Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.98549805

MACHINE LEARNING APPROACHES FOR EARTHQUAKE FORECASTING: A REVIEW

Manoj Kollam and Ajay Joshi
Department of Electrical & Computer Engineering, The University of West Indies, St.Augustine, Trinidad and Tobago
Correspondence should be addressed to Manoj Kollam; mkollam@gmail.com

Abstract
Seismology is crucial for saving as many lives as feasible through earthquake prediction. In order to better analyse and interpret seismic data, a precise and accurate algorithm is required to detect and locate future earthquakes. This paper aims to review the recent papers that have investigated the application of different machine learning (ML) approaches for earthquake forecasts, with the goals of grouping the utilised methods and analysing the major trends in earthquake forecasting and the different seismic indicators used in machine learning algorithm and the performance of the methods were reported. In this regard, we have analysed the different ML Algorithms used in different research paper considered in this paper for estimating magnitude, location, time of earthquakes and recommended an appropriate research approach.

Introduction
The study of earthquake prediction dates back to the late 1800s. Earthquakes is a natural calamities that cause loss of life, destruction to infrastructure, and significantly influence a country's economy. Every year, hundreds of people are killed by earthquakes all around the globe. Typically, earthquakes result when the tectonic plates abruptly slip and release energy. These devastating situations may be mitigated with earthquake prediction. Multiple computer algorithms have been used to predict earthquakes, but precise earthquake prediction (time and location) remains challenging. In earthquake prediction, ML-based methodologies and algorithms have demonstrated promising results over the previous decade. In comparison to conventional prediction approaches with significant false alarm rates, ML-based algorithms are more likely to successfully forecast earthquakes with greater precision [1]. This paper aims to understand and analyse seismic indicators used in antecedent research on earthquake forecasting using the Machine learning algorithms.

Survey of Different Methodology
In Yousefzadeh [2] study, they collected the data from USGS & IIEES over a period of 1973 to 2019. 16 seismic precursors and density parameter derived using a Kernel Density Estimation function (KDE) were utilised to create the models. During model building, both SNN and DNN were tuned to offer a elevated level of generalisation whereas minimising overfitting with the help of Decaying of weights and Dropout parameters were applied to the SNN and DNN, respectively. The models considered in this study, modified the number of nodes and layers used, activation function, dropout rates, and weight deteriorating. They are

regularly tweaked, trained, and evaluated to guarantee that they deliver high-quality results. For SVM, RBF kernel calibration provides the most accurate results. By cycling through several input values, they were able to generate parameter and width of kernel. During the DT calibration operation, the trial parameter, which helps in regulating the number of boosting rounds, was adjusted. According to the results, both DNN and SVM performed satisfactorily when it came to predicting high magnitude classes, with a sensitivity of 95.5% and 97.7%, respectively. Nevertheless, DT showed greater promise in its performance, obtaining a perfect accuracy rate when dealing with events of varying magnitudes.

Using spatial and temporal correlations and an artificial Neural network algorithm Long Short-Term Memory (LSTM) was considered by Wang [3] to possibility of foreseeing earthquakes. In order to find connections in space and time, the researchers made the data into matrix having two-dimensional structure from different regions with the same timestamps. This matrix was then fed into the LSTM layers. There were 128 neurons in the LSTM network's hidden layer, 256 in the first dense network layer, and 64 in the second. The use of SoftMax activation function and RMSprop Optimisation had helped the model to improve its performance. The model's efficiency was 63.50 % when using single dimensional matrix as a input (temporal correlation) & 87.59% when employing two-dimensional data (spatiotemporal correlation) with a decomposition strategy.

To anticipate earthquakes [4] developed an artificial neural network (ANN) model applying automated cluster- ing approaches. The study employed data from the US Geological Survey (USGS) and Meteorology, Climatology, and Geophysical Agency of Indonesia (BMKG), including 82,850 earthquake occurrences in Indonesia from 1910 to 2017. Seven seismic factors were added as input variables in the model. The clustering technique included subdividing the massive dataset into more manageable groups. This procedure consisted of three steps: first, deter- mining the optimal number of clusters by using a algorithm of valley tracing and hill-climbing, which determined that six was the best number of clusters. Using a hierarchical K-means method, the data was then separated into six groups. Ultimately, predictions of aftershocks were created using a neural network. According to the findings, the model performs better with an accuracy of 56 to 72% for earthquakes with a magnitude that is greater than six.

Cheraghi and Ghanbari (2017) [5] devised backpropagation (BP) algorithm functioned as the neural network's learning strategy, while mean squared error(MSE) is used to calculate the error. An ANN comprised of two hidden layers, the first of which included three neurons and the second of which contained two neurons. A sigmoid activation function introduced network nonlinearity. The calculated magnitude was used for energy projection by the researchers. The study employed an ANN to forecast the timing and earthquakes magnitude. The maximum error observed is 3.5%, while the average error for magnitude prediction was 0.5%. The prediction of earthquake timing exhibited an error margin of 10 days.

In order to find trends in the past that could indicate an impending earthquake of great size, [5] employed a tree- based algorithm for seismic characteristics calculated by using Gutenberg-Richter Law, b-value were considered to analyse earthquake patterns. Clustering, grouping, antecedent tree building, pattern extraction, and pattern selection were all part of the model's training procedure. For this clustering task, we used the K-means technique. From this cluster, the best K parameters were selected. With a tree-based algorithm, a search was conducted for earthquake precursor patterns in the considered data including great magnitude occurrences. The model accurately forecast an earthquake the following day and reached the maximum level of precision (93.59%) for the Santiago dataset.

Murwantara [6] conducted a comparative analysis of ML algorithm effectiveness in forecasting the seismic activities ranging from medium to long-term by employing multiple linear regression (MLR), support vector ma- chines (SVM), and Naive Bayes (NB). To predict earthquake magnitude, process is divided into two parts. First, earthquake power was determined based on latitude and longitude, and subsequently, magnitude was predicted using coordinates and depth information. Earthquake depth prediction was achieved by factorising the inverse of the magnitude prediction. The results revealed that SVM performed better than the other approaches in forecasting earthquake magnitude for a 30-year dataset with groupings. The prediction accuracy, as measured by a mean absolute error (MAE) of 0.598473, demonstrated that the SVM method produced more exact earthquake prediction results than the other techniques.

To forecast the seismic events, [7] utilised numerous machine learning (ML) techniques, including Bayesian networks (BN), regression trees (RT), simple logistic (SL) regression, random forests (RF), logistic model trees (LMT), ZeroR, and logistic regression (LR). Using nodes and directed edges, Bayesian networks may graphically describe the relationships between dependent and independent variables. The building blocks of a Bayes network include probabilistic distributions, conditional links, and random variables. Similar to linear regression, simple logistic regression employs a nominal dependent characteristic. The objective was to consider the likelihood that a particular nominal attribute value is connected to quantified. The tree based Logistic Model combines decision trees with logistic regression, expanding on the information provided by the prior tree-based algorithm. It's a decision tree where the branches represent different linear models and the leaves represent piecewise linear regression. LogitBoost creates Logistic regression(LR) model by using a every single node of a tree. ZeroR concentrates on the end attribute and disregards all other variables, generating predictions based on the majority class. ZeroR sets a baseline performance for comparison with other categorisation algorithms despite missing predictive ability. LR is utilised when the variable of interest is categorical. It is a technique for forecasting based on probability models and a type of regression model with the use of sigmoid activation function. Typically, the logistic regression hypothesis falls between 0 and 1. Logical regression includes determining the parameters of a logistic model, similar to linear regression. The outcome shows that SL accomplish with an accuracy of 99.94%.

In a study conducted by Asim [8] used four artificial neural network for forecasting earthquakes of magnitude greater than 5.4 magnitude based on a variety of seismic precursors. Eight seismic indicators were used to predict earthquakes, each of which reflected some aspect of the seismic situation or ground potential. In order to train the Piecewise Recurrent Neural Network(PRNN) algorithm quickly, the Levenberg-Marquardt backpropagation (LMBP) method was utilised instead of the standard backpropagation (BP) method. The PRNN network consisted of two hidden layers of 12 neurons, with the first layer employing the tan-sigmoid transfer function and the second layer of the log-sigmoid transfer function. Similar to the PRNN algorithm, but with six and seven neurons in each of the two layers with the capacity to store internal state as a directed cycle between units [9] . The Random Forest (RF) method comprised of 50 decision trees, whereas Linear Programming Boosting(LPBoost) was a mixture of several tree classifiers that were linearly added. The study of the performance of the four algorithms found that LPBoost had the best accuracy rate at 65%, followed by RNN at 64%. In contrast, PRNN's 58% accuracy rate was the lowest, but it created the fewest number of false alarms.

Asim [10] researched on the classification system as SVR and Hybrid Neural Network (HNN) to forecast magnitude largerthan 5, using USGS data for a period of 1980 to 2016 from Hindukush, South California, and Chile. Sixty different seismic features are computed for each earthquake occurrence. A two stage feature selection methodology was adopted, starting with the Maximum Relevance and Minimum Redundancy (MRMR) methods to identify the important informative features were chosen as a input to the SVR model to predict the trend, which was then passed along with features to the subsequent step to the predictive model, the Artificial Neural Network (ANN). The outcome of every ANN layer is given to the next ANN layer, along with the weight correction are then passed on from the previous ANN layer. The inclusion of Enhanced Particle Swarm Optimisation (EPSO) is considered for the Optimization of weights in ANN, which could become trapped in local minima, resulting in suboptimal solutions.

In their research, Khawaja M [11] suggested a robust earthquake prediction system, (Earthquake Predictor)EP- GPBoost, which utilises a blend of two algorithms to develop an ensemble algorithm capable of predicting earth- quakes 15 days prior to their occurrence. The dataset was obtained from USGS for years of 1980 to 2016. Boosting was used in the creation of the Genetic Programming (GP) algorithm, which uses many GP strings inside a class to function as a unified classifier. Boosting was employed for weight update during the evolution of the GP. For each class, P number of GP programs was generated using boosting, and the outputs of GP strings generated for each class were summed together. The class label for a test instance was determined by selecting the larger value from a weighted total of the generated results [9]. The EP-GPBoost performed really well across the board, especially with regards to reducing false positives. There was a small number of false positives reported in each of these regions: 74% in Hindukush, 80% in Chile, and 84% in Southern California. In addition, the model was successful in predicting weather patterns in Hindukush (with an accuracy of 78.7%), Chile (84.5%), and Southern California (86.6%).

Majhi et al. (2020) [12] suggested a novel earthquake magnitude prediction machine learning models using a Functional Link Artificial Neural Network (FLANN) with a Moth Flame optimization algorithm. The study em- ployed a dataset from USGS, covering global earthquake events, with 6 seismic precursors serving as input features. The FLANN model had no hidden layers, with nonlinearity achieved through the nonlinear function. Standard back propagation, least-square & gradient descent optimization algorithm, Levenberg-Marquardt backpropagation (LMBP), and Moth Flame optimization(MFO) were employed as learning algorithms, and their performances were evaluated to determine the optimal weight for the model. The data set was filtered to include only earthquakes with magnitudes greater than 5.5, and the time and date attributes of these events were merged into a single attribute. There was a general expansion and standardisation of all attributes. The proposed MFOFLANN model achieved a Root Mean Squared Error (RMSE) of 0.0565.

Using data from Asia tectonic plate, Bhandarkar [13] compared the accuracy of a Feed-forward Neural Network (FFNN) with that of a Long Short-Term Memory (LSTM) in predicting

earthquake trends. All of the nodes in the FFNN model had a sigmoid activation function, and the model's hidden layers included either 20 or 60 neurons. As a learning rule, we used the RMSprop algorithm. Nevertheless, the LSTM model used 15 steps of backpropagation in time and 40 hidden units per LSTM cell. A dropout layer was added in between the two hidden layers to reduce the likelihood of overfitting. To lessen the RMS loss, the Agadrad algorithm was used. With an R-score of -0.252, the LSTM model outperformed the FFNN model by 59%.
Using unbalanced classification methods and ensemble learning, Fernández-Gómez [14] presented a novel ap- proach to forecasting high-magnitude earthquakes within five days' notice. The analysis looked at seismic activity data that was skewed in four different ways, focusing on the towns of Santiago, Valparaiso, Talca, and Pichilemu in Chile. The pre-processing of data is done and then by using a classification algorithm for finding reliable pa-rameters. All classifiers developed up to this point were categorised as "basic" classifiers. Predictive accuracy was improved by selecting high-performance algorithms and combining them with a number of pre-processing tech- niques. The second step involved creating ensembles from the top basic classifiers to further boost their efficiency. Until all simple classifiers were chosen, or a perfect ensemble was established, the ensembles were created repeat- edly. The suggested model performed quite well; in Talca, it attained a Matthews Correlation Coefficient (MCC) of 0.96.

Corbi [15] used the Gradient Boosted Regression Trees (GBRT) technique to foretell the onset and magnitude of earthquakes on a small scale in the lab. Similar to real-world subduction zones, the research generated earth- quakes with magnitudes between 6.2 and 8.3 by simulating earthquakes in laboratory. In order to simulate a large earthquake in the lab, a gelatine wedge was used as the overriding plate, and a thick, stiff plate diving at 10 de- grees was used as the subducting plate. Two patches of equal size and friction were incorporated into the analogue megathrust, with a patch of increased velocity in between to approximate asperities. In contrast to its action on plastic contacts, gelatine slowed down when placed on sandpaper. At first, the model exhibited stick-slip action, with periods of stress building punctuated by the earthquake-like the development of frictional destabilisation at the gelatine-plate interface. The model generated cracks across either a single or double asperity in proportion to the length of the barrier relative to the asperities. Time to failure predictions were made by the GBRT model using information from a laboratory mode termed geodetic signals (TTF). Surface deformation may be defined by these 94 features. Using N seismic data and single consecutive cycles for time-to-first-fix, the model was trained individually at nine target locations parallel to the trench (TTF). Each target point had an R score between 0.7 and 0.8, which is encouraging.

After simulating an earthquake in the lab using a shear experiment, Rouet-Leduc [16] evaluated the data to determine how much time was left until the shear experiment failed. Fault gouge material was used in the experi- ment, and it was subjected to double direct shear along a two-fault configuration. When the gouge layer widened and deepened, the shearing block moved following a stick-slip frictional failure (lab earthquake). Rapid sonic emissions were produced as a result of several minor shear failures that occurred when the materials were about to collapse. When the gouge crumbled, the laboratory earthquake's shear stress and friction reduced rapidly, end- ing the unstable condition.The prediction model was built with information from a continuous acoustic time-series recorded at the fault. At each time frame, the RF model takes a weighted average of many decision trees, each of which makes a series of judgements based on statistical data. Around one hundred potentially important statistical characteristics were computed for each time interval, and the features employed in this study were chosen recur- sively from that set. The R-score for the Random Forest model, with a decision trees of 1000, was 0.89, indicating that it accurately anticipated the time to failure.

An electromagnetic signal-based earthquake magnitude prediction technique was proposed by Bao [17]. This algorithm makes use of Convolutional Neural Networks. To anticipate earthquakes on the 28th day, the study analysed data gathered every 27 days from January 1, 2017 through January 1, 2021 (a total of 6936 samples). To build a robust connection between the sensed data and seismicity, the proposed model incorporates a High- Dimensional-Feature-Extraction (HDFE) block that accepts a 3D feature matrix as input and a Temporal Correlation block that uses four convolution units. Imbalanced samples were improved by the application of noise modelling and SMOTE oversampling in this study.The proposed inductive electromagnetic sensor data was heavily utilised in the model assessment. The results show that the CNN model can accurately forecast the magnitude of earthquakes with a 97.88% success rate.

Seismic Precursor

Any observable geological or geophysical phenomena that happen before an earthquake and might offer hints or warnings that an earthquake is imminent is considered a seismic precursor. Understanding the underlying physical processes and creating more accurate and trustworthy forecasting models may be facilitated by identifying and studying seismic precursors, making them an integral aspect of earthquake monitoring and prediction efforts.

Time T t is one of the criteria that is defined as the time interval between the most recent L occurrences, where

L can be any number, and tt is the occurrence of earthquake time.

$$T^{t} = t^{t}{}_{L} - t^{t}{}_{1}$$

Time T' represents how frequently a foreshocks occurred before the months actual earthquake occurrence. The average magnitude of the most recent n occurrences is the second seismic indication taken into account. It correlates with the magnitudes of foreshocks because seismic activity's magnitude M' rises before a larger earthquake.

$$M^{t}{}_{mean} = \frac{\sum_{j} M^{t}}{L}$$

The seismic energy release rate, or SER, is an extra seismic indicator that may be used to link seismic activity. Low-magnitude seismic events signify the gradual release of seismic energy from faults; but, if this phenomenon is disrupted, it may result in a significant seismic event. The square root of the emitted seismic energy is depicted in Eq. (3)

$$\sqrt{SER} = \frac{\sum \sqrt{(10^{11.8+1.5M^{t}})}}{T^{t}}$$

The erg unit of measure for the energy released and derived from Eq. 3 is used. Another important seismic component known as the b value is produced by the well-known GR inverse power law rundle. It is described the incline of line connecting the magnitude of an earthquake to the log of its frequency of occurrence. In Eq. 4, the formula for the Gutenberg-Richter (GR) curve is provided, and the Least Squares method is used to determine the parameters a and b.

$$logttR = x - yN$$

Another seismic indication is the actual data's deviation from the GR inverse power curve (Eq. 4). As per Eq. 5 illustrates the mean square deviation's total.

Lower mean square deviation, higher data conformity, and increased likelihood of inverse power law prediction is all indicators of data quality.

$$\eta = \frac{\sum (logttR - x - yN)^2}{L - 1}$$

Another seismic indicator is the deviation between the highest recorded and highest occurring earthquake magni- tude. The catalogue contains the maximum observed event and Eq. 6 yields the maximum anticipated event.

$$Max_{AE} = \frac{x}{y}$$

Where an is the inverse power law's y-intercept, which may be found in Equation 4. Also taken into account as a precursor to seismicity is the time between distinctive occurrences (l) among the most recent events. The magnitude of a collection of related occurrences is referred to as a characteristic magnitude. For instance, the mean duration between two occurrences of magnitude 4.5–5.5 is determined using Eq. (7) and is referred to as the typical magnitude.

$$u^{t} = \frac{t^{t}_{char}}{l_{char}}$$

The goal would be for the mean time between typical events to be equal. Similar to how departure from this mean time mt is used, as per Eq. (8) seismic parameter is also taken into account.

$$m_t = \frac{a^{t}_{t}}{u^{t}}$$

where σt is the observable time's standard deviation.

The table 1 shows the different data sets used and input feature and the outcome discussed in considered work done by the different research papers.

Evaluation Metrics

Evaluation metrics performance for predictive models and algorithms are a fundamental concept in machine learning, data science, and statistical modelling, serving as the definitive measurement instruments. They allow practitioners to quantify the efficacy of a model and guide its enhancement by identifying its weak points. Depend- ing on the type of task, the data, and the business context, various models necessitate different metrics. Classifi- cation Metrics and Regression Metrics are two broad classifications for these metrics. In classification problems where the output is a categorical variable, classification metrics include, among others, accuracy, precision, recall, F1 score, and area under the receiver operating characteristic (ROC) curve. When the output variable is a real or continuous value, regression metrics include, among others, mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R2 score. Each of these metrics has its own strengths and weak- nesses, and the choice of which metric to use depends on the specific needs of the analysis. For example, if the focus is on accurately predicting small deviations from the mean, then metrics such as MSE and RMSE may be more appropriate. On the other hand, if it is more critical to accurately predict large deviations from the mean, MAE may be a better metric. Additionally, R-squared (R2) is typically used to assess the overall fit of the model, with higher values indicating a better fit. Moreover, Mean Percentage Error (MPE) is useful for determining the average magnitude of errors in percentage terms. Overall, selecting an appropriate regression model evaluation metric depends on the specific requirements of the analysis, and a combination of metrics may be necessary for a comprehensive evaluation, particularly when comparing multiple regression models. Both metrics evaluate the average magnitude of error in a case of predictions, with MAE less susceptible for the outliers and MSE more sensitive to them. Generally, MAE and MSE should have a lower value and higher values for R2suggest how good

Table 1: Overview of data Type and there input and output features

Type of Data used	Input to the model	Output
Lab generated Data	Acoustic time series [17]	Time to failure
USGS	Time, location, magnitude[3]	Earthquake event
USGS & BMKG	T, M_{mean}, $dE^{1/2}$, β, η, $M_{expected}$, μ, b-value, GR's law[4]	Estimating the of aftershock occurrence
USGS	Time, date, location, magnitude and depth[7]	Quake location, depth, magnitude
USGS/IIEES	a and b value, T, M_{mean}, $dE^{1/2}$, β, η,fault density $M_{expected}$, μ, c, ΔM, probability of occurrence, location, depth [2]	Earthquake magnitudes
USGS	Time, latitude, longitude, depth[8]	Earthquake magnitude with different ranges
USGS	T, M_{mean}, $dE^{1/2}$, β, η, $M_{expected}$, μ, c[9]	Magnitude greater than 5.4
USGS	T,a and b value, M_{mean}, $dE^{1/2}$, β, η, $M_{expected}$, μ, c, T_r, z, probability of quake occurrence,[10]	Magnitude more than 5.0
USGS	T, z, M_{mean}, $dE^{1/2}$,β,η,$M_{expected}$, μ,probability of occurrence, c, T_r, max magnitude, [11]	15 days before the earthquake
Lab generated Data	Geodetic signals [15]	Time and magnitude
USGS	Date, time. location, depth, magnitude	Earthquake magnitudes
Not stated	T, a and b value, M_{mean}, b-value, $dE^{1/2}$, β, η, $M_{expected}$, μ, M_{def}, probability of quake occurrence,c 6]	Next day earthquake occurrence
Not stated	Time, focal distance, fault distance, magnitude[5]	Time and magnitude of earthquake
Not stated	seismic occurrences	Earthquake trend
Not stated	Imbalanced datasets of seismic events [14]	5 days window
Electromag-netic sensor	51 features of electromagnetic disturbances[17]	Earthquake magnitude

the model. Table 2 provides a summary of the employed algorithms, their efficacy, and there evaluation metrics used in the previous section.

$$S_n = \frac{TP}{TP + FN}$$

$$S_p = \frac{TN}{TN + FP}$$

$$PPV = \frac{TP}{TP + FP}$$

$$NPV = \frac{TN}{TN + FN}$$

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TN + FN)(TP + FP)(TP + FN)(TN + FP)}}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$R = \frac{(TP * TN) - (FP * FN)}{(TP + FN) * (TN + FP)}$$

Table 2: Different algorithms vs performances with Evaluation metrics.

Algorithm	**Evaluation Metrics**	**Performance**
RF[16]	R-score	0.89
LSTM[3]	Accuracy	87.59%
ANN[4]	Accuracy	56% - 72%
Multiple Linear Regression, Support Vector Machines, Naive Bayes[7]	MAE	0.5985
DNN, SNN, SVM, DT[2]	Accuracy & Sensitivity	100% & 97.7%
BN, RT, SL, RF, LMT, ZeroR, LR[8]	Accuracy	99.933%
PRNN, Random forest, RNN and LPBoost[9]	Accuracy	65%
SVR-Hybrid Neural Network[10]	Accuracy	90.63%
GP-Adaboost[11]	Accuracy	86.62%
GBRT[15]	R-score	0.7 - 0.8
MFOFLANN[12]	RMSE	0.0565
Tree based algorithm[6]	Accuracy	93.69%
Artificial neural network[5]	Max_{error}	3.5%
FFNN, LSTM[13]	R-score	59%
Not stated[14]	MCC	0.96
CNN[17]	Accuracy	97.88%

Discussion

The algorithms utilised in the research have shown promising performance overall, with R-scores near 1.00, high accuracy, and low error. When applied to Iranian earthquake data and input characteristics, the Decision Trees algorithm attained a perfect accuracy of 100% [2] but when the different dataset is used the accuracy has dropped. In addition to various algorithms in this review, there are a number of other high-performance options that achieve accuracy rates of 90% or greater. Therefore, while attempting to forecast earthquakes, it is important to give considerable thought to the input characteristics, output variables, and prediction method chosen. For instance, Shodiq and colleagues (2018), Khawaja M., Asim et al. (2020) all utilised ANN for earthquake prediction but came to different conclusions. Underfitting can occur if the training dataset is too little, hence dataset size matters.

Conclusions

Researchers have looked into ways to lessen the devastation caused by earthquakes. In terms of earthquake fore- casting, machine learning (ML) has proven to be more accurate than previous approaches. Despite the widespread application of ML algorithms, no one is well suited to the wide variety of prediction issues that exist. This study summarises the characteristics, prediction variables, and algorithm performances of existing ML-based approaches for earthquake prediction. The study's primary objective is to provide scholars with a foundational understanding of the significance of ML in earthquake prediction. Although ANN has been successful, other algorithms have outperformed it. Suitable for forecasting the numerous sorts of earthquake prediction challenges, a hybrid model constructed from a combination of different ML algorithms can enhance results by helping tune, generalise, or adapt to new jobs.

References

[1] Mha Banna, K A Taher, M S Kaiser et al. "Application of Artificial Intelligence in Predicting Earthquakes: State-of-the-Art and Future Challenges". IEEE Access, vol. 8, 2020.

[2] M Yousefzadeh, S A Hosseini and Farnaghi M. "Spatiotemporally explicit earthquake prediction using deep neural network". Soil Dynamics and Earthquake Engineering, vol. 144, 2021.

[3] Q Wang, Y Guo, L Yu and Li P. "Earthquake Prediction Based on Spatio-Temporal Data Mining: An LSTM Network Approach". IEEE Transactions on Emerging Topics in Computing, vol. 8, 148–158, 2020.

[4] M N Shodiq, D H Kusuma, M G Rifqi, A R Barakbah and Harsono T. "Neural Network for Earthquake Predic- tion Based on Automatic Clustering in Indonesia". JOIV : International Journal on Informatics Visualization, vol. 2, 2018.

[5] E Florido, G Asencio-Cortés, J L Aznarte, C Rubio-Escudero and F Martínez-Álvarez. "A novel tree-based algorithm to discover seismic patterns in earthquake catalogs". Computers & Geosciences, vol. 115, 96–104, 2018.

[6] I M Murwantara, P Yugopuspito and Hermawan R. "Comparison of machine learning performance for earth- quake prediction in Indonesia using 30 years historical data". TELKOMNIKA (Telecommunication Computing Electronics and Control, vol. 18, 2020.

[7] P Debnath, P Chittora, T Chakrabarti et al., 2021.

[8] K M Asim, F Martínez-Álvarez, A Basit and T Iqbal. "Earthquake magnitude prediction in Hindukush region using machine learning techniques". Natural Hazards, vol. 85, 471–486, 2016.

[9] N S M Ridzwan and S H M Yusoff. "Machine learning for earthquake prediction: a review". Earth Sci Inform, vol. 16, 1133–1149, 2017.

[10] K M Asim, A Idris, T Iqbal, Martínez-Álvarez and F. “Seismic indicators based earthquake predictor system using Genetic Programming and AdaBoost classification”. Soil Dynamics and Earthquake Engineering, vol. 111, 2018.

[11] K M Asim, A Idris, T Iqbal, Martínez-Álvarez and F. “Earthquake prediction model using support vector regressor and hybrid neural networks”. PloS one, vol. 13, 2018.

[12] S K Majhi, S S Hossain and Padhi T. “MFOFLANN: moth flame optimized functional link artificial neural network for prediction of earthquake magnitude”. Evolving Systems, vol. 11, 45–63, 2019.

[13] T Bhandarkar, K V Satish, N Sridhar et al. “Earthquake trend prediction using long short-term memory RNN”. International Journal of Electrical and Computer Engineering (IJECE), vol. 9, 2019.

[14] M J Fernández-Gómez, G Asencio-Cortés, A Troncoso and F Martínez-Álvarez, 2017.

[15] F Corbi, L Sandri, J Bedford et al. “Machine Learning Can Predict the Timing and Size of Analog Earth- quakes”. Geophysical Research Letters, vol. 46, 2019.

[16] B Rouet-Leduc, C Hulbert, N Lubbers et al. Machine Learning Predicts Laboratory Earthquakes. Geophysi- cal Research Letters, vol. 44, 2017.

[17] Z Bao, J Zhao, P Huang, S Yong and Wang Xa. “A Deep Learning-Based Electromagnetic Signal for Earth- quake Magnitude Prediction”. Sensors, 2021.