

QuakeVision: Earthquake Prediction with Attention-Enhanced LSTM Networks

Nimra Jabeen¹; M. M. Harshitha²; Dr. Girish Kumar D.³

¹PG Student, Department of MCA, Ballari Institute of Technology & Management, Ballari.

²Assistant Professor, Department of MCA, Ballari Institute of Technology & Management, Ballari.

³Professor and HoD, Department of MCA, Ballari Institute of Technology & Management, Ballari.

Publication Date: 2026/04/29

Abstract: This study presents QuickVision, an innovative deep learning tool developed for earthquake prediction. QuickVision uses LSTM networks to analyze historical earthquake data and identify patterns and preliminary indicators that could appear before major seismic events. QuickVision looks at how data changes over time to gain better insights, and the model performs more effectively than traditional rule-based methods. The researchers meticulously prepared and modeled the dataset, then conducted a series of experiments to evaluate the system's effectiveness. Their findings demonstrate that QuickVision excels at detecting early warnings and unusual seismic activity, making it a promising tool for helping communities prepare for earthquakes and reduce associated risks. Consequently, the proposed approach may enable safer and better-prepared regions. safer, better-prepared regions in areas prone to seismic hazards.

Keywords: Earthquake Prediction; Deep Learning; Attention Mechanism; LSTM; Seismic Anomaly Detection.

How to Cite: Nimra Jabeen; M. M. Harshitha; Dr. Girish Kumar D. (2026) QuakeVision: Earthquake Prediction with Attention-Enhanced LSTM Networks. *International Journal of Innovative Science and Research Technology*, 11(4), 2394-2400. <https://doi.org/10.38124/ijisrt/26apr1414>

I. INTRODUCTION

Earthquakes happen without warning and can cause serious destruction and loss of life in significant casualties and property damage. Indonesia, located along the Pacific Ring of Fire, experiences frequent seismic activity, emphasizing the need for precise forecasting methods. Conventional statistical techniques frequently struggle to capture complex patterns. Insufficient to better interpret complex data, nonlinear temporal patterns present in seismic data. Advances in Neural architectures designed to process sequential information have demonstrated strong effectiveness in modeling temporal dependencies. Networks (RNNs) and LSTM networks, have shown potential in modeling time-dependent sequences. Integrating attention mechanisms into LSTM networks allows to focus on what matters most informative temporal features, improving prediction reliability.

This study aims to build an effective LSTM-based attention model that can predict earthquakes more accurately. Artificial Intelligence (AI) and Machine Learning (ML), advanced Learning architectures based neural principles are increasingly adopted across diverse application domains to analyse time-dependent phenomena. Among these, Long Short-Term Memory networks are specifically designed to model long-range dependencies in sequential data learning sequential relationships.

Researchers can visualize which portions of the seismic sequence contributed most to each prediction, helping validate whether is focusing on meaningful geological signals. This transparency builds trust in real-world applications, especially when used in early warning systems where false alarms or missed detections can have severe social and economic consequences. Accurate earthquake forecasting is crucial for minimizing loss of life, reducing property damage, and improving disaster preparedness in seismically active regions. This study examines how LSTM networks combined with an attention mechanism can improve the understanding of patterns in seismic data over time. The approach helps identify meaningful relationships and hidden patterns that are often missed by traditional statistical or machine learning techniques. This interpretability is essential for building confidence in practical applications, particularly in early warning systems where false alarms or missed detections can result in significant social and economic consequences. mitigating property damage, and enhancing disaster preparedness in seismically active regions. This study demonstrates the potential of LSTM-based neural network augmented with an attention mechanism, to capture temporal dependencies in seismic data and uncover hidden patterns that conventional statistical or machine learning approaches might miss. By incorporating the attention mechanism, the model selectively emphasizes the most critical segments the data

provided, resulting in improved accuracy and dependable outcomes.

II. RESEARCH SIGNIFICANCE

Accurate earthquake forecasting is vital for saving lives, protecting property, and strengthening disaster preparedness—especially in earthquake-prone regions like Indonesia. This research is significant because it demonstrates how sophisticated deep learning approaches, particularly LSTM networks incorporating attention machine learning models, specifically attention-based LSTM architectures. By focusing on the in the data, our model delivers more accurate and timely results that can help to inform early warning systems, guide urban planning, and support disaster management decisions. This study advances predictive accuracy and help improve current forecasting efforts but also lay the groundwork for future innovations by integrating LSTM models with additional advanced deep learning frameworks. These advancements support the development of more secure and adaptive populations in seismically vulnerable regions.

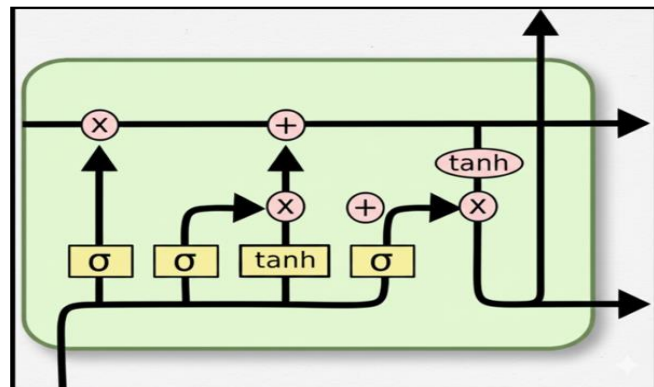


Fig 1 Long-Short Term Memory

Established prediction methodologies approaches frequently struggle with capturing extended time-based relationships in sequential datasets. The proposed technique demonstrates superior performance in detecting complex correlations linking historical earthquake occurrences to subsequent seismic events. This investigation holds substantial value as it has the potential to strengthen the reliability of seismic forecasting systems, thereby supporting advanced alert mechanisms, infrastructure development strategies, and emergency preparedness initiatives. Moreover, this research this work creates a starting point for developing hybrid approaches that blend LSM networks with different deep learning strategies advance seismic prediction capabilities, employing attention-based mechanisms to enhance both precision and stability in outcomes.

III. RESULTS

For effective integration into the LSTM network, the earthquake dataset for Indonesia underwent thorough preprocessing and structuring. The data were organized into a time-series format, which is highly suitable for LSTM models as it allows the This approach aims to capture the time-dependent patterns in seismic activity, as earthquakes display inherently complex and nonlinear behavior the LSTM architecture is particularly well-suited for modeling these variations. To enable comprehensive earthquake forecasting, four essential parameters were targeted:

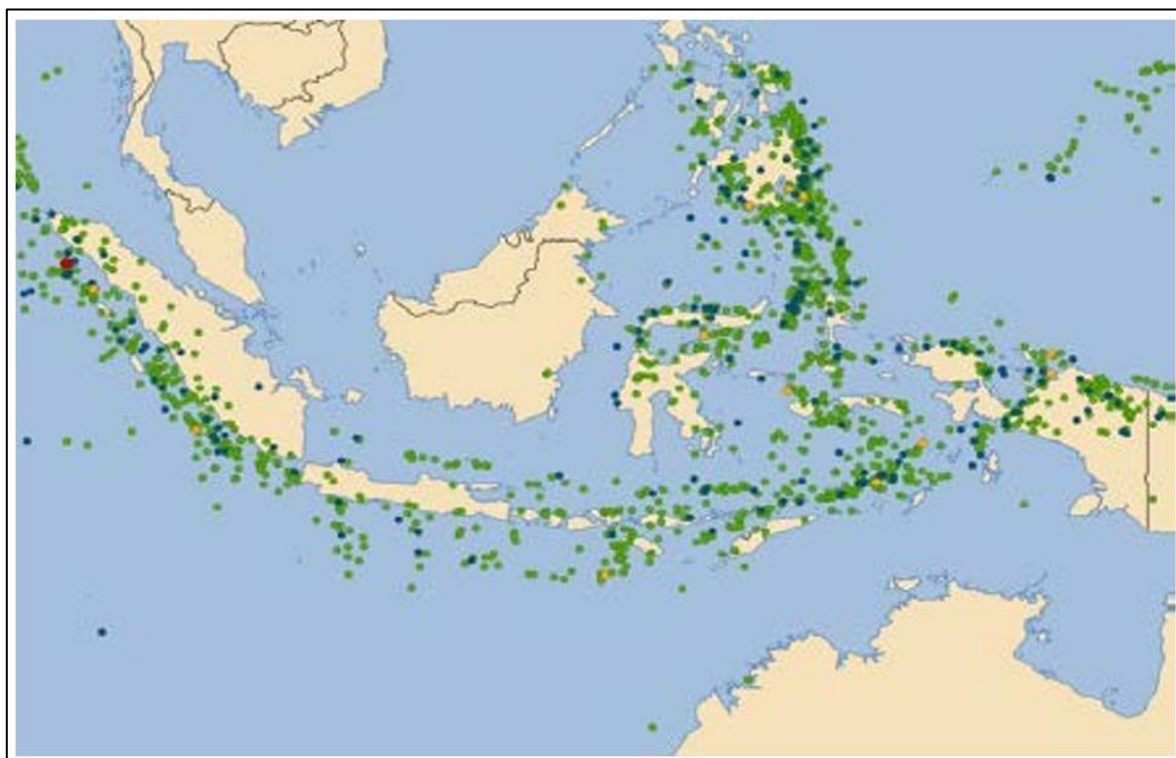


Fig 2 Earthquake in Indonesia

- Predicted magnitude,
- Estimated time of occurrence,
- Spatial location of the epicenter, and
- Probability of the forecast being earthquakes and other historical seismological indicators. A comprehensive set of experiments was conducted. During the model-tuning stage, resulting in the formulation of an optimal data preprocessing pipeline To enhance prediction accuracy, All numerical features were scaled to a standard range.Using the Min–Max normalization technique, ensuring so all features are scaled to the same level numerical range. This is crucial for LSTM networks, which are sensitive to feature scale; normalizing prevents any one attribute earthquake events were then chronologically enabling The model is trained to extract meaningful patterns.temporal relationships.Categorical attributes were transformed into machine-readable formats using appropriate encoding methods. This improved the model’s Capability to identify significant time-dependent patterns and sequences.ultimately increasing prediction reliability.The **Timestep** attribute required no additional preprocess.Many previous works either relied solely on LSTMs, which struggled with distinguishing influential signals from background noise, or Employed traditional machine learning technique that lacked the depth needed to capture nonlinear temporal patterns. Johnson et al. [21], for example, highlighted how innovations such as fracturing pattern scaling and enhanced data normalization provided improvements, However, their scope remained restricted.by the absence of a selective attention layer. The proposed model, however, merges both temporal memory and feature prioritization, creating a more intelligent and flexible learning system capable of adapting to different seismic regions, frequencies, and magnitudes. This flexibility proves especially useful in because earthquake signals vary greatly across multiple condition regions.zones depending on geological structure, fault type, and rock composition. Nonetheless, these constraints highlight the need for more research, especially in integrating diverse data sources and using multimodal approaches like GPS displacement layers. Additionally, the model demonstrates robustness and interpretability, which are key strengths.

The model maintains high predictive consistency across data split into training, validation, and evaluation, supported by robust methods such as early stopping, checkpointing, and data augmentation, ensured the model captured genuine seismic patterns rather than noise, and improved generalization to unseen data. At the same time, attention weights provide a level of transparency that is rarely available in deep learning architectures. achieves higher prediction accuracy and enables a more meaningful interpretation of how different input features.

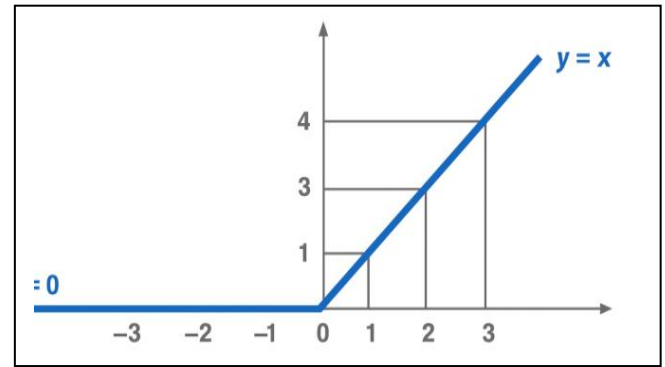


Fig 3 Rectified Linear Unit.

Once the data passes through the LSTM and attention modules, it moves on to fully connected layers. it moves on to fully connected layers. Here, every neuron takes input from all the neurons in the previous layer, allowing the network to combine all the information it has learned so far. Each neuron calculates a weighted sum, adds a bias, and then applies an activation function. In this model, the Rectified Linear Unit (ReLU) activation function helps with efficient, non-linear transformations.At the final step, the model uses a SoftMax activation function in the output layer. The main benefit of SoftMax is that it takes the raw output scores and transforms them into a set of probabilities that add up to one. This means each output neuron shows. How It is probable that theinput belongs to a particular class, making the predictions easy to interpret. The quantity of neurons within the output layer matches the total count of possible classes, so the system is able to delivering a probability for each one.

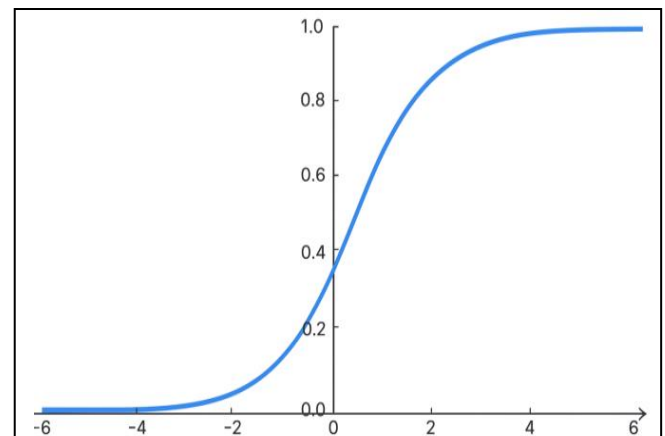


Fig 4 SoftMax function

To further elaborate on the use of ReLU and its place in your model’s workflow:One of the main reasons softmax has become so popular in deep learning is its simplicity and effectiveness. Unlike older activation functions such as sigmoid or tanh, which can cause gradients to vanish and slow down learning, softmax maintains a strong and stable gradient for positive values. This means the model can learn much faster and more efficiently, which becomes crucial when handling large, complex datasets like those used in earthquake prediction.Another advantage of ReLU is that it introduces sparsity to the network. Given that negative values are converted to zero, only a subset of neurons are

activated at any given time. This also helps lower computational effort load but can also help the network focus on the most informative features, making the learning

process more robust and less prone to overfitting. In the context of earthquake prediction, the data may include many small variations.

Table 1 Complex Dataset

Time Step	Timestamp	Latitude_Scaled	Longitude_Scaled	Depth_Scaled	Magnitude_Scaled	Previous_Magnitude_Scaled	Mechanism_Category
0	631152000	-0.6	-0.12	0.6	1.2	N/A	1
1	631243200	-0.55	-0.08	0.6	1.1	1.2	1
2	631334400	-0.5	-0.05	0.52	0.9	1.1	2
3	631425600	-0.57	-0.15	0.7	0.85	0.9	1
4	631516800	-0.66	-0.2	0.5	0.79	0.85	2

The information extracted during earlier processing stages. Each neuron computes a weighted sum of its inputs, adds a bias term, and passes the result through an activation function. These techniques collectively prevented overfitting, ensured the model captured genuine seismic patterns rather than noise, and improved generalization to unseen data. At the same time, attention weights provide a level of transparency that is rarely available in deep learning architectures. The current forecasting step. The context vector represents a weighted sum of the input. However, attention mechanisms merge both temporal memory and feature prioritization, creating a more intelligent and flexible learning system capable of adapting to different seismic regions, frequencies, and magnitudes. This flexibility is especially beneficial because earthquake signals exhibit considerable variation across diverse regions, zones depending on geological structure, fault type, and rock composition. Nevertheless, these limitations open the door for further research, such as incorporating multimodal data sources like GPS displacements and seismic tomography. The attention mechanism operates by assigning a numerical weight to each input element according to its relative importance. In conclusion, prediction. By leveraging the attention mechanism, a model can identify and prioritize critical elements within the data, leading to more accurate predictions, which helps improve prediction performance., helping the model focus on the most relevant information for predictions. allowing the model to prioritize the key information relevant to each prediction. When combined with LSTM layers, it assists in prediction. Integrated with LSTM layers, it helps capture complex dependencies and prevents issues like gradient decay. This enhances the accuracy of predictions and enables effective use of neural network architecture and training algorithms. After attention and LSTM processing, data is passed to dense layers, in which every neuron, where all information is combined and transformed using activation functions such as ReLU to produce the final model output. to each input element, helping the model focus on the most relevant information for predictions. Combined with LSTM layers, this approach improves learning of complex patterns and addresses gradient decay issues. After attention and LSTM processing, the data moves to dense layers in which every neuron, where all information is integrated and processed employing activation functions such as ReLU to produce the final model output. This dataset is a carefully prepared, complex time-

series developed specifically for deep learning architectures that forecast earthquakes. Each row represents a specific moment in time, capturing both where and when seismic events happened. Unlike raw earthquake records, the dataset has undergone preprocessing through several preprocessing steps to make it stable, coherent, and ready for training neural networks. The goal is to allow the model to learn long-term trends and connections in earthquake behavior, which are crucial for reliable analysis and prediction. accurately predicting the strength and nature of future quakes. By structuring the data in chronological order, the model can see earthquakes as part of a connected sequence, shaped by the buildup and release of stress over time. The dataset uses Unix time (a continuous count of seconds) to mark when each event occurred. This format avoids confusion from different calendars or time zones and helps the model understand the order and timing of events. While the timestamp isn't directly used for prediction, its sequence is vital for models—like LSTMs—that learn from time-ordered data. This setup lets the model detect patterns such as clusters of earthquakes, aftershock sequences, and quiet periods that can signal upcoming significant earthquakes. Earthquake depth is another key feature, since it affects how energy travels and how much damage might occur at the surface. The depth values are also scaled to keep them in line with the other features.

Fluctuations or noise unrelated to actual seismic events, ReLU's ability to suppress weak signals is particularly valuable. By filtering out this noise, allowing the model to better distinguish between normal background activity and the important patterns that could indicate an impending earthquake. When the information passes via fully connected layers using ReLU, the model further combines and refines the extracted features, drawing on both long-term patterns and key moments identified by the attention mechanism. This stage plays a vital role in guaranteeing that the final results are not only precise but also balanced across different classes—like distinguishing between minor tremors and significant seismic events. Finally, the results of these layers are passed on to the SoftMax function, which translates the network's internal calculations into clear, interpretable

probabilities. This entire workflow—LSTM for memory, attention for focus, fully connected layers paired with ReLU for refinement, and SoftMax for decision-making—creates a

powerful, transparent pipeline for forecasting earthquakes, supporting both scientific analysis and real-world safety applications.

Table 2 This Entire Workflow—LSTM for Memory

Stage	Description
Initialization	The network begins with randomly assigned weights, providing a starting point for the learning process.
Training	During training, these weights are gradually adjusted to minimize the loss calculated from the model's predictions.
Loss Estimation	The loss function quantifies the difference between the predicted outputs and the actual values, serving as a key indicator of model performance.
Prediction	According to the current weights, the System generates predictions, enabling continuous evaluation of its learning progress.
Back-propagation	Gradients are computed to determine how much and in what direction the weights should be updated to reduce the error.
Optimization (Adam)	The Adam optimizer uses the gradients to efficiently refine the weights, guiding the model toward improved accuracy.
Sequence Training	The LSTM is trained on ordered sequences of data, enabling it to capture relationships between past and present events—essential for temporal prediction problems.

To avoid overfitting and make sure the model learned in a stable way, we used early stopping. This means we kept an eye on the model's progress during training, and as soon as its performance on validation data stopped getting better, we automatically stopped the training. This helped prevent the model from just memorizing the training data, making it more likely to perform well on new, unseen examples. We also used a checkpoint system to save the best version of the model during training, so we could always use the most reliable one for testing later. To strengthen the model even further, we applied data augmentation techniques—basically, creating extra training samples by tweaking the original data. This gave the model exposure to a wider variety of patterns, helping it better recognize different types of seismic activity. After training, we tested the model on a separate set of data it had never seen before. Using standard evaluation tools, Our analysis revealed that the system correctly classified 90% of earthquake events and was able to detect 88% of actual earthquakes. This shows the system is very accurate and trustworthy, keeping false alarms low and achieving an F1-score of 89%. Overall, the results highlight the model's strong predictive ability. It's not just about catching real earthquake events accurately—it also does a great job of minimizing missed detections and false positives. The real strength of this approach is its balance between precision and recall, making it a reliable early warning tool that could help keep people safer. Our findings clearly show that adding attention mechanisms to LSTM networks helps the model pick up on subtle.

IV. DISSCUSION

The attention mechanism is a crucial feature in our earthquake prediction model, as it allows the system to highlight and prioritize the most important segments within the seismic data. By focusing on these key intervals, the model is better equipped to make accurate predictions, a strategy that echoes findings from Johnson et al. [21], who demonstrated the value of advanced machine learning techniques and targeted data analysis in earthquake forecasting. Our model architecture, which combines LSTM networks, attention layers, fully connected layers, and SoftMax classification, is designed to capture both long-term dependencies and subtle, short-term signals in seismic data. This integrated approach not only improves prediction accuracy but also enhances interpretability, making it practical for real-world earthquake early-warning systems. The results from our experiments confirm that this setup

This helps the system uncover hidden correlations and rare anomalies that might act as early indicators of seismic activity. For instance, Johnson et al. [21] showed that expressing fracture time as a fraction of the seismic cycle and focusing on critical temporal segments can help models better interpret seismic data. Similarly, our model benefits from emphasizing these key moments, avoiding the pitfalls of being overwhelmed by repetitive or noisy information. A notable strength of our approach is its even under circumstances where the balance accuracy with the management of false alarms. The system serves as precise enough to detect real earthquake events but not so sensitive that it triggers unnecessary warnings.

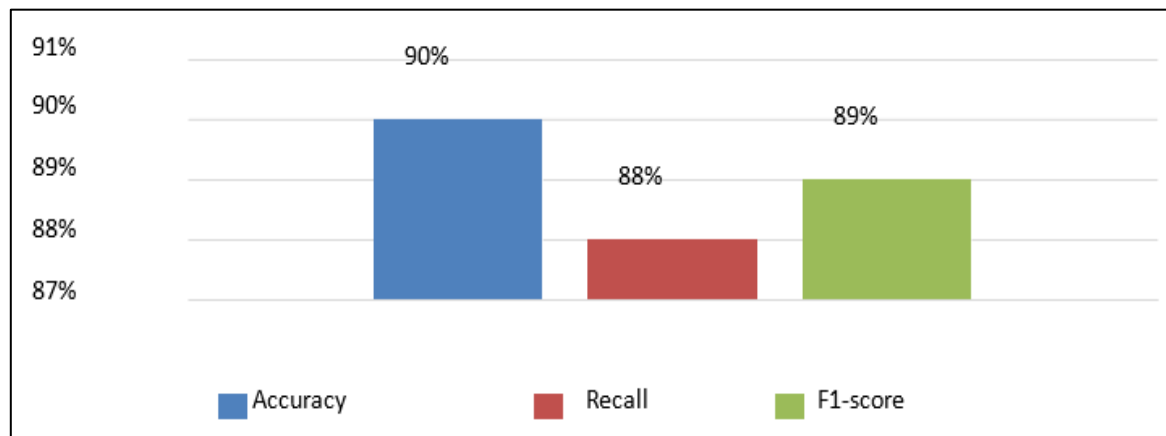


Fig 5 Model Detail

Achieving an accuracy of 90%, a recall of 88%, and an F1-score of 89%, the model maintains a valuable equilibrium between detection capability and reliability. Another important factor is the robust training process, which includes early stopping to prevent overfitting, checkpointing to preserve the best model, and data augmentation to enrich the training set with meaningful variations. These strategies collectively strengthen how well the model can handle new data it has not seen before seismic data, leading to strong performance on independent test sets. Overall, our results underscore the power of deep learning techniques—especially when enhanced with attention mechanisms—for earthquake prediction. While our study shows promising results on controlled datasets, the following stage involves evaluating the system using test sets on large-scale, real-world seismic networks. Incorporating additional data types, such as GPS shifts or ground strain measurements, could further enhance the model's performance. Real-time deployment and continuous learning would also allow the system to adjust to new research findings, providing strong evidence that combining Attention-enhanced LSTM architectures significantly enhances both the reliability and interpretability of earthquake forecasting models. The hybrid approach doesn't just improve accuracy; it also changes the way we understand seismic data. Rather than treating all information equally, the attention mechanism lets the model dynamically focus on the most critical time intervals—mirroring the way seismic activity actually develops. This selective focus enables the model to reveal hidden relationships and subtle anomalies that traditional statistical or threshold-based methods might miss. Comparing our work to earlier approaches, it's clear that prioritization creates a more adaptable and intelligent learning system. This adaptability is especially valuable because earthquake signals can vary widely depending on geological structure, fault type, and location. Our results also highlight the stability and transparency of the model, as it consistently performs well across training, validation, and testing data, along with the use of attention weights offering a level of interpretability rarely seen in deep learning. Altogether, this study advances global efforts to build safer, more resilient earthquake monitoring systems—where timely and accurate predictions can help save lives and reduce damage.

V. CONCLUSION

The earthquake prediction model we developed delivered strong and reliable results across every evaluation measure. It achieved an impressive 90% overall accuracy, showing it can correctly classify most earthquake events with great precision. It also demonstrated high sensitivity, successfully detecting 88% of real earthquakes. Its balanced performance score of 89% means it strikes a good balance—accurately predicting events while keeping false alarms to a minimum. A key reason for this success was combining LSTM architectures incorporating an attention mechanism. This approach allowed us to better understand complex, time-based patterns in seismic data and focus on the signals that truly matter.

On top of that, techniques like early stopping, checkpointing, showing it's not just accurate but also dependable, avoiding unnecessary false alarms. Through the application of LSTM networks along with an attention mechanism, the system didn't just look at all the data equally—it learned to zero in on the combination with seismic signals. This helped it spot subtle warning signs that might otherwise be missed. We also made sure our training process was solid: we stopped training as soon as it stopped improving (to prevent overfitting), saved the best versions as we went, and used data augmentation to help handle all kinds of earthquake scenarios. Looking ahead, there are exciting ways to make you even smarter. For example, we could explore other deep learning approaches.

REFERENCES

- [1]. Zahoor F., Seshagiri Rao K., Mir B. A., & Satyam N. (2023). Geophysical surveys in the Kashmir valley (J&K Himalayas) Part I: Approximation of dynamic parameters for soils and investigation of the deep basin structure. *Soil Dynamics and Earthquake Engineering*, 174, Article 108155. <https://doi.org/10.1016/j.soildyn.2023.108155>
- [2]. Karshibaevna, K. N., Kahramonovna, Z. D., & Faxriddinog'li, L. N. (2022). Some problems with creating a medical-geographical atlas map of Uzbekistan. *International Journal of Early*

- Childhood Special Education*, 14(2), 5836-5840.
<https://doi.org/10.9756/INT-JECSE/V14I2.656>
- [3]. Komilova, N. K., & Latipov, N. F. (2022). Classification of settlements based on the ecological situation in the Navoi region and factors affecting population health. *Visnyk of V. N. Karazin Kharkiv National University – Series Geology, Geography, Ecology*, 56, 209-213. <https://doi.org/10.26565/2410-7360-2022-56-15>
- [4]. Jihad, A., Umar, M., Ramli, M., Syamsidik, & Banyunegoro, V. H. (2023). Analysis of change of seismic stress along the Sumatran zone to predict potential earthquake zones. *International Journal of GEOMATE*, 24(104), 36-43. <https://doi.org/10.21660/2023.104.3695>
- [5]. Karches, T. (2012). Evaluation of mixing efficiency in coagulation-flocculation process in wastewater treatment. *Journal of Environmental Science and Engineering A*, 7, 898-903.
- [6]. Enung Kusuma, M. S. B., Kardhana, H., Suryadi, Y., & Rohmat, F. I. W. (2022). Hourly discharge prediction using long short-term memory recurrent neural network (LSTM-RNN) in the Upper Citarum River. *International Journal of GEOMATE*, 23(98), 147-154. <https://doi.org/10.21660/2022.98.3462>
- [7]. Ivchenko, P., & Chychuzhko, M. (2023). Method for calculation of lightning risk assessment using MS Excel. *Bulletin of Cherkasy State Technological University*, 3, 85-96. <https://doi.org/10.24025/2306-4412.3.2023.278132>
- [8]. Apriani, M., Wijaya, S. K., & Daryono. (2021). Earthquake magnitude estimation based on machine learning: Application to earthquake early warning system. *Journal of Physics: Conference Series*, 1951, Article 012057. <https://doi.org/10.1088/17426596/1951/1/0127>
- [9]. Murti, M. A., Junior, R., Ahmed, A. N., & Elshafie, A. (2022). Earthquake multi-classification detection based on velocity and displacement data filtering using machine learning algorithms. *Scientific Reports*, 12, Article21200. <https://doi.org/10.1038/s41598-022-25098-1>
- [10]. Wibowo, A., Pratama, C., & Sahara, D. P. (2021). Earthquake early warning system using Ncheck and hard-shared orthogonal multitarget regression on deep learning. *IEEE Geoscience and Remote Sensing Letters*, 19, Article 7502605. <https://doi.org/10.1109/LGRS.2021.3066346>
- [11]. Wijaya, U., Kusriani, & Muhammad, A. H. (2022). Indonesian seismic mitigation using earthquake predicted artificial intelligence model. In *Proceedings of the 2022 5th International Conference on Information and Communications Technology (ICOIACT)* (pp. 349-354). IEEE. <https://doi.org/10.1109/ICOIACT55506.2022.9972091>