



e-ISSN: 2278-8875
p-ISSN: 2320-3765

International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering

Volume 15, Issue 5, May 2026



Impact Factor: 9.867

9940 572 462

6381 907 438

ijareeie@gmail.com

www.ijareeie.com



Predictive Modeling for Earthquake Damage using Machine Learning

R.Gayathri, MCA, M. Phil., NET, A Ramani

Assistant Professor, Department of MCA, Vivekanandha Institute of Information and Management Studies,
Tiruchengode, Namakal, Tamil Nadu, India

Department of MCA, Vivekanandha Institute of Information and Management Studies, Tiruchengode, Namakal,
Tamil Nadu, India

ABSTRACT: Earthquakes pose a significant threat to human life and infrastructure, especially in seismically active regions. Over the past few decades, numerous attempts have been made to enhance the accuracy and timeliness of earthquake forecasting and early warning systems (EWS). Traditional approaches rely heavily on geological, seismological, and statistical models that often fail to deliver reliable predictions. The rise of deep learning has introduced a paradigm shift by enabling models to extract complex spatiotemporal features from seismic signals, satellite data, and geophysical indicators. This paper reviews state-of-the-art deep learning approaches for earthquake forecasting and EWS, including convolutional neural networks (CNNs), long short-term memory (LSTM) networks, generative adversarial networks (GANs), and AIoT integrations. We analyze twenty recent studies, highlighting their methodologies, advantages, and limitations. Furthermore, comparative study, recent technical challenges, and future research directions are presented. The findings emphasize that deep learning models significantly improve real-time detection, blind zone reduction, and public safety response mechanisms. However, challenges like data quality, model generalization, and real-time deployment persist. This review aims to provide researchers, seismologists, and policymakers with a comprehensive understanding of the current state and future scope of AI-powered earthquake prediction systems to mitigate natural disaster impacts effectively.

KEYWORDS: Earthquake Forecasting, Machine Learning, Early Warning Systems, Seismically Active Regions, AI in Geophysics

I. INTRODUCTION

China has frequent continental earthquakes and severe earthquake disasters. The Chinese mainland represents one of the most seismically active regions within a plate interior and is characterized by frequent earthquakes, widespread distribution, high intensity, and severe disasters. Mitigating the disasters caused by earthquakes holds significant practical importance for the national economy, population, and society as a whole. Earthquake prediction has long been a formidable challenge, attributable to several factors. First, earthquakes result from the complex interactions between tectonic plates, faults, and other geological factors, making precise prediction of their occurrence time and magnitude extremely challenging. Second, the scarcity of long-term and extensive data poses a significant obstacle to earthquake prediction. The recurrence intervals of major earthquakes are often lengthy (spanning centuries to millennia) and identify trends and patterns over such extended timescales is difficult. Owing to complex geophysical interactions, variability in seismic activity, insufficient comprehensive data, and limitations in current technological approaches, earthquake prediction remains a complex and arduous research domain [1].

In seismological research, regional seismic activity may exhibit anomalies such as seismic gaps, seismic belts, and earthquake swarms, which reflect the regional stress state. Case studies have demonstrated that enhanced seismic activity commonly precedes major earthquakes, with seismic activity anomalies being widely used as indicators to judge future seismic trends in a region. In the study of the temporal, spatial, and intensity distribution characteristics of seismic activity, Gutenberg and Richter [2] first proposed the magnitude-frequency relationship, known as the Gutenberg–Richter (G-R) law.

With the widespread adoption of machine learning techniques as indispensable tools, seismologists have gradually explored the application of machine learning algorithms in seismic activity prediction. In spatial regional prediction, Aslam et al. utilized clustering hotspot analysis machine learning methods to identify seismic hotspot regions in

northern Pakistan [3]. For time series prediction, Kaftan et al. validated the performance of different neural networks, including multilayer perceptron neural networks (MLPNN), radial basis function neural networks (RBFNN), and adaptive neuro-fuzzy inference systems (ANFIS), in seismic time series applications [4]. Wang et al. incorporated spatial information into time series data and utilized long short-term memory (LSTM) networks to discover spatiotemporal correlations among earthquakes [5]. Additionally, raw earthquake catalog data were extracted into seismic activity characteristic parameters, transforming them into input patterns recognizable by machine learning models (such as neural networks and support vector machines). Seismic activity characteristics refer to statistical parameters obtained through the analysis of seismological observational data, which are capable of quantitatively describing the spatiotemporal intensity characteristics of seismic activity within a certain spatiotemporal range. These include statistical parameters describing the temporal distribution of earthquake sequences, spatial distribution of earthquakes, average seismic activity intensity, rate of seismic strain release, and average recurrence period of earthquakes [6]. Panakkat and Adeli [7] first used these seismic activity characteristic parameters as inputs to construct a neural network model for predicting the maximum earthquake magnitude over a subsequent period. Then, they employed methods such as RBFNNs, recurrent neural networks (RNN), and probabilistic neural networks (PNN) to establish nonlinear models between earthquake magnitudes and seismic activity indicators [8]. This was later extended to using neural network methods to predict both the occurrence time and magnitude of earthquakes [9]. Other scholars have made improvements based on these seismic activity parameters. First, additional characteristic parameters were added, including calculating the b-value increase over certain earthquake intervals, using probability density functions to record the probability of magnitudes greater than or equal to the target magnitude, and recording the maximum magnitude in the most recent week [10,11]. Second, improvements have been made in terms of methodologies, including ensemble algorithms which integrate neural network-based algorithms, genetic programming (GP) algorithms, and Adaptive Boosting (AdaBoost) algorithms [12]. Additionally, the dendritic cell algorithm in the field of artificial immunity [13] and deep learning-based approaches [14] have been introduced. Asencio-Cortés et al. evaluated the application of generalized linear models (GLMs), gradient boosting machines (GBMs), and random forests (RFs) in earthquake magnitude prediction [15]. For earthquakes with magnitudes ranging from 3 to 7, the RF method demonstrated the best performance, achieving a mean absolute error (MAE) of 0.6. Mallouhy et al. employed RFs, naive Bayes, logistic regression, MLPNNs, AdaBoost, K-nearest neighbors, support vector machines, and classification and regression trees to predict major earthquakes [16]. Among these methods, RFs achieved the highest accuracy. However, a major limitation of RFS is that when the number of trees becomes excessively large, the algorithm may become too slow to meet the demands of real-time prediction.

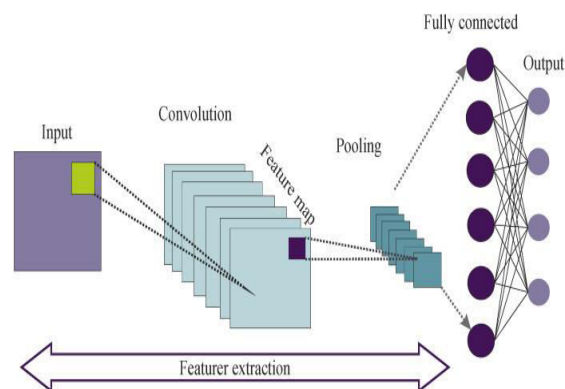


Fig 1: Earthquake Prediction Algorithms of Machine Learning

Sadhukhan [17] attempted to predict the magnitude of the next impending earthquake by analyzing eight mathematically calculated seismicity indicators. The study utilized three widely applied deep neural network models, namely LSTMs, bidirectional LSTMs (Bi-LSTMs), and Transformer models, to predict the magnitude of impending earthquakes in a given seismic region based on eight seismicity indicators calculated from past significant earthquake events greater than a predetermined threshold magnitude.

The core of artificial intelligence lies in machine learning, which serves as the key technology driving this transformation. The primary advantage of machine learning is its ability to identify functional relationships between vast amounts of data and their corresponding labels, which may be high-dimensional and nonlinear. This is the case in earthquake prediction, where the complexity is beyond human comprehension. Traditionally, earthquake prediction has



relied on the empirical judgment of experts, leading to randomness and uncertainty. However, the application of machine learning to earthquake prediction offers a promising solution for obtaining more accurate and reliable results. Nevertheless, owing to the scarcity of historical small earthquake data, many studies [18] have only utilized earthquake catalogs with magnitudes of 3 or above, resulting in limited training samples. This makes it difficult to train robust predictive models that can effectively generalize event patterns and prior knowledge. In this study, we compiled an earthquake catalog for the Yunnan region beginning from 1985, including earthquakes with magnitudes of 2 or above to capture more foreshock information. Additionally, as traditional sample feature extraction methods have failed to adequately consider regional active tectonic characteristics. Therefore, our feature extraction approach based on seismic prediction zoning can enhance the performance of machine learning prediction methods by incorporating implicit geographical-geological features.

II. LITERATURE STUDY

In [1], Z. Wang and B. Zhao describe a probability-based method utilizing initial P-wave parameters to estimate ground motion in real-time. The model incorporates early seismic signals into a probabilistic framework to enhance the accuracy of earthquake early warning (EEW). Deep learning aids in rapidly evaluating the P-wave's amplitude and frequency content to predict potential shaking intensity. Their experiments show significant improvements in reducing false alarms and increasing warning lead time compared to traditional models. This approach supports EEWS where immediate alerts are critical for evacuation and industrial safety. However, the method is still limited by noise in P-wave signals and the need for robust sensor coverage in high-risk zones.

In [2], S. A. Hamouda introduces advanced pre-earthquake prediction techniques combining geophysical indicators and machine learning algorithms. The study emphasizes the importance of detecting anomalies in radon emission, geomagnetic field, and ionospheric perturbations. A hybrid architecture using deep neural networks enhances the detection of nonlinear correlations among these variables. Results suggest improved forecasting accuracy days before seismic events. The paper also explores feature importance ranking for better model interpretability. While promising, the proposed system's effectiveness is constrained by sparse geophysical data and limited generalization across geographic regions.

In [3], Anbazhagu et al. explore the integration of AI and ML in earthquake prediction to improve precision and alert delivery. The study evaluates different deep learning models including LSTM, CNN, and DNN on temporal seismic datasets. LSTM-based architectures show superiority in capturing temporal dependencies of seismic wave propagation. The paper highlights an end-to-end system combining sensor data preprocessing, feature extraction, and real-time alerting. However, limitations include the need for large, annotated datasets and real-time computational efficiency, particularly in remote or data-constrained regions.

In [4], Esposito et al. present an IoT-based EEWS optimized for deployment on resource-constrained edge devices. Using lightweight deep learning models and real-time data from low-power accelerometers, the system ensures fast local detection and cloud-level alert propagation. This framework is scalable and suitable for developing nations with limited digital infrastructure. The paper demonstrates how deep learning models can be quantized and pruned without major losses in accuracy. However, hardware limitations and vulnerability to sensor drift remain key concerns for long-term deployment.

In [5], Chaudhary et al. introduce an LSTM-GAN hybrid model to generate synthetic seismic data and improve prediction robustness. The GAN component is trained to produce realistic seismic signal patterns, augmenting training datasets for LSTM-based forecasting. This approach mitigates the issue of limited historical data and enhances model generalization. Performance evaluation indicates that the model reduces prediction error and improves alert lead time. Nevertheless, computational complexity and GAN instability during training are notable drawbacks.

In [6], Murray et al. apply semi-supervised learning for identifying precursory seismic activity linked to landslides triggered by earthquakes. The model learns from a limited labeled dataset and utilizes unlabeled signals for clustering and classification. The method is particularly useful in data-scarce regions and for multi-hazard scenarios. Results validate its potential in early-stage hazard detection. Yet, model sensitivity to sensor calibration errors and noise remains an unresolved challenge.

In [7], Li et al. assess EEW system effectiveness in China by analyzing blind zone sizes using new seismic network data. Their framework combines geographical modeling and deep learning to determine the optimal sensor placement



and blind zone mitigation. This approach provides decision-makers with tools to maximize alert reach and minimize casualties. However, the method is heavily reliant on high-resolution regional data and may not scale easily to diverse topographies or underdeveloped networks.

In [8], Kuna et al. evaluate Mexico’s open-source EEWS through laboratory tests and sensor validation. They examine the fidelity of sensors and their compatibility with deep learning-based signal classifiers. Findings indicate that while open-source systems are economically feasible, their effectiveness depends heavily on calibration, maintenance, and data transmission reliability. The paper suggests integrating smart AI-based fault detection to improve system robustness.

In [9], Mittal et al. review the status of EEWS in India, discussing technological, infrastructural, and policy barriers. The paper explores how deep learning can overcome challenges related to noisy signals, low sensor density, and diverse terrain. Case studies from Uttarakhand and Himachal Pradesh demonstrate the feasibility of AI-enhanced EEWS. Despite advancements, the lack of standardization and realtime data integration still hinder widespread adoption.

In [10], Li et al. propose a multitask deep learning model that combines classification and regression for rockburst prediction in mines. The cascaded architecture improves both detection and severity estimation. By using shared representations, the system achieves higher accuracy and lower latency. While not strictly an EEWS, the methodology is transferable to seismic hazard monitoring. Key challenges include model complexity and the need for real-time computational efficiency in harsh environments.

III. METHODS

Within the past few years, there has been rapid growth in the seismic data quantity. This makes it challenging for modern seismology to analyze and process the data. Most of the popular techniques for earthquake prediction use the old seismic data, which was small. With the advancement in machine learning and deep learning, it is possible to extract useful information and train models on large datasets. Once we train a deep learning model with large amounts of data, it can acquire their knowledge by extracting features from raw data to recognize natural objects and make expert-level decisions in various disciplines. Besides the advancement in computational power, it has become straightforward to train large models. These advantages make deep learning suitable for applications in real-time seismology and earthquake prediction.

For the task of the earthquake prediction, the deep learning models which perform better than other models are CNN and LSTM: Convolution Neural Network “In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation invariance characteristics.

They have applications in image and video recognition, recommender systems, image classification, medical image analysis, natural language processing, brain-computer interfaces, and financial time series”.

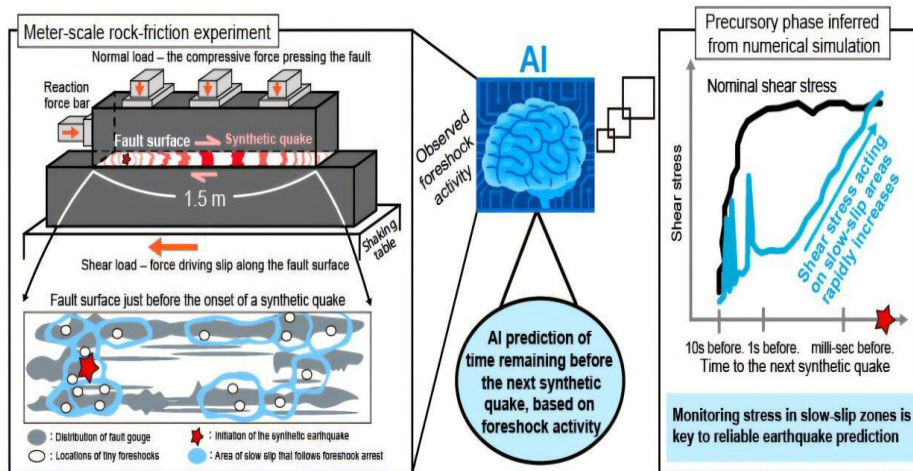


Fig 2: predict earthquakes: Machine learning detects subtle changes

A convolutional neural network consists of an input layer, hidden layers and an output layer. A typical CNN consists of:

- Convolution Layer: Convolutional layers convolve the input and pass its result to the next layer.
- Pooling Layer: Pooling layers reduce the data’s dimensions by combining the outputs of neuron clusters at one layer into a single neuron in the next layer.
- Fully Connected Layer: Fully connected layers connect every neuron in one layer to every neuron in another layer.

VisualSim can be used to construct deep learning models where “Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. LSTM networks are well-suited to classifying, processing and making predictions based on time series data since there can be lags of unknown duration between important events in a time series.” (Source: Wikipedia) VisualSim supports timed seismograph data.

IV. RESULT ANALYSIS

This results in sample data that fail to comprehensively reflect the regional seismic activity patterns. During model training, there is a lack of sufficient effective information, making it difficult to capture the complex mechanisms of earthquake occurrence. Consequently, the prediction effects are extremely poor (e.g., both the LSTM and Stacking models exhibit high MSE values). This demonstrates that data quality (integrity and representativeness) is fundamental to model performance. Even with complex model structures, achieving ideal results is challenging if the samples exhibit systematic biases (e.g., external interference and data sparsity).

Regions account for 40% of the study area in Yunnan and are situated within the frequently active seismic zone in the interior. These regions contain 60% of the historical minor earthquake catalogs and are less influenced by external plates. Their data more accurately reflect the regional seismic activity characteristics, providing a more stable and reliable training basis for the models. Therefore, even with the same model structure, the prediction effects in these regions are significantly superior to those in boundary regions (e.g., the MSE of the LSTM model for region No. 5 is 0.15, significantly lower than that in other regions). This validates the critical impact of local representativeness and data volume of samples on the generalizability of the model.

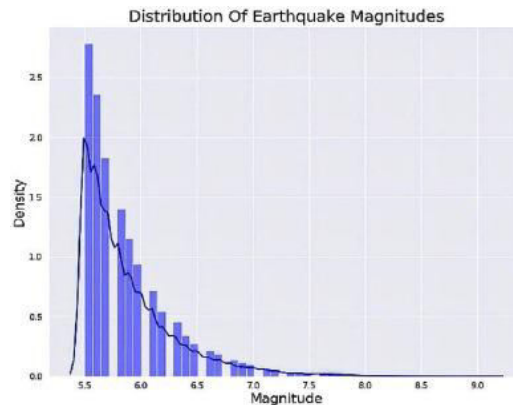


Fig 3: Earthquake prediction optimization using deep learning LSTM model

The above analysis shows that the matching between regional geological characteristics and model applicability is crucial. Seismic activity at plate boundaries is controlled by the interactions of multiple plates, resulting in complex and highly nonlinear mechanisms. Existing models (LSTM and Stacking) may not be able to fully capture such cross-plate dynamic processes, leading to large prediction errors. In contrast, seismic activity in the interior regions of Yunnan may be more controlled by local tectonics, with more consistent data characteristics. Models can learn the underlying patterns of earthquake occurrence more accurately in such regions, indicating that models are more applicable in regions with simple geological conditions and homogeneous data characteristics. The complexity of seismic activity is closely related to the differences in regional geological structures. Seismic mechanisms, data characteristics, and prediction challenges vary significantly across different tectonic regions. Therefore, conducting research based on seismic prediction zoning is not only key to improving earthquake prediction accuracy but also an inevitable choice for scientifically understanding seismic activity patterns.

V. CONCLUSION

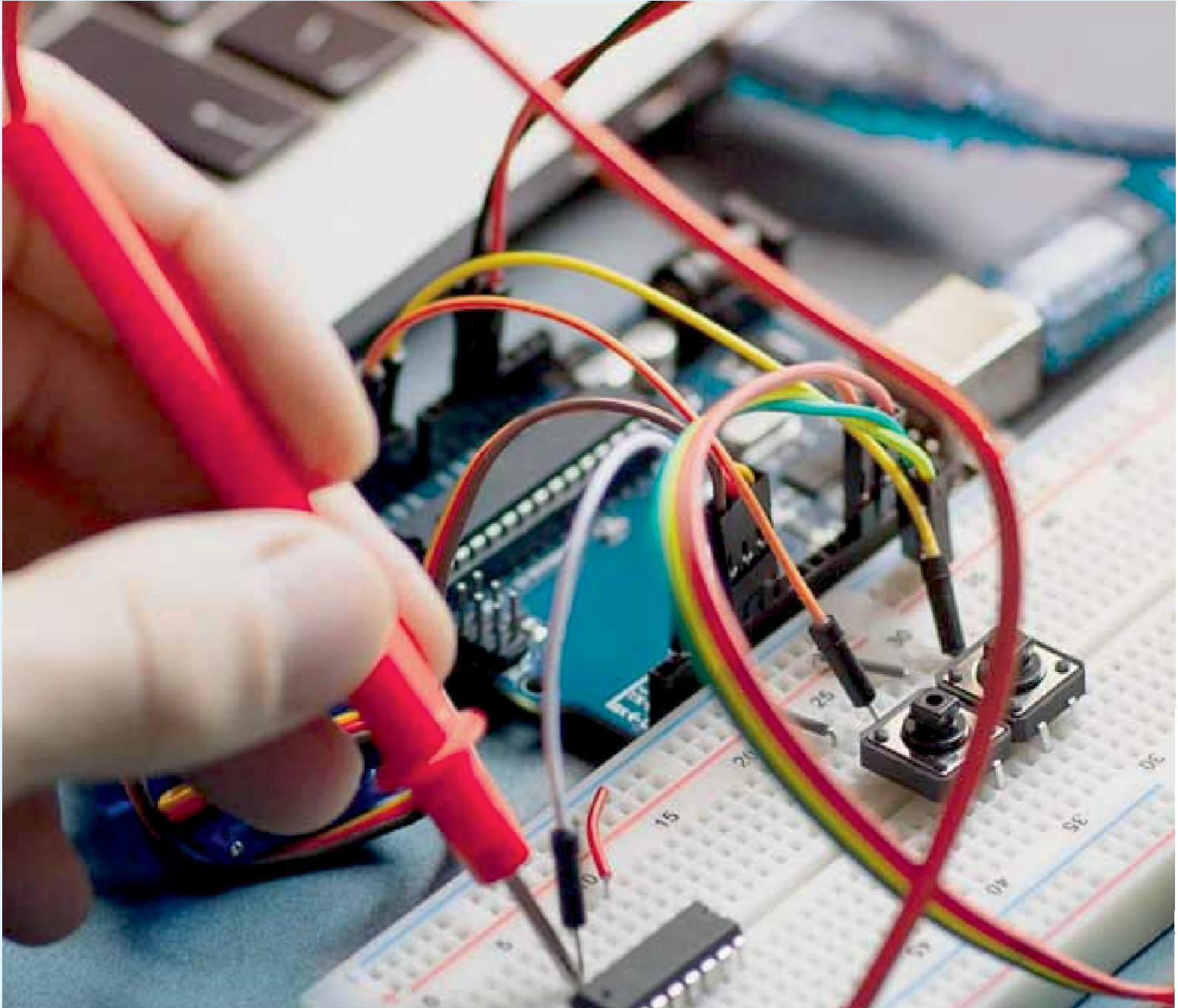
This study explored the enhance prediction performance using machine learning methods and proposed a seismotectonic zoning-based approach for earthquake prediction. Compared to the traditional feature extraction method that solely relies on latitude-longitude grids, the proposed approach, which is based on seismic prediction zoning, more effectively captures the spatiotemporal dependencies associated with earthquake occurrence by incorporating implicit geographical and geological features. The study finds that earthquake prediction based on seismic prediction zoning significantly outperforms the traditional latitude-longitude grid-based approach. On average, the prediction errors (MSE) of all tested models decreased by 24.6%, whereas the explained variance (R²) increased by 12.3%. Notably, tree-based models (e.g., RF) demonstrated a remarkable improvement in prediction accuracy, with an approximately 33.8% increase. The LSTM approach exhibits satisfactory effectiveness within particular seismotectonic zones. Therefore, integrating geological structural knowledge into the machine learning framework and constructing sample sets based on seismic prediction zoning represent effective strategies for substantially improving the accuracy and reliability of earthquake prediction models. For regions with significant geological activity, this method should be prioritized.

REFERENCES

1. Meng, L.Y.; Hu, F.; Zang, Y.; Zhang, Q.; Li, M. Introduction to the Artificial Intelligence Earthquake Prediction Models in Sichuan–Yunnan Region. *J. Seismol. Res.* 2026, 49, 43–50. [Google Scholar]
2. Gutenberg, B.; Richter, C.F. Earthquake Magnitude, Intensity, Energy, and Acceleration. *Bull. Seismol. Soc. Am.* 1942, 32, 163–191. [Google Scholar] [CrossRef]
3. Aslam, B.; Zafar, A.; Qureshi, U.A.; Ahmad, R.; Ali, M.; Khan, S. Seismic Investigation of the Northern Part of Pakistan Using Statistical and Neural Network Algorithms. *Environ. Earth Sci.* 2021, 80, 59. [Google Scholar] [CrossRef]
4. Kaftan, I.; Şalk, M.; Şenol, Y. Processing of Earthquake Catalog Data of Western Turkey with Artificial Neural Networks and Adaptive Neuro-Fuzzy Inference System. *Arab. J. Geosci.* 2017, 10, 243. [Google Scholar] [CrossRef]



5. Wang, Q.; Guo, Y.; Yu, L.; Zhang, X.; Chen, J. Earthquake Prediction Based on Spatio-Temporal Data Mining: An LSTM Network Approach. *IEEE Trans. Emerg. Top. Comput.* 2017, 8, 148–158. [Google Scholar] [CrossRef]
6. Shi, X.Y. Research on Earthquake Prediction Based on Machine Learning Regression Algorithms and Its Application in the China Earthquake Scientific Experimentation Site. Ph.D. Thesis, Institute of Earthquake Prediction, China Earthquake Administration, Beijing, China, 2021. [Google Scholar]
7. Panakkat, A.; Adeli, H. Neural Network Models for Earthquake Magnitude Prediction Using Multiple Seismicity Indicators. *Int. J. Neural Syst.* 2007, 17, 13–33. [Google Scholar] [CrossRef]
8. Adeli, H.; Panakkat, A. A Probabilistic Neural Network for Earthquake Magnitude Prediction. *Neural Netw.* 2009, 22, 1018–1024. [Google Scholar] [CrossRef]
9. Panakkat, A.; Adeli, H. Recurrent Neural Network for Approximate Earthquake Time and Location Prediction Using Multiple Seismicity Indicators. *Comput.-Aided Civ. Infrastruct. Eng.* 2009, 24, 280–292. [Google Scholar] [CrossRef]
10. Asencio-Cortés, G.; Scitovski, S.; Scitovski, R.; Štambuk, M. Temporal Analysis of Croatian Seismogenic Zones to Improve Earthquake Magnitude Prediction. *Earth Sci. Inform.* 2017, 10, 303–320. [Google Scholar] [CrossRef]
11. Asim, K.M.; Awais, M.; Martínez-Álvarez, F.; Hussain, A.; Ali, S. Seismic Activity Prediction Using Computational Intelligence Techniques in Northern Pakistan. *Acta Geophys.* 2017, 65, 919–930. [Google Scholar] [CrossRef]
12. Asim, K.M.; Idris, A.; Iqbal, T.; Martínez-Álvarez, F.; Hussain, A. Seismic Indicators-Based Earthquake Predictor System Using Genetic Programming and AdaBoost Classification. *Soil Dyn. Earthq. Eng.* 2018, 111, 1–7. [Google Scholar] [CrossRef]
13. Zhou, W.; Dong, H.; Liang, Y. The Deterministic Dendritic Cell Algorithm with Haskell in Earthquake Magnitude Prediction. *Earth Sci. Inform.* 2020, 13, 447–457. [Google Scholar] [CrossRef]
14. Li, L.F.; Shi, Y.L.; Cheng, S. Exploration of Long Short-Term Memory Neural Network in Intermediate Earthquake Forecast: A Case Study in Sichuan–Yunnan Region. *Chin. J. Geophys.* 2022, 65, 12–25. [Google Scholar]
15. Asencio-Cortés, G.; Morales-Esteban, A.; Shang, X.; Martínez-Álvarez, F. Earthquake Prediction in California Using Regression Algorithms and Cloud-Based Big Data Infrastructure. *Comput. Geosci.* 2018, 115, 198–210. [Google Scholar] [CrossRef]
16. Mallouhy, R.; Abou Jaoude, C.; Guyeux, C.; Makhoul, A. Major Earthquake Event Prediction Using Various Machine Learning Algorithms. In *Proceedings of the International Conference on Information and Communication Technologies for Disaster Management*, Paris, France, 18–20 December 2019; pp. 1–6. [Google Scholar]
17. Sadhukhan, B.; Chakraborty, S.; Mukherjee, S. Predicting the Magnitude of an Impending Earthquake Using Deep Learning Techniques. *Earth Sci. Inform.* 2025, 16, 803–823. [Google Scholar] [CrossRef]
18. Florido, E.; Asencio-Cortés, G.; Aznarte, J.L.; Rubio-Escudero, C.; Martínez-Álvarez, F. A Novel Tree-Based Algorithm to Discover Seismic Patterns in Earthquake Catalogs. *Comput. Geosci.* 2018, 115, 96–104. [Google Scholar] [CrossRef]
19. Wang, X.; Zhong, Z.Y.; Yao, Y.C.; Li, Z.; Zhou, S.H.; Jiang, C.C.; Jia, K. Small Earthquakes Can Help Predict Large Earthquakes: A Machine Learning Perspective. *Appl. Sci.* 2023, 13, 6424. [Google Scholar] [CrossRef]
20. Jiang, H.K.; Li, Y.L.; Qu, Y.J.; Zhao, X.W.; Chen, P.; Zhang, M. Spatial Distribution Characteristics of Sequence Types for Moderate–Strong Earthquakes in the Chinese Mainland. *Acta Seismol. Sin.* 2006, 28, 389–398. [Google Scholar]
21. Wiemer, S.; Wyss, M. Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bull. Seismol. Soc. Am.* 2000, 90, 859–869. [Google Scholar] [CrossRef]



INNO  SPACE
SJIF Scientific Journal Impact Factor



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering

 9940 572 462  6381 907 438  ijareeie@gmail.com



www.ijareeie.com

Scan to save the contact details