Journal of Theoretical and Applied Information Technology
31st May 2024. Vol.102. No. 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A PROTOTYPE FOR THE DETECTION AND CLASSIFICATION OF SEISMIC EVENTS USING STA/LTA AND MACHINE LEARNING

FRIDY MANDITA[1,2] , AHMAD ASHARI[3*] , M. EDI WIBOWO[4] , WIWIT SURYANTO[5]

[1]Informatic Engineering, Department of Engineering University 17 Agustus 1945, Surabaya, Indonesia

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

[3,4]Department of Computer Science and Electronics Universitas Gadjah Mada, Yogyakarta, Indonesia

[5]Department of Physics Universtas Gadjah Mada, Yogyakarta, Indonesia

E-mail: [1]fridymandita@untag-sby.ac.id, [2]fridymandita@mail.ugm.ac.id, [3*]ashari@ugm.ac.id, [4]mediw@ugm.ac.id, [5]ws@ugm.ac.id

ABSTRACT

In this study, the detection and classification of seismic events is a significant concern of this research. A volcano eruption is one of the natural disasters on Earth. Monitoring volcano activities is essential to analyzing and monitoring volcanoes before their eruption. This activity is beneficial in interpreting signals from a volcano before an eruption from the volcano can cause damage. Based on that, a tool has been developed to detect and classify volcanic seismic events. The combination of algorithm time series, which is STA/LTA and machine learning (LSTM), is being used to analyze data of seismic events. A dataset was collected from one of Indonesia's mountains during 2019 – 2021. The dataset will be classified into different classes based on the type of seismic events. Noise detection is implemented to classify true or false seismic events before continuing to detect and classify them. STA/LTA is used to remove noise signals from data seismic events. The next step is to use machine learning to classify labelling signals based on the type of seismic events. The experiments use a learning rate of 0.001 and 0.01. They show that tools can detect and classify signals of seismic events with an accuracy of around 0,70 – 0,80.

Keywords: *Seismic events, Detection, Classification, STA/LTA, Machine learning*

1. INTRODUCTION

Volcanic eruptions are one of the natural disasters in Indonesia, and they can cause losses to many parties. The geographical location of Indonesia, which is at the confluence of three tectonic plate points, is known as the Ring of Fire [1], resulting in many active volcanoes in Indonesia. From Sabang to Merauke, Indonesia is surrounded by active volcanoes totalling roughly 130 [2]. The impact of this geographic location includes volcanic eruptions that can occur at any time. One of the active and dangerous volcanoes in Indonesia is Mount Merapi [3], [4].

Mount Merapi is a type of stratovolcano with an altitude of 2980 meters above sea level, located 25 - 30 km north of the city of Yogyakarta at 7'32.5' south latitude and 110'26.5' east longitude, administratively located in 4 districts, namely Sleman Regency in Yogyakarta Province, and Magelang Regency, Boyolali Regency, and Klaten Regency in Central Java Province [5].

From the end of the 20th century until the beginning of the 21st century, there was an eruption every 2 - 5 years at Mount Merapi, where the most significant eruption occurred in 2010, the last 100 years. The 2010 eruption started with a solid phreatomagmatic event on 26 October. It reached a climax eruptive phase on the night of 4 – 5 November by producing vertical ash and pyroclastic density currents through the area on the volcano’s south side towards Yogyakarta [6].

Monitoring volcanic activity is an initial activity to assess risk and provide early warning if there is an initial volcanic eruption, primarily intended to warn people living in areas around high-risk volcanoes [7]. Seismic waveform data is time series data, which can be detected using time series analysis. Time series analysis can be classified into three types: 1) Short-Term Average / Long-Term

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Average (STA/LTA), 2) template matching, and 3) autocorrelation/cross-correlation [8].

Machine learning (ML) has been implemented in seismology in recent years, with various applications used to identify invisible signals and patterns and extract information features to increase understanding related to seismology [9]. On the other hand, the use of ML in seismology cannot stop volcanic eruptions, but it plays an essential role in processing seismic signal data to convey information related to volcanic activity. Although the application of ML has progressed, there are still challenges in its implementation in seismology.

In this study, the ML model proposed by [10], [11] automatically classifies seismic signals related to volcanic activity by labelling the dataset to detect an event. Other studies related to monitoring volcanic activity using several ML models to analyze seismic signal data were carried out by [12, 13, 14], resulting in different levels of accuracy for each model applied using seismic signal data using the time waveform, frequency spectrum, and cepstrum methods with various frequencies ranging from 0.5 H_Z – 15 H_Z to sort out the arrival of the P wave signal with class features that have been classified in advance by the type of vibration.

The challenges of implementing ML in seismology include processing volcano observation data, which is mainly done manually; this process should be done automatically. In the detection phase, it is generally carried out semi-automatically or automatically. In contrast, the classification phase is usually carried out manually, which is time-consuming and, depending on the user, will result in the level of accuracy [10], [12], [15]. Another thing that needs to be considered is the poor quality of the geophysical dataset for use as sampling data, causing interference and incomplete data, which is difficult to apply using standard ML techniques and can result in poor detection and classification performance [9], [13].

Based on the challenge of ML, the motivation of the research is to build a prediction model for detecting and classifying seismic events using seismic signal data for volcanic eruption prediction. This research aims to detect and classify seismic events using seismic signal data with noise detection and classification in time series analysis and machine learning with more accurate results than other proposed ML mentioned before when classified data.

Specifically, this research will test the performance of automatically predicting volcanic eruption status using a modified ML architectural model and measure the level of performance and accuracy. The model shows us that limited data can give better results than other ML methods, which generally use extensive data for detection and classification. Data taken within a certain period continuously comes from raw seismic data to obtain data completeness, including the number of events from daily observations.

2. DATA ANALYSIS

The use of machine learning in seismology has experienced significant developments, which can be proven by various research studies related to this field. Research in seismology using ML has been applied in multiple ways, including monitoring volcanoes, earthquake prediction, predicting volcanic eruption status, detecting, and classifying seismic vibration [10]. The data used in this study comes from seismic signal data that is processed using time series analysis and the ML model.

Although ML provides a tool to extract and process information from seismic signal data, distinguishing true or false and the noise of an event requires a trial phase, testing, and application of the ML model to differentiate it. The experiments used a data primer obtained from one of Indonesia's mountains, namely Mount Merapi, as mentioned in chapter 1. The data collected comes from stations around Mount Merapi every day from 2019 to 2021 for collection.

Seismic signal preprocessing data removes poor signal quality from this study's seismic signal data database. To find inconsistencies in the seismic signal database, it is necessary to re-check with the help of people with knowledge of seismology. Not all events recorded by observation stations can adequately be observed about volcanic activity.

From this, it is necessary to carry out a visual check to determine whether the event has been appropriately labelled or not, and this process will be more straightforward if the signal quality is more robust when compared to the noise level around the seismic signal. Patterns in seismic signal classification are made by dividing the dataset manually into smaller pieces (with different durations). Each extracted segment is then classified into a particular class according to the nature of the underlying physical event (reference class).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Seismic signal data contains information about volcanic activity (events) and can be classified into different event types based on waveform and spectrum. We examine seismic event data before categorizing them based on data type with the help of experts in the field of seismology. From the results of the data categorization, seismic events were classified into eight classes, as shown in Table 1. The total data used was around 2500 from monitoring stations around Mount Merapi after data categorization.

Table 1: Type Of Data Seismic Events

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

3. RELATED WORKS

3.1. Event Detection

In recent years, research on event detection for volcanic eruptions has been carried out to detect an event on a volcano, which can detect true or false events using noise detection and classification of volcanic seismic events on volcanoes.

In 2019, research was carried out regarding the detection and classification of continuous seismic signals for seismic events with the help of the ML static model to process data from waveforms using several parameters [15]. The research only worked for five classes of seismic events data type. Classifying eruptive and non-eruptive data is a research topic [14] based on volcanic time series data by comparing the level of accuracy of the four statistical models from the ML used but still not implemented with large datasets.

Research conducted by [16] modified the ML model to process seismic signal data with the help of the polarization analyzer feature for use in areas with moderate levels of seismicity. The accuracy of the proposed method is still weak. Accurate seismic phase detection and identification is essential for detecting and estimating seismic events parameters. In identifying seismic phases, distinguishing between identification for noise and accurate seismic signals is very important [17]. The research only used a small dataset for the experiments.

An experiment to detect volcanic ash was implemented to detect it using a machine-learning model [18]. The process needs more time when used to detect volcanic ash. It can still be improved for time processing. Therefore, an automatic detection method is needed to distinguish true or false volcanic seismic events efficiently and accurately.

3.2. Event Classification

Volcano event classification classifies a volcano's seismic events, whether included in the *normal, waspada, siaga*, or *awas* categories.

Research by [10] uses data from waveform seismic events with six parameters to classify a signal automatically. The experiments only used six classes of type seismic events. Time series analysis is applied to process seismic signal data with the wavelet transform model using four parameters for seismic signal classification carried out in research [19]. The research only aims to detect and classify four classes of type seismic events.

Classification of seismic signals in research [12] uses the features of the time waveform, the spectrum, and the cepstrum from the waveform data for signal classification. The automatic classification of a seismic signal is a challenge in research development, where most of it is still done manually in the process, while for automatic recognition, an automated recognition model related to events from a volcano is needed [12]. The results of the experiments show lower accuracy when used for detection and classification.

4. METHODOLOGY

4.1. Workflow Model for Detection and Classification

Figure 1 shows a model's workflow for predicting a volcano eruption's status. Data from seismic signals will be detected first to distinguish true or false events using noise detection at the event detection stage and at low seismic signals and medium seismic signals with time series analysis before proceeding to the following process, namely seismic signal data preprocessing and signal data feature extraction process seismic at the classification stage of a seismic signal to classify seismic signal data and followed by a learning process with machine learning to process seismic signal data. The training and model testing phase is carried out next to try a model for predicting the status of volcanic eruptions. After all the processes have been carried out, it is expected to produce a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

model that has been validated to detect and classify the status of data seismic events with more optimal results daily.

Figure 1: Workflow Model Detection and Classification

4.2. STA/LTA Algorithm

Short Term Amplitude (STA) / Long Term Amplitude (LTA) is a typical technique for earthquake detection, shown in Figure 2. This is derived from the results of human detection (experts), where a fundamental change in amplitude indicates a potential earthquake and can be identified visually. STA/LTA has two critical parameters: the length of the short term window and the length of the long-term window. Standard parameter selections may have a short-term window of three seconds and a long-term window of thirty seconds. The choice of the third parameter can be used to change the overlap of the short-term window and the tail end of the long-term window [8].

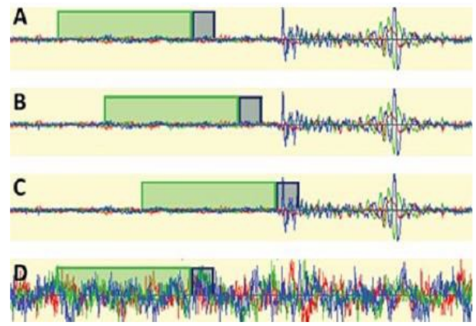


Figure 2: STA/LTA Algorithm

STA / LTA (with proper alignment) has the following equation:

$$STA\,(x_i) = \frac{1}{ns}\;\sum_{j=i-ns}^{i} x_i^2 \qquad (1)$$

$$LTA\,(x_i) = \frac{1}{nl}\;\sum_{j=i-nl}^{i} x_i^2$$

$$r_i = \frac{STA_i}{LTA_i}$$

STA/LTA does not rely on previous data, making it useful when a new station is created and no prior data is stored. STA/LTA is beneficial because it has no prerequisites, is linear with a time complexity of O(n) and can recognize signals with unique characteristics. However, it relies on a high signal-to-noise ratio to be effective.

4.3. Detection of Signal Seismic

Detection of seismic signal data is the initial stage for processing seismic waveform data from a volcano observation station. Effective detection can save time and effort in obtaining data related to the stages and locations of events, especially in areas with moderate seismic activity on a local or regional scale [16]. Detecting a seismic signal is a time series analysis process or matching a pattern from seismic waveform data [8] to detect or distinguish true or false a volcanic seismic event in the presence of noise.

The process of detecting (event detection) seismic signals is usually carried out by applying a threshold equal to the seismic amplitude ratio

Journal of Theoretical and Applied Information Technology
31st May 2024. Vol.102. No. 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

between the short-term average (STA) and the long-term average (LTA) of the seismic signal amplitude, where STA/LTA is one approach to time series analysis [20]. The number of disturbance samples (noise) must be several times greater than the number of occurrences within an incident, and the number of disturbances for each station must be the same [15].

To detect an event, it is necessary to continuously obtain the probability of occurrence from time to time from each observation station around Mount Berapau so that all activities can be combined to declare an event. The goal is to establish that a genuine seismic event occurred. The original seismic event is sufficient to determine whether the event is true or false (true or false) of a seismic event. It can be defined as grouping an event with the noise level of seismic signal data [21].

4.4. Data Preprocessing of Data Seismic Events

Seismic signal preprocessing data removes poor signal quality from the seismic signal data database for this study. To find inconsistencies in the seismic signal database, it is necessary to re-check with the help of people who know the field of seismology. Not all events recorded by observation stations can adequately be observed about volcanic activity.

From this, it is necessary to carry out a visual check to determine whether the event has been appropriately labeled or not, and this process will be more straightforward if the signal quality is more robust when compared to the noise level around the seismic signal. Patterns in seismic signal classification are done by dividing the dataset manually into smaller pieces (with different durations). Each extracted segment is then classified into a particular class according to the nature of the underlying physical event (reference class). Data seismic events are shown in Table 1, chapter data analysis.

4.5. Data Signal Seismic

Volcanic activity data collection was obtained from observation stations around the volcano. This study will use primary data from the Geological Disaster Technology Research and Development Center (BPPTKG) at the Mount Merapi observation station in Yogyakarta at several observation stations.

Seismic signal data related to its activity is collected from observation stations around Mount Merapi within a certain period. Data is taken with conditions at a frequency between 0.5 Hz – 50 Hz. The data collected in the form of related images from the activity of Mount Merapi, which is used as the object of this research, is shown in Figure 3.

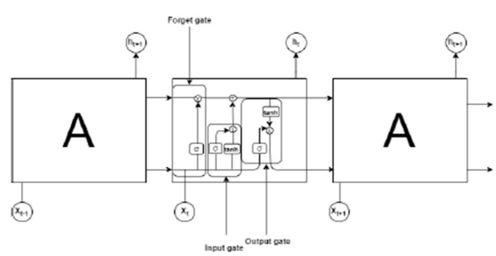
Figure 3: Example Data Seismic

4.6. Long Short Term Memory (LSTM)

An RNN called the Long Short Term Memory (LSTM) algorithm was put out by [21]. An artificial neural network with memory states included is called an RNN. RNN may handle time series data as they often include interdependencies. Still, Conventional RNN are untrainable because of their disappearing and exploding gradients, which renders them subpar solutions. An LSTM can solve the RNN puzzle. It has three essential parts that are referred to as gates. The input gate, forget gate, and output gate are these. There are also two memory cells: internal and concealed [23].

Figure 4: Long Short-Term Memory (LSTM) Architecture

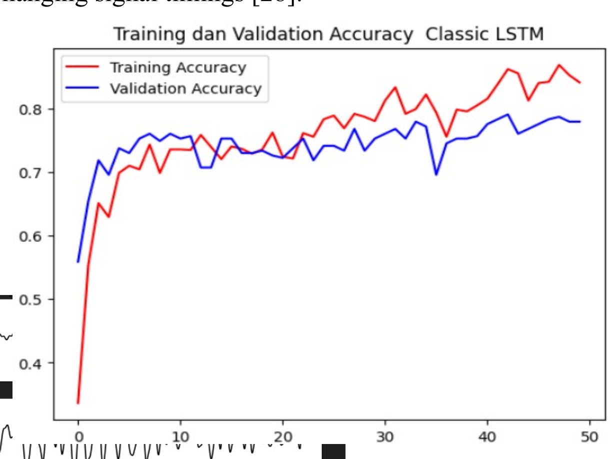
In the area of seismic events, LSTM has been used for a variety of purposes, such as the introduction of seismic events [24], the detection of

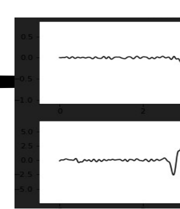


P waves in seismic event data [25], the prediction of seismic events [26], and the introduction of seismic events [27].

The technique was selected due to its ability to simulate long-term dependencies, delete or add information from memory cell states, and represent the dynamics of changing signal timings [28].

5. RESULTS AND





Journal of Theoretical and Applied Information Technology
31st May 2024. Vol.102. No. 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DISCUSSIONS

In this stage, detection and classification will be built to detect true or false events using noise detection and automatically classify the classification of seismic signal data for volcano eruption prediction. At this stage, data identification and feature extraction will be carried out. The data obtained is analyzed to determine whether it is sufficient to build the expected machine learning model or not. Conducting data cleansing to train the best machine learning model to get the best detection and classification of data seismic events. For this experiment, the seismic data will be divided into eight classes of seismic event data.

For input, the ML model input can be defined as follows:

$$x_t = \begin{matrix} AP \\ DG \\ LF \\ MP \\ RF \\ TR \\ VTA \\ VTB \end{matrix} \qquad (2)$$

Data analysis is also carried out related to the results of data classification by involving people who are experts in the field of seismology. When testing data that has been carried out repeatedly, it is necessary to test data validation to get the most optimal results from the model that is built. The stages in training and testing the model detection and classification of seismic events.

Results of training and validation accuracy of Classic LSTM using a learning rate of 0.001 are shown in Figure 5, Figure 6, Figure 7, and Figure 8

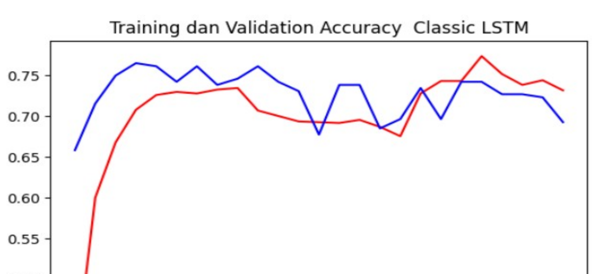


Figure 5: Training and Validation Accuracy Epochs 25 Classic LSTM

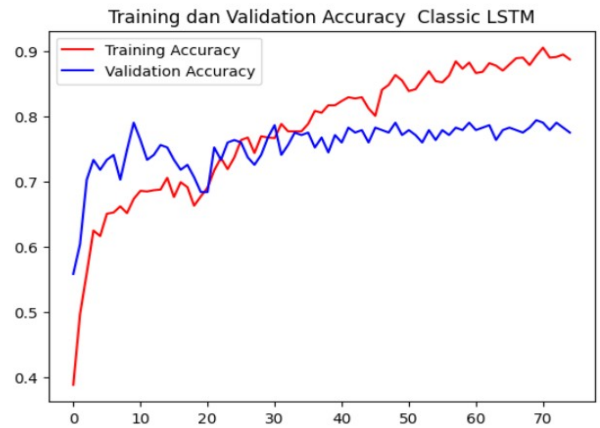


Figure 6: Training and Validation Accuracy Epochs 50 Classic LSTM

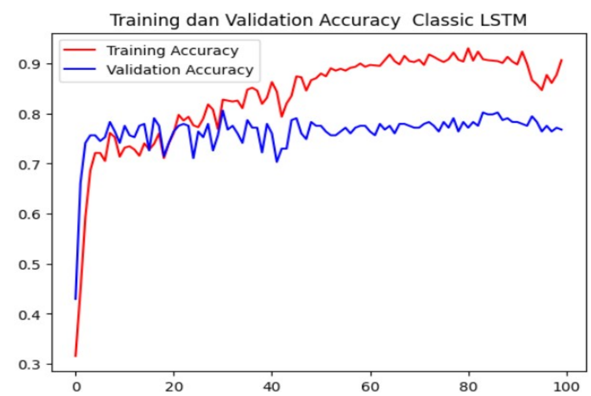


Figure 7: Training and Validation Accuracy Epochs 75 Classic LSTM

Figure 8: Training and Validation Accuracy Epochs 100 Classic LSTM

Figure 5, Figure 6, Figure 7, and Figure 8 show the training and accuracy results of Classic LSTM using epochs between 25 and 100, producing an accuracy between 0.73 and 0.77.

The training and validation accuracy results for Vanilla LSTM with a learning rate of 0.001 are shown in Figure 9, Figure 10, Figure 11, and Figure 12.

Figure 9: Training and Accuracy Loss Epochs 25 with Vanilla LSTM

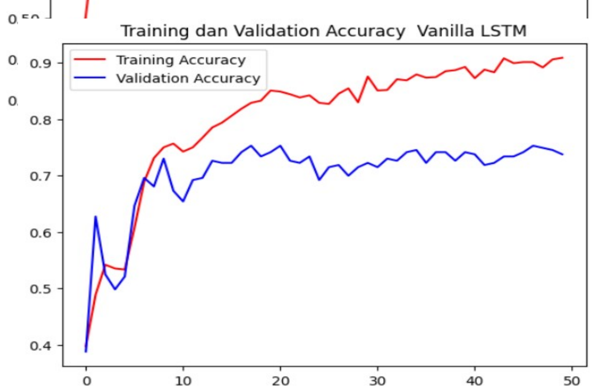
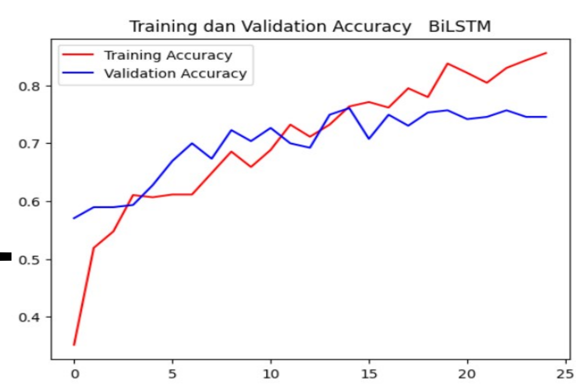


Figure 10: Training and Validation Accuracy Epochs 50 Vanilla LSTM



Journal of Theoretical and Applied Information Technology
31st May 2024. Vol.102. No. 10
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

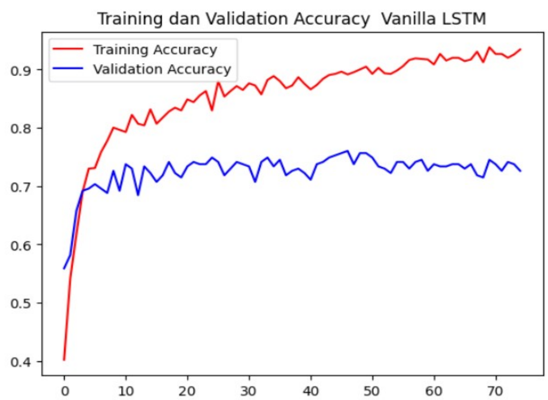


Figure 11: Training and Validation Accuracy Epochs 75 Vanilla LSTM

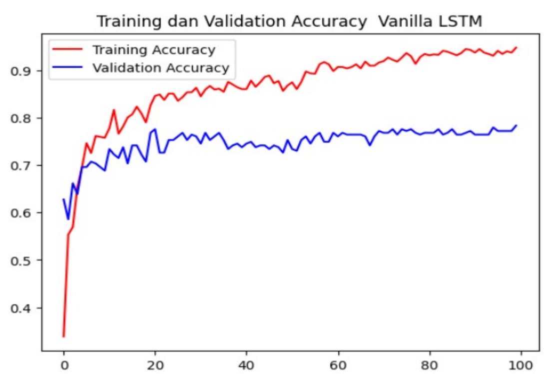


Figure 12: Training and Validation Accuracy Epochs 100 Vanilla LSTM

Figure 9, Figure 10, Figure 11, and Figure 12 show the training and accuracy results for Vanilla LSTM with epochs from 25 to 100, with accuracies ranging from 0.76 to 0.78.

BiLSTM with a 0.001 learning rate training and validation accuracy results are displayed in Figure 13, Figure 14, Figure 15, and Figure 16.

Figure 13: Training and Accuracy Loss Epochs 25 with BiLSTM

Figure 14: Training and Accuracy Loss Epochs 50 with BiLSTM

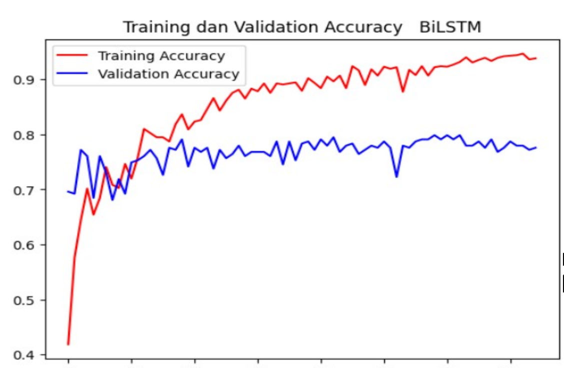


Figure 15: Training and Accuracy Loss Epochs 75 with BiLSTM

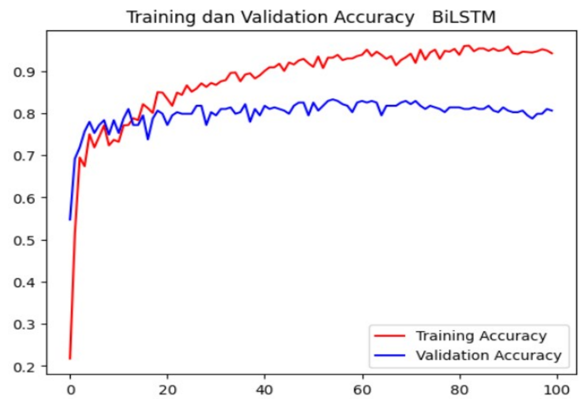


Figure 16: Training and Accuracy Loss Epochs 100 with BiLSTM

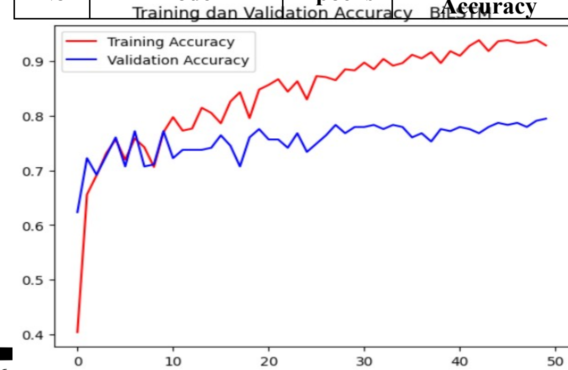
The training and accuracy results for the BiLSTM with epochs ranging from 25 to 100 and accuracies between 0.77 and 0.80 are displayed in Figure 13, Figure 14, Figure 15, and Figure 16.

The results of detection and classification are shown in Table 2.

Table 2 shows the results of experiments on the detection and classification of seismic signals for volcano eruption prediction for training and validation accuracy of the data used. The Classic LSTM model at epochs 25 produces an output of 0.73 compared to the Vanilla LSTM model, which produces a result of 0.77, which is the same as the BiLSTM model. At epochs 50, Classic LSTM gave the lowest results at 0.76, a difference of 0.1 points when compared to the results from the Vanilla LSTM model at 0.77, but this was still below the results from the BiLSTM model at 0.78.

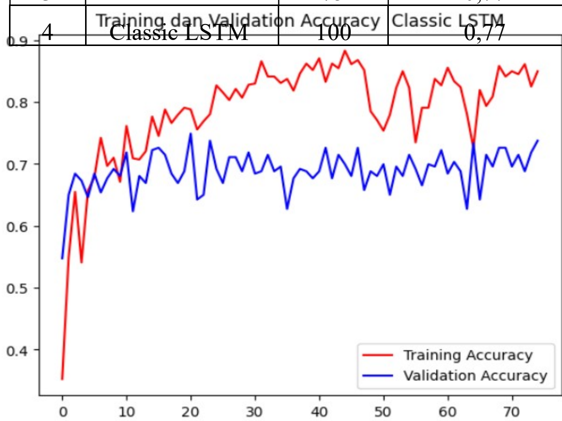
Table 2: Average Test Accuracy for Detection and Classification with Learning Rate 0,001

No	Model	Epochs	Average Test Accuracy



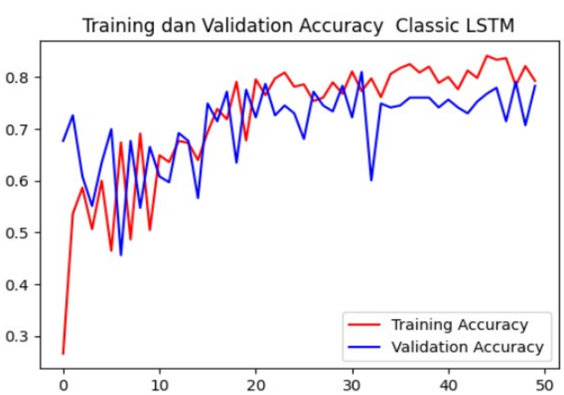
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

1	Classic LSTM	25	0,73
2	Classic LSTM	50	0,76
3	Classic LSTM	75	0,77
4	Classic LSTM	100	0,77
5	Vanilla LSTM	25	0,76
6	Vanilla LSTM	50	0,77
7	Vanilla LSTM	75	0,78
8	Vanilla LSTM	100	0,78
9	BiLSTM	25	0,77
10	BiLSTM	50	0,78
11	BiLSTM	75	0,79
12	BiLSTM	100	0,8



Meanwhile, in epochs 75 Classic LSTM gave a result of 0.77, which is slightly different from the results with the Vanilla LSTM and BiLSTM models with a difference between 0.1 - 0.2, for Vanilla LSTM it gave a result of 0.78 and BiLSTM at 0, 79. For epochs 100, BILSTM gives the best results compared to Classic LSTM and Vanilla LSTM with results of 0.80 compared to other models at 0.77 for Classic LSTM and 0,78 for Vanilla LSTM.

Another experiment we used for detection and classification using a learning rate of 0,01 to compare with a learning rate of 0,001 using mode



Classic LSTM, Vanilla LSTM, and BiLSTM.

The Classis LSTM with a learning rate of 0.01 results are shown in Figure 17.

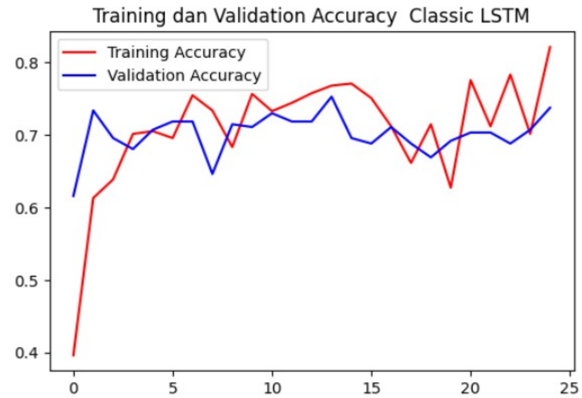


Figure 17: Training and Validation Accuracy Epochs 25 Classic LSTM

The training and accuracy results of the Classic LSTM employing epochs ranging from 25 to 100 are displayed in Figure 17, Figure 18, Figure 19, and Figure 20, with accuracy of 0.72 to 0.76.

Figure 18: Training and Validation Accuracy Epochs 50 Classic LSTM

Figure 19: Training and Validation Accuracy Epochs 75 Classic LSTM

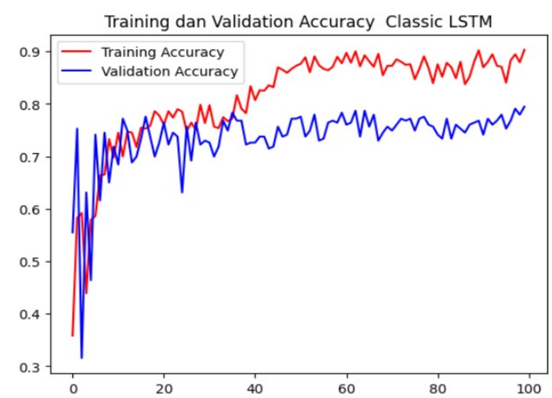


Figure 20: Training and Validation Accuracy Epochs 100 Classic LSTM

Figure 21, Figure 22, Figure 23, and Figure 24 display the training and validation accuracy results for a Vanilla LSTM with a learning rate of 0.01.

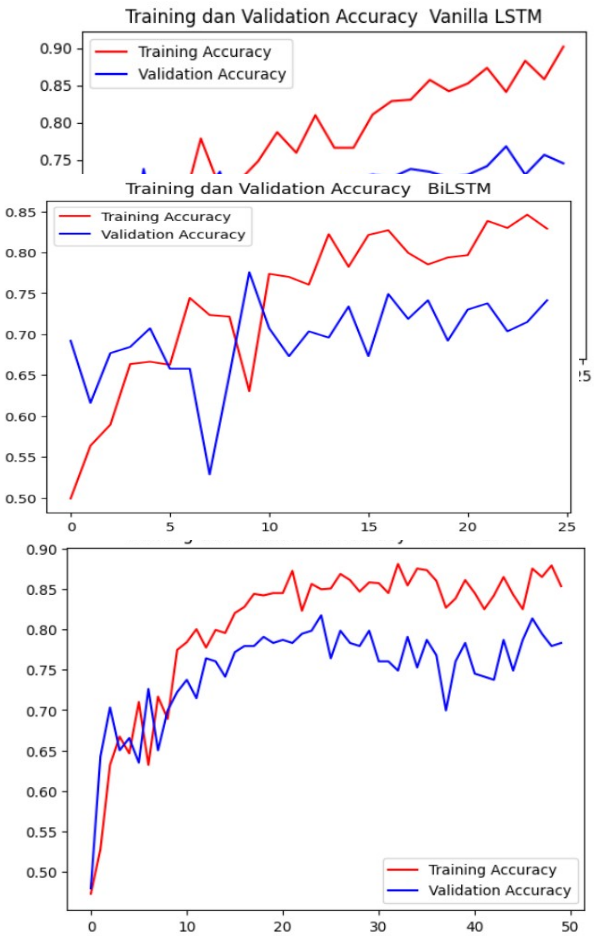


Figure 21: Training and Validation Accuracy Epochs 25 Vanilla LSTM

Figure 22: Training and Validation Accuracy Epochs 50 Vanilla LSTM

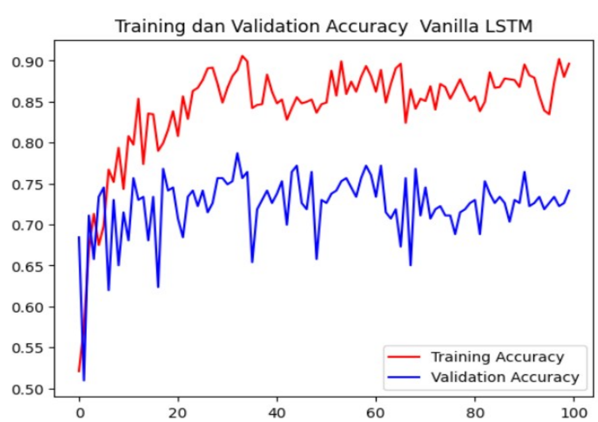


Figure 23: Training and Validation Accuracy Epochs 75 Vanilla LSTM

Figure 24: Training and Validation Accuracy Epochs 100 Vanilla LSTM

Figure 17, Figure 18, Figure 19, and Figure 20 show the training and accuracy results of the Classic LSTM with epochs ranging from 25 to 100, with an accuracy of 0.75 to 0.78.

Figure 25, Figure 26, Figure 27, and Figure 28 display the accuracy results for training and validation of the BiLSTM model with a learning rate of 0.01.

Figure 25: Training and Validation Accuracy Epochs 25 BiLSTM

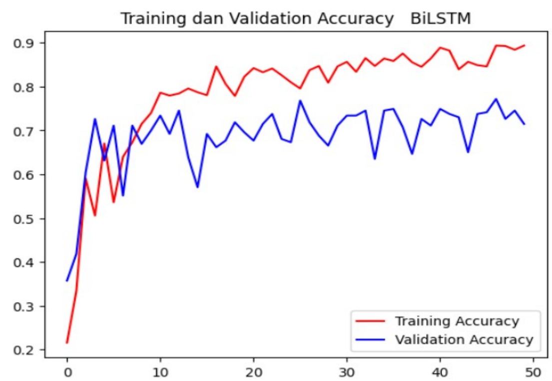


Figure 26: Training and Validation Accuracy Epochs 50 BiLSTM

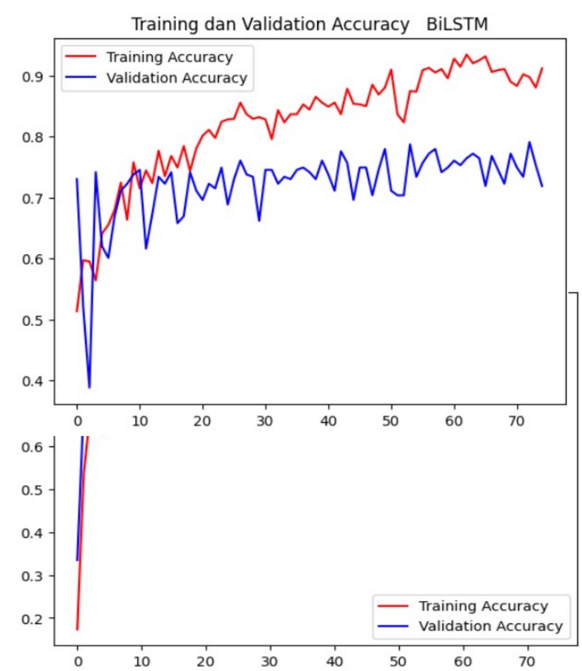


Figure 27: Training and Validation Accuracy Epochs 75 BiLSTM

Journal of Theoretical and Applied Information Technology
31st May 2024. Vol.102. No. 10
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

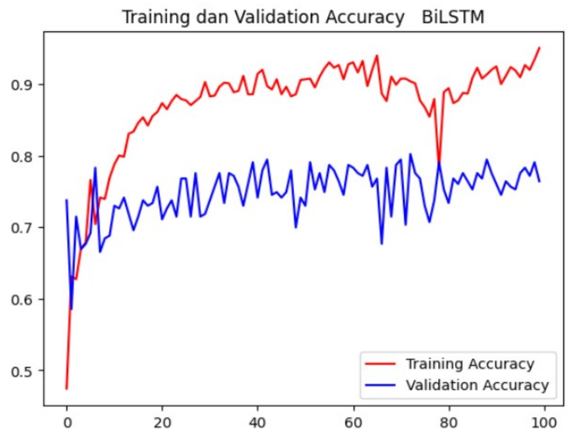


Figure 28: Training and Validation Accuracy Epochs 100 BiLSTM

Figure 25, Figure 26, Figure 27, and Figure 28 show the training and accuracy results for the BiLSTM with epochs ranging from 25 to 100 and accuracies between 0.77 and 0.80.

Table 3 displays the outcomes of the detection and classification processes.

Table 3 displays the findings from tests on the detection and categorization of seismic signal data for training and validation. At epoch 25, the conventional LSTM model provides an output of 0.72, while the vanilla LSTM model produces a result of 0.75, which is identical to the BiLSTM model. At epoch 50, the Classic LSTM model had the lowest results at 0.74, a 0.2 point difference from the Vanilla LSTM model at 0.76. However, this was still lower than the BiLSTM model at 0.78, with a value difference of roughly 0.2. Meanwhile, in epochs 75, Classic LSTM produced a result of 0.75, which differs significantly from the results of the Vanilla LSTM and BiLSTM models by 0.01 - 0.02; Vanilla LSTM produced a result of 0.77, while BiLSTM produced 0.79. For epochs 100, BILSTM outperforms Classic LSTM and Vanilla LSTM with results of 0.80, whereas Classic LSTM and Vanilla LSTM produce 0.77 and 0.78, respectively.

Table 3: Average Test Accuracy for Detection and Classification with Learning Rate 0,01

No	Model	Epochs	Average Test Accuracy
1	Classic LSTM	25	0,72
2	Classic LSTM	50	0,74
3	Classic LSTM	75	0,75
4	Classic LSTM	100	0,76
5	Vanilla LSTM	25	0,75
6	Vanilla LSTM	50	0,76
7	Vanilla LSTM	75	0,77
8	Vanilla LSTM	100	0,78
9	BiLSTM	25	0,77
10	BiLSTM	50	0,78
11	BiLSTM	75	0,79
12	BiLSTM	100	0,8

Table 3 displays the findings from tests on the detection and categorization of seismic signal data for training and validation. At epoch 25, the conventional LSTM model provides an output of 0.72, while the vanilla LSTM model produces a result of 0.75, which is identical to the BiLSTM model. At epoch 50, the Classic LSTM model had the lowest results at 0.74, a 0.2 point difference from the Vanilla LSTM model at 0.76. However, this was still lower than the BiLSTM model at 0.78, with a value difference of roughly 0.2. Meanwhile, in epochs 75, Classic LSTM produced a result of 0.75, which differs significantly from the results of the Vanilla LSTM and BiLSTM models by 0.01 - 0.02; Vanilla LSTM produced a result of 0.77, while BiLSTM produced 0.79. For epochs 100, BILSTM outperforms Classic LSTM and Vanilla LSTM with results of 0.80, whereas Classic LSTM and Vanilla LSTM produce 0.77 and 0.78, respectively.

The research is focused on building a model for detection and classification to predict volcano eruption using time series analysis, which is STA/LTA, and machine learning (LSTM). Based on the experiments, the model that we built gives better results when detecting and classifying seismic events using a small dataset. The accuracy reached point 0,80.

The model we build combines time series analysis and machine learning, which is LSTM, to detect and classify signals of seismic events, showing that the model can recognize seismic events that are already classified in the eight major categories better using the model we build. The experiments show that for events using a small dataset, the model gives an accuracy of around 0,80.

6. CONCLUSIONS

From the experimental results, we can conclude that the model we propose, namely STA LTA and BiLSTM, shows better results when used for detecting and classifying seismic events for volcanic eruption prediction using epochs 25 – 100. It can be concluded that although it can be recognized using small seismic event data, the model we propose, namely STA/LTA and BiLSTM, shows better results. At epochs 25, it gives a result of 0.77, whereas at epochs 50, it shows a result of 0.78. Meanwhile, for epoch 75, it was at 0.79, and the accuracy at epoch 100 was 0.80. the level of accuracy may be further improved by increasing the amount of data and modifying the machine learning models we use. The open issue of this research is how to use another time series analysis and machine learning to detect and classify seismic events and how accurate it is when using large datasets.

Future work of this research is to improve the accuracy of the detection and classification model of seismic event data using more datasets or different datasets to classify and detect seismic events. Other work in the future is to build a parallel volcanic eruption status model by adding larger seismic event data to obtain more precise results in terms of determining the volcanic eruption status.

Acknowledgements

This research is funded by the DRPM scheme PDD for the budget year 2022.

REFERENCES:

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

[2] A. Ratdomopurbo *et al.*, "Overview of the 2006 eruption of Mt. Merapi," *Journal of Volcanology and Geothermal Research*, vol. 261, pp. 87–97, Jul. 2013, doi: 10.1016/j.jvolgeores.2013.03.019.

[3] S. S. Sholihat, S. W. Indratno, and U. Mukhaiyar, "The role of parameters in Bayesian Online Changepoint Detection: detecting early warning of mount Merapi eruptions," *Heliyon*, vol. 7, no. 7, Jul. 2021, doi: 10.1016/j.heliyon.2021.e07482.

[4] Surono *et al.*, "The 2010 explosive eruption of Java's Merapi volcano-A '100-year' event," *Journal of Volcanology and Geothermal Research*, vol. 241–242, pp. 121–135, Oct. 2012, doi: 10.1016/j.jvolgeores.2012.06.018.

[5] A. Aji, A. I. Benardi, W. Setyaningsih, and K. D. I. A. P Yohanes, "Study of the merapi volcano eruption and the impact on community agricultural landuse in sleman regency," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing Ltd, Mar. 2021. doi: 10.1088/1755-1315/683/1/012137.

[6] P. Jousset, J. Pallister, and Surono, "The 2010 eruption of Merapi volcano," *Journal of Volcanology and Geothermal Research*, vol. 261. pp. 1–6, Jul. 01, 2013. doi: 10.1016/j.jvolgeores.2013.05.008.

[7] S. S. S. K. B. S. F. J. C. V.-B. et al. Susan C. Loughlin, "Book Review," *American Mineralogist*, vol. 103, no. 4, pp. 664–664, Apr. 2018, doi: 10.2138/am-2018-678.

[8] J. Martin and L. Audretsch, "Earthquake Detection using Deep Learning Based Approaches," 2020.

[9] Q. Kong, D. T. Trugman, Z. E. Ross, M. J. Bianco, B. J. Meade, and P. Gerstoft, "Machine learning in seismology: Turning data into insights," *Seismological Research Letters*, vol. 90, no. 1, pp. 3–14, Feb. 2019, doi: 10.1785/0220180259.

[10] M. Malfante, M. Dalla Mura, J. I. Mars, J. P. Métaxian, O. Macedo, and A. Inza, "Automatic Classification of Volcano Seismic Signatures," *J Geophys Res Solid Earth*, vol. 123, no. 12, pp. 10,645-10,658, Dec. 2018, doi: 10.1029/2018JB015470.

[11] M. Malfante, M. Dalla Mura, J. P. Metaxian, J. I. Mars, O. Macedo, and A. Inza, "Machine Learning for Volcano-Seismic Signals: Challenges and Perspectives," *IEEE Signal Process Mag*, vol. 35, no. 2, pp. 20–30, Mar. 2018, doi: 10.1109/MSP.2017.2779166.

[12] A. Falcin *et al.*, "A machine-learning approach for automatic classification of volcanic seismicity at La Soufrière Volcano, Guadeloupe," *Journal of Volcanology and Geothermal Research*, vol. 411, Mar. 2021, doi: 10.1016/j.jvolgeores.2020.107151.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[13] Z. Li, M. A. Meier, E. Hauksson, Z. Zhan, and J. Andrews, "Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning," *Geophys Res Lett*, vol. 45, no. 10, pp. 4773–4779, May 2018, doi: 10.1029/2018GL077870.

[14] G. F. Manley *et al.*, "Understanding the timing of eruption end using a machine learning approach to classification of seismic time series," *Journal of Volcanology and Geothermal Research*, vol. 401, Sep. 2020, doi: 10.1016/j.jvolgeores.2020.106917.

[15] M. Titos, A. Bueno, L. García, M. C. Benítez, and J. Ibañez, "Detection and Classification of Continuous Volcano-Seismic Signals with Recurrent Neural Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 1936–1948, Apr. 2019, doi: 10.1109/TGRS.2018.2870202.

[16] J. Wiszniowski, B. Plesiewicz, and G. Lizurek, "Machine learning applied to anthropogenic seismic events detection in Lai Chau reservoir area, Vietnam," *Comput Geosci*, vol. 146, Jan. 2021, doi: 10.1016/j.cageo.2020.104628.

[17] J. Li *et al.*, "Gan-lstm joint network applied to seismic array noise signal recognition," *Applied Sciences (Switzerland)*, vol. 11, no. 21, Nov. 2021, doi: 10.3390/app11219987.

[18] F. Torrisi, C. Corradino, S. Cariello, and C. Del Negro, "Enhancing detection of volcanic ash clouds from space with convolutional neural networks," *Journal of Volcanology and Geothermal Research*, vol. 448, Apr. 2024, doi: 10.1016/j.jvolgeores.2024.108046.

[19] J. P. Canário, R. Mello, M. Curilem, F. Huenupan, and R. Rios, "In-depth comparison of deep artificial neural network architectures on seismic events classification," *Journal of Volcanology and Geothermal Research*, vol. 401, Sep. 2020, doi: 10.1016/j.jvolgeores.2020.106881.

[20] G. Saccorotti and I. Lokmer, "A review of seismic methods for monitoring and understanding active volcanoes," in *Forecasting and Planning for Volcanic Hazards, Risks, and Disasters*, Elsevier, 2020, pp. 25–73. doi: 10.1016/B978-0-12-818082-2.00002-0.

[21] A. Reynen and P. Audet, "Supervised machine learning on a network scale: Application to seismic event classification and detection," *Geophys J Int*, vol. 210, no. 3, pp. 1394–1409, Sep. 2017, doi: 10.1093/gji/ggx238.

[22] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[23] A. Berhich, F. Z. Belouadha, and M. I. Kabbaj, "LSTM-based Models for Earthquake Prediction," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Mar. 2020. doi: 10.1145/3386723.3387865.

[24] A. Falcin *et al.*, "A machine-learning approach for automatic classification of volcanic seismicity at La Soufrière Volcano, Guadeloupe," *Journal of Volcanology and Geothermal Research*, vol. 411, Mar. 2021, doi: 10.1016/j.jvolgeores.2020.107151.

[25] W. Zhu, X. Li, C. Liu, F. Xue, and Y. Han, "An STFT-LSTM System for P-Wave Identification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 3, pp. 519–523, Mar. 2020, doi: 10.1109/LGRS.2019.2922536.

[26] O. Nicolis, F. Plaza, and R. Salas, "Prediction of intensity and location of seismic events using deep learning," *Spat Stat*, vol. 42, Apr. 2021, doi: 10.1016/j.spasta.2020.100442.

[27] S. M. Mousavi, W. Zhu, Y. Sheng, and G. C. Beroza, "CRED: A Deep Residual Network of Convolutional and Recurrent Units for Earthquake Signal Detection," *Sci Rep*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/s41598-019-45748-1.

[28] N. B. Yoma *et al.*, "End-to-end LSTM based estimation of volcano event epicenter localization," *Journal of Volcanology and Geothermal Research*, vol. 429, Sep. 2022, doi: 10.1016/j.jvolgeores.2022.107615.