTA & Template Matching and Classic	0.89	0.90	0.88	0.889
5.	Proposed Method Hybrid STA/LTA & Autocorrelation and Classic	0.90	0.91	0.89	0.899
6.	Proposed Method Hybrid STA/LTA & Template Matching and Vanilla	0.91	0.92	0.90	0.909
7.	Proposed Method Hybrid STA/LTA & Autocorrelation and Vanilla	0.93	0.94	0.92	0.929
8.	Proposed Method Hybrid STA/LTA & Template Matching and BiLSTM	0.95	0.96	0.94	0.949
9.	Proposed Method Hybrid STA/LTA & Autocorrelation and BiLSTM	0.95	0.96	0.94	0.949

As illustrated in the table, the Classic LSTM model achieved an accuracy of 85%, with a precision of 86% and a recall of 84% with an F1 score of 0.849, establishing a solid baseline for comparison. The Vanilla LSTM slightly outperformed this, with an accuracy of 86%, precision of 87%, recall of 85%, and an F1 score of 0.859. The BiLSTM model further improved these results, achieving 88% accuracy, 89% precision, 87% recall, and an F1 score of 0.879, benefiting from the bidirectional architecture, enhancing the model's ability to process contextual information.

Moreover, the proposed hybrid methods demonstrated significant performance by combining STA/LTA techniques (Template Matching or Autocorrelation) with LSTM-based models. Specifically, when paired with Classic LSTM, the hybrid techniques achieved 89% and 90% accuracy, with F1 scores of 0.889 and 0.899, respectively, reflecting improved accuracy in predicting positive cases while maintaining intense precision and recall. Higher performance was observed in the use of Vanilla LSTM as the baseline in these hybrid approaches. The Hybrid STA/LTA & Template Matching with Vanilla LSTM achieved 91% accuracy, while the Hybrid STA/LTA & Autocorrelation with Vanilla LSTM achieved 93% accuracy, delivering F1 scores above 0.90. This underscores the significant impact of the hybrid approach in enhancing model performance.

As equally important, the most remarkable results were observed in hybrid methods utilizing BiLSTM. Hybrid STA/LTA & Template Matching with BiLSTM and Hybrid STA/LTA & Autocorrelation with BiLSTM, achieved 95% accuracy, with a precision of 96%, recall of 94%, and an F1 score of 0.949. The results highlight the robustness and effectiveness of combining STA/LTA techniques with BiLSTM for advanced classification tasks.

Overall, the hybrid methods consistently enhanced the baseline models' performance, particularly when integrated with BiLSTM, establishing the methods as promising solutions for applications requiring high prediction accuracy and reliability.

Figure 9 illustrates the output of volcanic eruption prediction utilizing the proposed method, in which the volcanic eruption status is classified as “waspada”, and the type of seismic signal is identified as mp with a magnitude of 4.9.

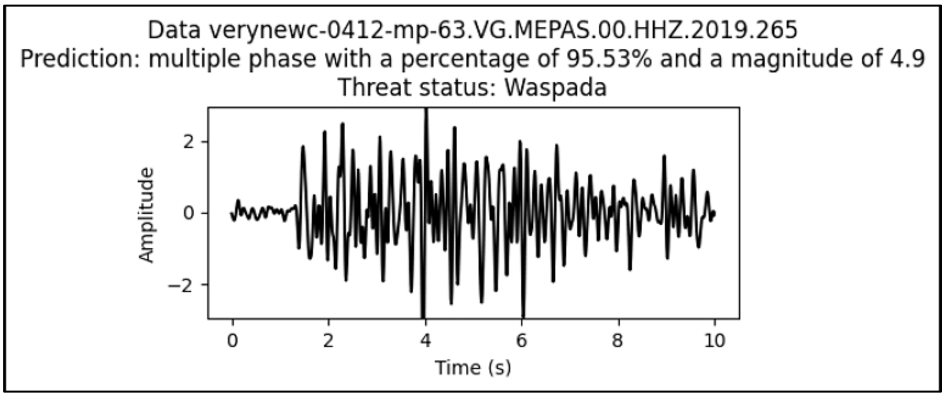

Figure 9. The results of Volcano Activity

Based on experimental analysis with eight and four classes, the proposed methods included hybrid STA/LTA & Template Matching and STA/LTA & Autocorrelation with BiLSTM, which were observed to be more accurate in detecting and classifying seismic events, as well as in predicting volcano eruptions. Notably, the proposed method achieved higher-performing results among all.

6. Conclusion

As detailed in the preceding sections, this study proposes a model for detecting, classifying, and predicting volcanic eruptions by employing a hybrid time series method and ML. Multiple models have been developed to compare the accuracy levels in analyzing a seismic signal with the proposed model. To compare the level of accuracy at the experimental stage, seismic events are divided into two classes: eight and four classes. The primary data is collected from one of the active volcanoes in Indonesia, Mount Merapi, which has been actively emitting lava in recent decades. Moreover, the study has been performed to detect and classify seismic events, as well as predict volcanic eruptions, exhibiting varying levels of accuracy.

Furthermore, the comparative analysis between the employed methods with the proposed method achieves an accuracy level between 0.84 to 0.89, whereas the proposed model achieves an accuracy level from 0.90 to 0.95. The proposed method provides higher accuracy than other methods, with an accuracy between 0.93 to 0.95 for the STA/LTA & Template Matching and STA/LTA & Autocorrelation with BiLSTM models—the proposed model. This signifies that the proposed method presents high-performing results compared to other methods. Correlating to this, future research involves tuning the model to detect and classify seismic event data by employing other datasets or data from other volcanoes, including a larger volume of datasets. Moreover, a classification of seismic events will be added to the proposed method's accuracy. Additionally, another future research will improve the status of volcanic eruptions for seismic events, thus producing more precise results.

7. Declarations

7.1. Author Contributions

Conceptualization, F.M., A.A., M.E.W., and W.S.; methodology, F.M. and W.S.; software, F.M.; validation, A.A.; formal analysis, F.M.; investigation, F.M. and M.E.W.; resources, F.M. and W.S.; data curation, F.M.; writing—original draft preparation, F.M.; writing—review and editing, F.M. and A.A.; visualization, F.M.; supervision, A.A.; project administration, A.A.; funding acquisition, A.A. All authors have read and agreed to the published version of the manuscript.

7.2. Data Availability Statement

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

7.3. Funding

The research is funded by the Doctoral Dissertation research schema for the 2022 budget year.

7.4. Institutional Review Board Statement

Not applicable.

7.5. Informed Consent Statement

Not applicable.

7.6. 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.

8. References

[1] Supriyo, B., Hidayat, S. S., Suharjono, A., Anif, M., & Koesuma, S. (2015). Design of real-time gas monitoring system based-on wireless sensor networks for Merapi volcano. 2014 1st International Conference on Information Technology, Computer, and Electrical Engineering: Green Technology and Its Applications for a Better Future, ICITACEE 2014 - Proceedings, 30–34. doi:10.1109/ICITACEE.2014.7065709.

[2] Ratdomopurbo, A., Beauducel, F., Subandriyo, J., Agung Nandaka, I. G. M., Newhall, C. G., Suharna, Sayudi, D. S., Suparwaka, H., & Sunarta. (2013). Overview of the 2006 eruption of Mt. Merapi. Journal of Volcanology and Geothermal Research, 261, 87–97. doi:10.1016/j.jvolgeores.2013.03.019.

[3] Sholihat, S. S., Indratno, S. W., & Mukhaiyar, U. (2021). The role of parameters in Bayesian Online Changepoint Detection: detecting early warning of mount Merapi eruptions. Heliyon, 7(7), e07482. doi:10.1016/j.heliyon.2021.e07482.

[4] Surono, Jousset, P., Pallister, J., Boichu, M., Buongiorno, M. F., Budisantoso, A., Costa, F., Andreastuti, S., Prata, F., Schneider, D., Clarisse, L., Humaida, H., Sumarti, S., Bignami, C., Griswold, J., Carn, S., Oppenheimer, C., & Lavigne, F. (2012). The 2010 explosive eruption of Java's Merapi volcano-A "100-year" event. Journal of Volcanology and Geothermal Research, 241–242, 121–135. doi:10.1016/j.jvolgeores.2012.06.018.

[5] Chan, H. P., Konstantinou, K. I., & Blackett, M. (2021). Spatio-temporal surface temperature variations detected by satellite thermal infrared images at Merapi volcano, Indonesia. Journal of Volcanology and Geothermal Research, 420. doi:10.1016/j.jvolgeores.2021.107405.

[6] Antriyandarti, E., Ferichani, M., & Ani, S. W. (2013). Sustainability of Post-Eruption Socio Economic Recovery for the Community on Mount Merapi Slope through Horticulture Agribusiness Region Development (Case Study in Boyolali District). Procedia Environmental Sciences, 17, 46–52. doi:10.1016/j.proenv.2013.02.010.

[7] Sunardi, S., Dede, M., Wibowo, S. B., Prasetyo, Y., Astari, A. J., Lukman, L., Lavigne, F., Gomez, C., Nurani, I. W., Sakai, Y., & Kamarudin, M. K. A. (2024). Preliminary assessment of river ecosystem services in the volcanic area of Mount Merapi, Indonesia. Aquatic Ecology, 58(3), 819–832. doi:10.1007/s10452-024-10107-4.

[8] Falcin, A., Métaxian, J. P., Mars, J., Stutzmann, É., Komorowski, J. C., Moretti, R., Malfante, M., Beauducel, F., Saurel, J. M., Dessert, C., Burtin, A., Ucciani, G., de Chabalier, J. B., & Lemarchand, A. (2021). A machine-learning approach for automatic classification of volcanic seismicity at La Soufrière Volcano, Guadeloupe. Journal of Volcanology and Geothermal Research, 411. doi:10.1016/j.jvolgeores.2020.107151.

[9] Espinosa-Ortega, T., Budi-Santoso, A., Sulistiyani, Win, N. T. Z., Widiwijayanti, C., & Costa, F. (2022). Probabilistic analysis to correlate seismic data with lava extrusion phases at Merapi volcano (Indonesia). Journal of Volcanology and Geothermal Research, 426. doi:10.1016/j.jvolgeores.2022.107537.

[10] Audretsch, J. (2020). Earthquake Detection using Deep Learning Based Approaches. KAUST Research Repository, KAUST: King Abdullah University of Science and Technology, Saudi Arabia. doi:10.25781/KAUST-52098.

[11] Wiszniowski, J., Plesiewicz, B., & Lizurek, G. (2021). Machine learning applied to anthropogenic seismic events detection in Lai Chau reservoir area, Vietnam. Computers and Geosciences, 146. doi:10.1016/j.cageo.2020.104628.

[12] Titos, M., Bueno, A., Garcia, L., & Benitez, C. (2018). A Deep Neural Networks Approach to Automatic Recognition Systems for Volcano-Seismic Events. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(5), 1533–1544. doi:10.1109/JSTARS.2018.2803198.

[13] Kubo, H., Naoi, M., & Kano, M. (2024). Recent advances in earthquake seismology using machine learning. Earth, Planets and Space, 76(1), 36. doi:10.1186/s40623-024-01982-0.

[14] Lara, F., Lara-Cueva, R., Larco, J. C., Carrera, E. V., & León, R. (2021). A deep learning approach for automatic recognition of seismo-volcanic events at the Cotopaxi volcano. Journal of Volcanology and Geothermal Research, 409, 107142. doi:10.1016/j.jvolgeores.2020.107142.

[15] 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), 10,645-10,658. doi:10.1029/2018JB015470.

[16] Titos, M., Bueno, A., García, L., Benítez, M. C., & Ibañez, J. (2019). Detection and Classification of Continuous Volcano-Seismic Signals with Recurrent Neural Networks. IEEE Transactions on Geoscience and Remote Sensing, 57(4), 1936–1948. doi:10.1109/TGRS.2018.2870202.

[17] Vaezi, Y., & van der Baan, M. (2015). Comparison of the STA/LTA and power spectral density methods for microseismic event detection. Geophysical Journal International, 203(3), 1896–1908. doi:10.1093/gji/ggv419.

[18] Qu, S., Guan, Z., Verschuur, E., & Chen, Y. (2020). Automatic high-resolution microseismic event detection via supervised machine learning. Geophysical Journal International, 222(3), 1881–1895. doi:10.1093/gji/ggaa193.

[19] Pantobe, L., Burtin, A., Chanard, K., & Komorowski, J. C. (2024). Evolution of shallow volcanic seismicity in the hydrothermal system of La Soufrière de Guadeloupe following the April 2018 MLV 4.1 earthquake. Journal of Volcanology and Geothermal Research, 447. doi:10.1016/j.jvolgeores.2023.107989.

[20] Ma, H., Chu, R., Sheng, M., & Yu, Z. (2020). Template matching for simple waveforms with low signal-to-noise ratio and its application to icequake detection. Earthquake Science, 33(5–6), 256–263. doi:10.29382/eqs-2020-0256-01.

[21] Yang, D. H., Zhou, X., Wang, X. Y., & Huang, J. P. (2021). Mirco-earthquake source depth detection using machine learning techniques. Information Sciences, 544, 325–342. doi:10.1016/j.ins.2020.07.045.

[22] Ikeda, H., & Takagi, R. (2019). Coseismic changes in subsurface structure associated with the 2018 Hokkaido Eastern Iburi Earthquake detected using autocorrelation analysis of ambient seismic noise. Earth, Planets and Space, 71(1), 1-11. doi:10.1186/s40623-019-1051-5.

[23] Nurtas, M., Zhantaev, Z., & Altaibek, A. (2024). Earthquake time-series forecast in Kazakhstan territory: Forecasting accuracy with SARIMAX. Procedia Computer Science, 231, 353–358. doi:10.1016/j.procs.2023.12.216.

[24] Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., & Gerstoft, P. (2019). Machine learning in seismology: Turning data into insights. Seismological Research Letters, 90(1), 3–14. doi:10.1785/0220180259.

[25] Liu, H., Song, J., & Li, S. (2022). Seismic Event Identification Based on a Generative Adversarial Network and Support Vector Machine. Frontiers in Earth Science, 10. doi:10.3389/feart.2022.814655.

[26] Ross, Z. E., Trugman, D. T., Hauksson, E., & Shearer, P. M. (2019). Searching for hidden earthquakes in Southern California. Science, 364(6442), 767-771.. doi:10.1126/science.aaw6888.

[27] Jiao, P., & Alavi, A. H. (2020). Artificial intelligence in seismology: Advent, performance and future trends. Geoscience Frontiers, 11(3), 739–744. doi:10.1016/j.gsf.2019.10.004.

[28] Centeno, R., Gómez-Salcedo, V., Lazarte, I., Vilca-Nina, J., Osores, S., & Mayhua-Lopez, E. (2024). Near-real-time multiparametric seismic and visual monitoring of explosive activity at Sabancaya volcano, Peru. Journal of Volcanology and Geothermal Research, 451, 108097. doi:10.1016/j.jvolgeores.2024.108097.

[29] Ozkaya, S. G., Baygin, M., Barua, P. D., Tuncer, T., Dogan, S., Chakraborty, S., & Acharya, U. R. (2024). An automated earthquake classification model based on a new butterfly pattern using seismic signals. Expert Systems with Applications, 238. doi:10.1016/j.eswa.2023.122079.

[30] Lara, F., León, R., Lara-Cueva, R., Tinoco-S., A. F., & Ruiz, M. (2022). Detection of volcanic micro earthquakes based on homomorphic deconvolution and STA/LTA. Journal of Volcanology and Geothermal Research, 421. doi:10.1016/j.jvolgeores.2021.107439.

[31] Song, Y., Lee, J., Yeeh, Z., Kim, M., & Byun, J. (2023). Improved lithospheric seismic velocity and density model of the Korean Peninsula from ambient seismic noise data using machine learning. Journal of Asian Earth Sciences, 254. doi:10.1016/j.jseaes.2023.105728.

[32] Coombs, M. L., Wech, A. G., Haney, M. M., Lyons, J. J., Schneider, D. J., Schwaiger, H. F., Wallace, K. L., Fee, D., Freymueller, J. T., Schaefer, J. R., & Tepp, G. (2018). Short-Term Forecasting and Detection of Explosions During the 2016–2017 Eruption of Bogoslof Volcano, Alaska. Frontiers in Earth Science, 6. doi:10.3389/feart.2018.00122.

[33] Dempsey, D. E., Cronin, S. J., Mei, S., & Kempa-Liehr, A. W. (2020). Automatic precursor recognition and real-time forecasting of sudden explosive volcanic eruptions at Whakaari, New Zealand. Nature Communications, 11(1), 3562. doi:10.1038/s41467-020-17375-2.

[34] Manley, G. F., Pyle, D. M., Mather, T. A., Rodgers, M., Clifton, D. A., Stokell, B. G., Thompson, G., Londoño, J. M., & Roman, D. C. (2020). Understanding the timing of eruption end using a machine learning approach to classification of seismic time series. Journal of Volcanology and Geothermal Research, 401. doi:10.1016/j.jvolgeores.2020.106917.

[35] Saad, O. M., & Chen, Y. (2021). Earthquake Detection and P-Wave Arrival Time Picking Using Capsule Neural Network. IEEE Transactions on Geoscience and Remote Sensing, 59(7), 6234–6243. doi:10.1109/TGRS.2020.3019520.

[36] Mandita, F., Ashari, A., Wibowo, M., & Suryanto, W. (2024). A Prototype for The Detection and Classification of Seismic Events Using STA/LTA And Machine Learning. Journal of Theoretical and Applied Information Technology, 102(10), 4559-4570.

[37] Mustafa, A. H., Abdelmoneim, F. M., Matta, M. G., Barghash, T. O., & Gomaa, W. (2022). Automatic Forecasting of Volcanoes Eruption Time. Proceedings of the 2022 16th International Conference on Ubiquitous Information Management and Communication, IMCOM 2022, 1-8. doi:10.1109/IMCOM53663.2022.9721804.

[38] Abebe, E., Kebede, H., Kevin, M., & Demissie, Z. (2023). Earthquakes magnitude prediction using deep learning for the Horn of Africa. Soil Dynamics and Earthquake Engineering, 170. doi:10.1016/j.soildyn.2023.107913.

[39] Sandhya, G., Babu, J. B., Babu, G. P., & Yallamandaiah, S. (2023). CNN Based Automatic Fault Detection in 3D Seismic Images. 2023 7th International Conference On Computing, Communication, Control And Automation, ICCUBEA 2023, 1-6. doi:10.1109/ICCUBEA58933.2023.10392122.

[40] Lestari, I. (2021). 4 Levels of Volcano Status. Available online: https://ilmugeografi.com/ilmu-bumi/gunung/tingkatan-status-gunung-berapi (accessed on February 2025).