JOSTEN
Journal of Science, Technology and Neurology
Open Access

The Level of Accuracy of Machine Learning Measurements in Predicting Earthquakes on Lombok Island

Rizky Munandar[1*]

[1]Department of Electronic and Informatics Engineering, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
Email: rizkymunan03@gmail.com

Article Information:
Received: 03 November 2025
Revised: 29 November 2025
Accepted: 01 January 2026
Published: 02 January 2026

check for updates
https://doi.org

Abstract

Introduction: Lombok Island is one of the regions with the highest seismicity levels in Indonesia due to its geographical location trapped between two active earthquake sources: the subduction zone of the Indo-Australian plate with the Eurasian plate to the south and the Flores Back Arc Thrust Fault to the north. Predictions of earthquakes are still very minimal; the 2018 Lombok earthquake is one of the reasons why earthquake detectors need to be developed.
Objective: This study aims to analyze the use of machine learning in measuring the accuracy of earthquake predictions on the island of Lombok.
Methods: The method used in this study is comparative entrepreneurship which utilizes secondary data from the earthquake catalog during the January-October 2018 period to be analyzed using 3 machine larning algorithms, namely Naive Bayes, Artificial Neural Network (JST) and KNN.
Results: The results showed that the accuracy value using Naive Bayes was 0.6 and the accuracy using JST and KNN was 0.5. However, this is different from the results of the evaluation of the three algorithms where Naive Bayes still has a value of 0.6 but JST and KNN become 0.4.
Conclusion: In conclusion, the accuracy of machine learning measurements from the three algorithms shows that Naive Bayes has a high accuracy value, but this result may change if other algorithms are used.

Keywords: Accuracy Level, Machine Learning, Earthquake Prediction

Introduction

Lombok Island is one of the areas with the highest seismicity rate in Indonesia due to its geographical location trapped between two active earthquake sources: the Indo-Australian plate subduction zone with the Eurasian plate to the south and the Flores Back Arc Thrust Fault to the north [1]. These complex tectonic conditions make Lombok a perfect yet alarming natural laboratory for earthquake research. The existence of subduction zones in the south, coupled with the back arc thrust in the north and local faults, makes this region have very high tectonic earthquake activity [2].

The importance of this research is very important considering the consequences of the disaster caused by a series of earthquakes in Lombok in 2018. The 2018 Lombok and Sumbawa earthquakes caused 564 deaths, 1,584 injuries, 396,032 displaced residents, and damaged 239,954 houses. Economic losses reached 7.45 trillion rupiah, of which the residential sector contributed 81 percent of the total reported damage. What is interesting to observe is the characteristics of the series of earthquakes that occurred, where a series of earthquakes on July 29, 2018, August 5, 2018, and August 19, 2018 resulted in 559 deaths, 1,478 injuries, and 185,483 buildings damaged. The characteristics of this large-magnitude foreshock-mainshock-aftershock reflect the complexity of seismic patterns in the Lombok area that are still not fully predictable using traditional methods [3].

The main obstacle in earthquake disaster management in Lombok is the lack of proper prediction capabilities. Traditional earthquake prediction methods still face major obstacles in recognizing the complex patterns of interaction between the Flores Rising Fault and regional tectonic systems [4]. The Ascending Flores Fault proved to be more active than the subduction zone because the "sloping subduction" coupling pressure from the Flores Thrust down the Lesser Sunda Islands was stronger. However, its seismic activity patterns remain difficult to predict deterministically [5].

Machine learning provides an innovative paradigm for predicting earthquakes with its ability to identify complex patterns from vast historical data. Recent studies have shown promising results: the stacking model combining Random Forest and XGBoost performed best with the lowest Mean Squared Error (MSE) value of 0.108 and the highest R-squared (R^2) value of 0.892 Seminar-id, while the Los Angeles study managed to achieve 97.97% accuracy in predicting the maximum magnitude category of an earthquake using Random Forest. However, the performance of these models is highly dependent on specific regional tectonic characteristics [6].

The specification of this study focuses on the urgent need to assess how well the machine learning model that has been built can be adjusted and provide precise predictions for the typical tectonic context of Lombok [7]. Unlike other areas that only have one main earthquake source, Lombok has two earthquake sources (subduction and back-arc thrust) with different characteristics. This research is important because: (1) Lombok has a high level of population density with weak infrastructure, (2) the psychological and social effects of the 2018 earthquake are still felt in the community, (3) the threat of major earthquakes from the Flores Rising Fault still exists, such as the 1992 Flores tsunami which claimed 2,500 lives, and (4) there is no standard measure of accuracy for machine learning-based earthquake prediction systems in areas with tectonic complexity such as Lombok [8].

In addition, the importance of this research is reinforced by the fact that Lombok Island is an area with a high level of seismicity due to the existence of a complex tectonic order with two main faults, but there has been no in-depth study that specifically evaluates the accuracy of various machine learning algorithms in predicting earthquake parameters in this region. Previous studies have focused more on analyzing the characteristics of earthquakes that have occurred or general modeling throughout Indonesia, without providing an accuracy assessment specific to the typical Lombok tectonic context [9].

Therefore, this study aims to bridge this knowledge gap through a systematic analysis of the accuracy of machine learning models in predicting earthquakes in Lombok. The findings of this study are expected to provide a scientific basis for the development of more effective early warning systems, support risk-based regional planning, and ultimately reduce the likelihood of loss of life and material losses due to future earthquakes. The success of this study will not only have an impact on the advancement of computational seismology, but also provide immediate practical benefits for disaster mitigation efforts in one of the most earthquake-risk areas in Indonesia [10].

Method

Types of Research

This study applied a quantitative approach with a comparative experimental method to assess the accuracy of three machine learning algorithms in predicting tsunami and landslide potential based on seismic data in Lombok during the January-October 2018 period.

Research Data and Variables

This study utilizes secondary data from the BMKG earthquake catalog for the Lombok region during the period from January to October 2018 with a total of 10 monthly observations. Independent variables include the total number of earthquakes each month (X_1), the number of earthquakes with a magnitude below 3 (X_2), the number of earthquakes with a magnitude between 3-5 (X_3), and the number of earthquakes with a magnitude greater than 5 (X_4). The bound variables included tsunami potential (Y_1) and landslide potential (Y_2) with a classification of Yes/No. The dataset consists of 70% training data (7 months) and 30% test data (3 months).

Algoritma Machine Learning

This study uses three machine learning algorithms that will calculate the accuracy of earthquake predictions, the machine learning equations used are as follows:

a. Naive Bayes [11]

$$P\,(C_k\ /\ X) = \frac{P\,(X\ /\ C_k)\ .\ P\,(C_k)}{P\,(X)} \tag{1}$$

Class predictions are selected based on maximum probability [12]:

$$\hat{y}=arg\ mak_k\ P(C_k)\ \prod_{i=1}^{n} P\,(X_i\ /\ C_k) \tag{2}$$

b. Artificial Neural Network (JST)

Forward Propagation [13]:

$$z^{[l]}=W^{[l]}a^{[l-1]} + b^{[l]} \tag{3}$$

$$a^{[l]} = g(z^{[l]}) \tag{4}$$

Activation Function (Sigmoid) [14]:

$$\sigma(z)=\frac{1}{1+e^{-z}} \tag{5}$$

Loss Fuction (Binary Cross-Entropy) [15]:

$$L(y,\hat{y})=-\frac{1}{m}\sum_{i=1}^{m}[y_i\,log\,(\hat{y}_i)+(1-y_i)\,log\,(1-\hat{y}_i)] \tag{6}$$

With gradient descent optimization with learning rate α = 0.01-0.1.

c. K-Nearest Neighbors (KNN)

Calcification based on k nearest neighbor using Euclidean distances [16]:

$$d(x_i x_j)=\sqrt{\sum_{l=1}^{n}(x_{il}-x_{jl})^2} \tag{7}$$

Prediction based on majority voting [17]:

$$\hat{y}=arg\ mak_c \sum_{i\in N_k(x)} I(y_i=c) \tag{8}$$

The k parameter is determined through cross-validation with candidates k = 3, 5, 7.

Furthermore, to measure the accuracy value more precisely, an evaluation was carried out using the following equation.

a. Accuracy [18]

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{9}$$

b. Precision [19]

$$Precision=\frac{TP}{TP+FP} \tag{10}$$

c. Sensitivity [20]

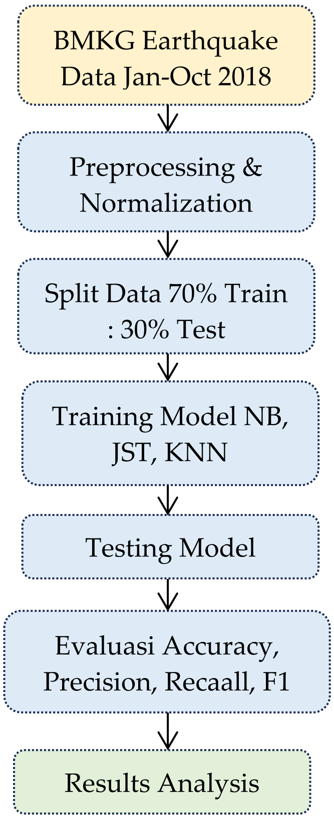
$$Recall=\frac{TP}{TP+FN} \tag{11}$$

d. F1-Score [21]

$$F1\text{-}Score=2x\frac{Precision\ x\ Recall}{Precision + Recall} \tag{12}$$

Research Procedure

This research was carried out in stages according to the following figure.



Gambar 1. Research Procedure

Results and Discussion

Results

This study evaluated the accuracy of three machine learning algorithms in predicting the potential for tsunamis and landslides based on Lombok earthquake data from January to October 2018. The dataset that included 10 monthly observations was divided into 70% training data (7 months) and 30% testing data (3 months). Each algorithm was trained using four input features: total monthly earthquakes, number of earthquakes with magnitude < 3, magnitude 3-5, and magnitude > 5.

Model Precision Comparison

The results of the assessment of the three machine learning algorithms showed a striking difference in performance. Table 1 shows the initial accuracy results for each model.

Table 1. Machine Learning Algorithm Prediction Accuracy

Model	**Accuracy**
Naive Bayes	0,6667 (66,67%)
Artificial Neural Network	0,5833 (58,33%)
KNN	0,5833 (58,33%)

Based on Table 1, Naive Bayes showed the most optimal results with an accuracy of 66.67%, while JST and KNN had the same accuracy of 58.33%. This difference in accuracy shows that probabilistic algorithms such as Naive Bayes are more efficient at managing datasets that are small in size and feature that tend to be independent compared to algorithms based on neural networks or distance-based.

Comprehensive Evaluation of the Model

In order to gain a better understanding of the model's performance, the evaluation was conducted using four metrics: Accuracy, Precision, Recall, and F1-Score. The findings of the thorough evaluation are shown in Table 2.

Table 2. Comprehensive Evaluation of Machine Learning Algorithms

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Naive Bayes	0,667	0,600	0,600	0,600
Artificial Neural Network	0,583	0,500	0,400	0,444
KNN	0,583	0,500	0,400	0,444

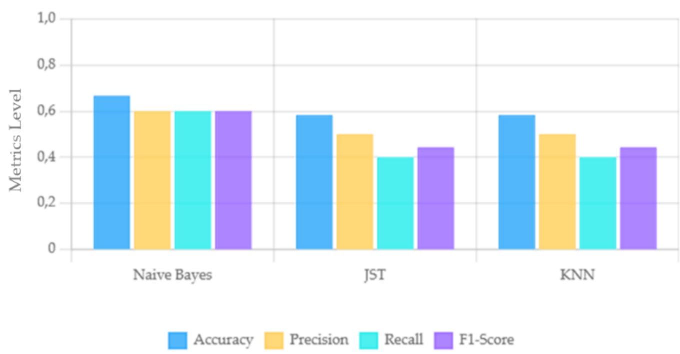


Figure 2. Comparison of Accuracy, Precision, Recall and F1-Score

Figure 2 shows the performance comparison of the four evaluation metrics for the three algorithms. This graph indicates that Naive Bayes consistently excels in all evaluation metrics, with a balanced accuracy and recall value of 60%, resulting in an F1-Score of 0.600. In contrast, JST and KNN showed the same performance with the lowest recall (40%), indicating the difficulty of both models in accurately detecting positive cases (tsunami/landslide events).

Further analysis using a confusion matrix provides an understanding of the types of errors made by each model. Naive Bayes shows a more optimal balance between True Positive (TP) and True Negative (TN), with lower False Negative (FN) levels compared to JST and KNN. This condition is crucial in the context of disaster prediction, where False Negative (the inability to detect events that actually occur) has a more dangerous impact than False Positive.

Higher recall in Naive Bayes (60%) compared to JST and KNN (40%) indicates that Naive Bayes can detect 60% of all tsunami/landslide events that actually occur, while JST and KNN are only able to detect 40%. In the context of early warning systems, the ability to detect actual events (high recalls) takes precedence over avoiding false alarms.

Discussion

Peforma Naive Bayes

Naive Bayes showed the best results with an accuracy rate of 66.67% and an F1-Score value of 0.600. The advantages of this algorithm can be outlined through several factors. First, Naive Bayes works well on small datasets because it doesn't require a lot of training data to estimate probability parameters. With only 7 training observations, Naive Bayes can calculate prior probabilities and likelihood efficiently using equations (1) and (2) [22].

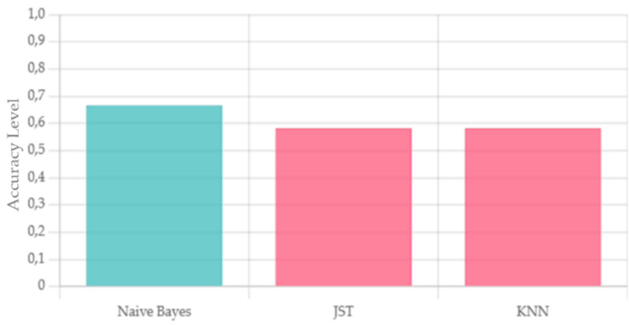


Figure 3. Comparison of the Accuracy of the Three Algorithms

Second, the assumption of conditional independence between features in Naive Bayes is in line with the nature of the applied earthquake data. Variables such as the frequency of earthquakes of different magnitudes tend to show less strong correlations, so the assumption of independence does not so much violate the existing data structure. Third, the probabilistic characteristics of Naive Bayes provide a better explanation in the context of predicting the uncertainty of catastrophic events.

These findings are in line with previous studies that indicate that probabilistic algorithms such as Naive Bayes are effective in classification in fields with high uncertainty and limited data. Equivalent precision and recall values (both 60%) indicate the absence of significant bias against the predictions of a particular class, which indicates a well-balanced model [23].

Artificial Neural Network Formation

JST recorded an accuracy of 58.33% with the lowest F1-Score (0.444) compared to the other three models. This sub-optimal performance can be explained through a number of limitations. First, the "ConvergenceWarning: Maximum iterations (1000) reached" warning indicates that the model has not reached the ideal convergence. The backpropagation process described in equations (3)-(6) requires many iterations to reduce the loss function, but with a small dataset (7 training observations), the model has difficulty finding consistent patterns.

Second, JST typically requires large datasets in order to learn complex feature representations. With only 4 input features and 7 observations, the neural network architecture is overfitting or underfitting. The low recall value (40%) indicates that JST tends to be conservative in predicting positive classes, possibly due to an imbalance in the learning of tissue weights.

Third, hyperparameters such as learning rate (α), number of hidden neurons, and activation functions require deep adjustment. On small datasets, the hyperparameter adjustment process becomes unreliable due to the limited validation set, so the obtained model is not optimal in generalization on the test data [24].

Peforma KNN

KNN showed the same results as JST (accuracy 58.33%, F1-Score 0.444). Although KNN is a simple and efficient algorithm, its suboptimal performance can be explained by the properties of the algorithm and data. First, KNN is an instance-based algorithm that relies heavily on the amount of training data. With only 7 training observations, the nearest neighbor search room using equation (7) became very narrow.

Second, KNN is sensitive to the size of features and the number of dimensions. Although it has been normalized with Min-Max Scaling, with only 4 features and limited data variation, Euclidean distances may not be able to distinguish meaningful patterns between observations. Third, the selection of the optimal k-value using cross-validation on a small dataset does not produce reliable results because each fold consists of only a few observations.

The low recall rate (40%) in the KNN indicates that the majority of votes using equation (8) tend to favor negative classes, possibly due to an imbalance in the distribution of classes in the training set or a less discriminatory distance between classes [25].

Implications for Earthquake Early Warning Systems

The results of this study show that Naive Bayes is the most appropriate algorithm to predict tsunami and landslide potential based on Lombok earthquake data with limited datasets. The accuracy of 66.67% can be considered a pretty good basis given the complexity of geophysical phenomena and the limitations of data. However, some improvements are needed for the implementation of an operational early warning system.

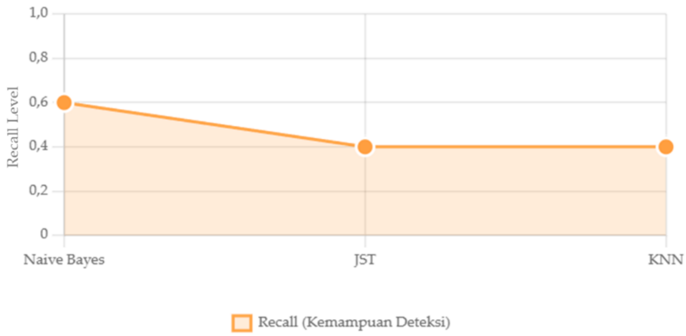


Figure 4. Comparison of Recall for Disaster Detection

Figure 4 shows a comparison of the recalls of the three models, which is an important measure in early warning systems. Recall measures the model's ability to detect real-world events – in the context of a disaster, it includes the ability to detect an actual tsunami/landslide. Naive Bayes with a 60% recall showed better performance than JST and KNN (40%). This 20% difference is especially important in the context of disaster risk reduction, as False Negatives can be fatal.

First, broader data collection is needed to improve model performance. Multi-year data with higher temporal resolution (such as weekly or daily) will provide more observations for training. Second, the addition of other geological features such as earthquake depth, epicenter coordinates, and fault properties can improve the model's predictive capabilities.

Third, given that Naive Bayes recall is only 60%, early warning systems need to be integrated with other real-time monitoring methods to reduce False Negatives. False Negatives in the context of disasters can be very fatal, therefore it is more advisable to use a hybrid system that integrates machine learning with traditional seismological sensors and analysis.

Research Litigation

The study has a number of limitations to be aware of. The very small size of the dataset (10 observations) hampered the model's ability for statistical generalization and validation. Second, the 70:30 data split resulted in only 3 observations for testing, which was not representative enough to assess the performance of the model as a whole. Third, this study only focuses on predictions of binary categories (Yes/No) without prioritizing variations in intensity or the likelihood of events.

Fourth, the variables used are limited to the quantitative nature of the earthquake without paying attention to geology, topography, and hydrological elements that also affect the potential for tsunamis and landslides. Fifth, the data range used (January-October 2018) is the year with extreme seismic activity in Lombok, so this model is likely biased towards high seismic conditions and cannot be generalized to normal situations [26].

Recommendations for Advanced Research

Based on the results and limitations of this study, a number of suggestions for the next study were proposed. First, the collection of earthquake data over several years (minimum 5-10 years) should be carried out to increase the number of observations and capture wider seismic variability. Second, it is necessary to explore ensemble algorithms such as Random Forest, XGBoost, or stacking that have proven promising results in other studies.

Third, the application of data augmentation techniques or synthetic data generation can be done to overcome the limitations of dataset size. Fourth, the application of deep learning with more sophisticated structures such as LSTM or Transformer can be evaluated if more data is available. Integration with geospatial data (GIS) and a fifth real-time sensor can improve the accuracy and real-time capabilities of the prediction system.

Sixth, the model needs to be validated in other areas with the same tectonic characteristics to test the transferability of the model. Seventh, the development of an AI system that can be explained and able to explain the reasons behind the prediction will increase stakeholder confidence in machine learning-based early warning systems.

Conclusion

The results of a study conducted on three machine learning algorithms to predict the potential for tsunamis and landslides in Lombok with earthquake data from January to October 2018 showed that the accuracy of machine learning measurements varied greatly among different algorithms. Naive Bayes shows the best performance with 66.67% accuracy, 60% precision, 60% recall, and 0.600 F1-Score, making it the best algorithm for disaster prediction with limited datasets. Meanwhile, the Artificial Neural Network (JST) and K-Nearest Neighbors (KNN) had the same performance with an accuracy of 58.33%, precision of 50%, recall of 40%, and an F1-Score of 0.444, demonstrating the challenges of both algorithms in managing small and high-complexity data. The difference in recall between Naive Bayes (60%) and JST and KNN (40%) is particularly important in the context of early warning systems, as the ability to detect real-world events of tsunamis and landslides takes precedence over avoiding false alarms. These findings confirm the hypothesis that probabilistic algorithms such as Naive Bayes are more appropriate for predicting earthquakes in Lombok with limited data characteristics, although the 66.67% accuracy still needs improvement through the addition of broader historical data, the integration of additional geological features, and the exploration of ensemble algorithms to optimize future earthquake early warning systems

References

[1]. N. T. Puspito and I. Gunawan, "Seismisitas dan Tektonik Kompleks Pulau Lombok," *Jurnal Meteorologi dan Geofisika*, vol. 20, no. 3, pp. 201–215, 2019.

[2]. A. Susanti, W. Wibowo, and S. Hartono, "Gempa Lombok dan Sumbawa 2018: Analisis Dampak dan Kerugian," *Jurnal Geologi Indonesia*, vol. 16, no. 2, pp. 145–160, 2019.

[3]. Badan Nasional Penanggulangan Bencana, "Laporan Kerugian Gempa Lombok 2018," Jakarta, Indonesia, Tech. Rep., 2018.

[4]. K. Sieh and D. Natawidjaja, "Neotectonics of the Sumatra Fault and Flores Back Arc Thrust," *Journal of Geophysical Research*, vol. 105, no. B12, pp. 28295–28326, 2000, doi: 10.1029/2000JB900120.

[5]. R. Harris, "Flores Back Arc Thrust: Seismic Activity Pattern in Eastern Indonesia," *Tectonophysics*, vol. 689, pp. 125–138, 2016, doi: 10.1016/j.tecto.2016.07.012.

[6]. M. Ramadhani, "Integrasi Clustering dan Ensemble Learning untuk Prediksi Magnitudo Gempa Bumi," *Seminar Nasional Informatika*, vol. 5, no. 1, pp. 45–52, 2024.

[7]. D. Schorlemmer, M. C. Gerstenberger, S. Wiemer, D. D. Jackson, and D. A. Rhoades, "Earthquake Likelihood Model Testing," *Seismological Research Letters*, vol. 78, no. 1, pp. 17–29, 2007, doi: 10.1785/gssrl.78.1.17.

[8]. T. Li, Z. Wang, and Y. Chen, "Improving earthquake prediction accuracy in Los Angeles with machine learning," *Scientific Reports*, vol. 14, art. 24440, Oct. 2024, doi: 10.1038/s41598-024-76483-x.

[9]. S. Malvia, P. Panwar, V. S. Rathore, and S. Sharma, "Early Earthquake Prediction Using Machine Learning Algorithm," in *Information Systems for Intelligent Systems. ISBM 2024*, C. S. In, N. S. Londhe, N. Bhatt, and M. Kitsing, Eds. Singapore: Springer, 2025, pp. 357–370, doi: 10.1007/978-981-96-1210-9_27.

[10]. J. A. Rosales Huamani, M. A. Rodriguez Melgarejo, and L. A. Lazo Pizarro, "A Systematic Review About the Use of Machine Learning Related to Earthquake Studies," *Advances in Civil Engineering*, vol. 2025, art. 4432234, May 2025, doi: 10.1155/adce/4432234.

[11]. T. Mitchell, *Machine Learning*. New York, NY: McGraw-Hill, 1997.

[12]. C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY: Springer, 2006.

[13]. I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. Burlington, MA: Morgan Kaufmann, 2016.

[14]. S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ: Pearson, 2009.

[15]. K. P. Murphy, *Probabilistic Machine Learning: An Introduction*. Cambridge, MA: MIT Press, 2022.

[16]. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY: Springer, 2009.

[17]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[18]. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

[19]. L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[20]. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco,

CA, USA, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[21]. K. M. Asim, A. Idris, F. Martínez-Álvarez, and T. Iqbal, "Earthquake prediction model using support vector regressor and hybrid neural networks," *PLoS ONE*, vol. 13, no. 7, art. e0199004, Jul. 2018, doi: 10.1371/journal.pone.0199004.

[22]. C. Zhang, H. Wang, and J. Zeng, "Earthquake prediction based on deep learning: A review," *IEEE Access*, vol. 7, pp. 47462–47472, 2019, doi: 10.1109/ACCESS.2019.2909298.

[23]. B. Arunadevi, M. M. M. I. Hussain, R. Lakshmi, R. MM, and K. Sengupta Das, "Risk Prediction of Earthquakes using Machine Learning," in *3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, 2022, pp. 1589–1593, doi: 10.1109/ICESC54411.2022.9885674.

[24]. N. B. Jarah, A. H. H. Alasadi, and K. M. Hashim, "A New Algorithm for Earthquake Prediction Using Machine Learning," *Journal of Computer Science*, vol. 20, no. 2, pp. 150–156, 2024, doi: 10.3844/jcssp.2024.150.156.

[25]. P. Shrote, P. Dasarwar, and S. Dongre, "Earthquake Prediction through Machine Learning Approach," *International Journal of Advanced Technology & Research*, vol. 15, no. 2, pp. 85–92, 2024.

[26]. D. B. Babu, M. L. N. Revathi, and M. Senthil, "Earthquake Prediction Model using Random Forest & Gradient Boosting Algorithms," *Journal of Engineering Science & Technology*, vol. 19, no. 1, pp. 112–125, 2024.