Dicle University
Journal of Engineering
https://dergipark.org.tr/tr/pub/dumf
duje.dicle.edu.tr

Research Article

Enhanced Earthquake Magnitude Prediction Using Hybrid Machine Learning and Deep Learning Models

Kerem GENCER[1], İnayet Hakkı ÇİZMECİ[2*]

[1] Afyon Kocatepe University, Computer Engineering Department, keremgencer09@hotmail.com, Orcid No: 0000-0002-2914-1056
[2] Afyon Kocatepe University, Computer Engineering Department, icizmeci@aku.edu.tr, Orcid No: 0000-0001-6202-4807

ARTICLE INFO

Article history:

Received 24 March 2025
Received in revised form 21 May 2025
Accepted 16 June 2025
Available online 30 June

Keywords:

Deep Learning, Artificial Intelligence, Transformers, Earthquake Prediction

DOI: 10.24012/dumf.1663473

* Corresponding author

ABSTRACT

This study evaluates the performance of machine learning and hybrid deep learning models for predicting earthquake magnitudes using historical seismic data. Five models, including Random Forest (RF), ARIMA, Long Short-Term Memory (LSTM), CNN+LSTM, and Transformer + Gaussian Processes (GP), were compared using metrics such as Root Mean Squared Error (RMSE) and R^2. The RF model was quite efficient, with an RMSE of 0.072 and an R^2 of 0.30. However, it did not incorporate temporal analysis. ARIMA was also better, with an RMSE of 0.065 and R^2 of 0.42, which is best suited for linear relationships. LSTM identified the sequential relations well and provided an RMSE of 0.097 and R^2 of 0.51. The hybrid CNN+LSTM model outperformed standalone approaches with an RMSE of 0.090 and R^2 of 0.58 by combining spatial and temporal features. The Transformer + GP model achieved the highest accuracy, with an RMSE of 0.063 and R^2 of 0.62, offering robust uncertainty quantification through confidence intervals. These results highlight the superiority of hybrid models in seismic forecasting, demonstrating their potential to improve predictive accuracy and support better risk management strategies.

Introduction

Earthquakes are natural disasters that cause societies to suffer socioeconomic losses, stop human activities, and disrupt physical structures [1]. When we look at Turkey's history, it is known that it has witnessed many destructive earthquakes. The main reason for destructive earthquakes is that the country is in the 1st degree earthquake zone. In addition, the fact that the city centers, where people live densely, are established on or near the live fault lines known as fractures have also increased the destructive effect [2] (Fig. 1).

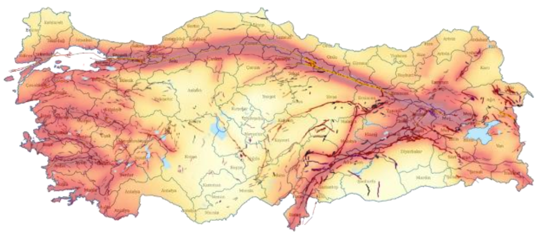

Figure 1. Active fault lines in Turkey (AFAD)

AFAD's report shows that earthquakes cause 60 percent of disaster-related deaths. In 2023, a devastating earthquake of 7.6 Mw occurred in Kahramanmaraş [3]. In the earthquake, 44,218 people died, 80,278 people were injured, and 528,146 people were evacuated to other cities [4]. Knowing the earthquake's magnitude is very important to prepare for such situations. Traditional techniques such as ARIMA are used to detect linear trends in forecasting procedures. However, the proposed methods cannot cope with the difficulties associated with irregular and non-linear seismic data. New forecasting methods can be developed using machine learning and deep learning techniques to extract features and learn temporal patterns. The main work in this field in the last five years has been presented in Table 1.

This study investigates the performance of Machine learning models (Random Forest, ARIMA), Deep learning models (LSTM), and Hybrid architecture (CNN+LSTM, Transformer + GP).

The specific contributions of this research include:

1. Comparing classical statistical models like ARIMA with modern machine learning techniques.

2. Implementing hybrid models such as CNN+LSTM and Transformer + Gaussian Processes for improved accuracy.

3. Incorporating real-time uncertainty estimation using Gaussian Processes.

We aim to evaluate these models using historical earthquake data and provide insights into their strengths and limitations.

Table 1. Major studies for earthquake prediction in the last five years

Researcher (Year)	Method	Objective	Results
Bhandarkar et al. (2019) [5]	LSTM, FFNN	Comparison of earthquake prediction methods	LSTM achieved a 59% better R^2 score than FFNN.
Doğan and Demir (2020) [6]	Structured Recurrent Neural Network (SRNN)	Predicting earthquakes with magnitude four or above in Turkey in a one-month period.	0.72 accuracy and 0.74 precision were obtained.
Li et al. (2020) [7]	CNN, Attention Mechanism (DLEP Model)	Improving earthquake prediction by combining explicit and implicit seismic properties	The DLEP model outperformed the baseline methods.
Berhich et al. (2020) [8]	LSTM, YSA	Performance comparison of ANN and LSTM for earthquake prediction	ANN: MAE 0.30, MSE 0.13; LSTM: MAE 0.11, MSE 0.02 error rates were achieved.
Karcı and Şahin (2022) [9]	LSTM	Estimation of earthquake magnitude and timing	Kalaba-Sivrice and Izmir earthquakes were predicted with 91% and 93% accuracy.
Bhargava and Pasari [10]	LSTM, YSA	Earthquake forecast in the Himalayan region	It is effective in small-scale regions, but insufficient results were obtained for large earthquakes.
Demirelli et al. (2023) [3]	Random Forest, Extreme Gradient Boosting, Decision Trees, k-NN	Earthquake prediction with seismic, geological and geodetic data.	RF and XGB gave the best results with 0.09 MSE.
Doğan (2023) [11]	SVM, Linear Regression, Gradient Boost, Elastic Net, Bayesian Ridge, XGBoost	Estimating earthquake locations and depths in north-western Turkey	Potential high-magnitude earthquake zones were evaluated with RMSE, MAE, and Adjusted R^2.
Kavianpour et al. (2023) [12]	CNN, BiLSTM, Attention Mechanism	Estimating the maximum magnitude and number of earthquakes in different regions in China	The proposed model gave better results compared to other methods.
Ridzwan and Yusoff (2023) [13]	Machine Learning Algorithms	Comparing the effectiveness of different algorithms by analysing 31 studies in 2017-2021	It provided insights into the effectiveness of seismic features and the performance of algorithms.

When the studies presented in Table 1 are examined, certain trends and common findings are observed in the research on earthquake prediction. Deep learning-based methods, especially LSTM models, consistently demonstrate superior performance compared to traditional artificial neural networks and statistical methods. In 2019, Bhandarkar et al. reported that LSTM achieved 59% better R^2 score than FFNN, and in 2020, Berhich et al. reported that LSTM offered significantly lower error rates (MAE 0.11, MSE 0.02) than ANN. Hybrid and advanced models, such as CNNs and attention mechanisms [7][12], provide better results than single-method models. However, the effectiveness of the models varies according to their geographical scope. Although they perform better in small-scale regions, they can also predict large earthquakes in large regions [10]. Ensemble methods such as Random Forest and XGBoost have shown strong performance in studies combining different data types (seismic, geological, geodetic) [3]. Consequently, the most promising approaches for earthquake prediction are deep learning methods, but the prediction accuracy remains highly dependent on the regional context, data quality, and estimated parameters (magnitude, timing, or location).

Materials and methods

Data description

In this study, the earthquake data set recorded by the Disaster and Emergency Management Presidency (AFAD) between 1990 and 2024, with magnitudes between 3 and 6.5 (3447 earthquakes), epicenter in Afyonkarahisar (38.7580, 30.5387), and epicenter distances between 100 km and 150 km (Figure 2). The dataset obtained from AFAD includes various descriptive features for each earthquake event. Specifically, it contains the earthquake ID, date and time of occurrence, latitude, longitude, and depth (in kilometers). Additionally, it provides multiple magnitude scales such as MD (Duration Magnitude), ML (Local Magnitude), Mw (Moment Magnitude), Ms (Surface-Wave Magnitude), and Mb (Body-Wave Magnitude). The dataset also includes the earthquake type and the location name. These variables were used to construct a time-series dataset for training the models, and the ML (Local Magnitude) value was selected as the primary target variable for prediction.

Afyonkarahisar contains 31 live fault lines. The maximum earthquake magnitudes that these fault lines can produce are between 6.18 and 6.86 M_w [14].

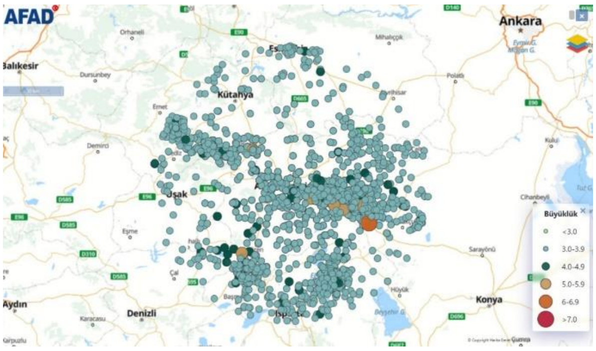

Figure 2. Map of earthquakes within a radius of 150 km, centered in Afyonkarahisar

Moment magnitude scale M_w and Richter local magnitude scale M_L indicate earthquake magnitudes. It has been observed that the scale reaches saturation in high-magnitude earthquakes ($M_L \geq 6$). Conversely, Mw is a scale widely used in large earthquakes, and its structure is more stable [15]. Since there are no earthquakes exceeding magnitude 6 in this dataset, the local magnitude scale M_L was selected for use in the calculations.

The dataset includes earthquake records with magnitude (M_w), converted to Local Magnitude (M_L) using the Equation (1).

$$M_L = 0.67 \, x \, M_w + 1.45 \qquad (1)$$

The data spans multiple years and provides magnitudes for various earthquake events, processed into a time series format [16], [17].

Preprocessing

Missing values in magnitude data were handled using mean imputation. The data was normalized for deep learning models using Min-Max scaling. Features (x) and labels (y) were generated with a rolling window of five years for sequential models.

Models implemented

- **Random forest (RF)**

Random Forest is an ensemble learning method based on decision trees. It was optimized using GridSearchCV to find the best parameters for estimators, depth, and leaf samples. Evaluation metrics include $RMSE$ and R^2 [18].

- **ARIMA**

The ARIMA model analyzes time series data using autoregressive and moving average components. After performing a stationarity (ADF) test, differences were applied for non-stationary series. The model parameters (p,d,q) were selected based on AIC [19].

- **LSTM**

LSTM networks are recurrent neural networks designed to model sequential dependencies: Input shape (timesteps, features), hidden layers including 50 LSTM units with dropout for regularization. The model was trained using the Adam optimizer and mean squared error as the loss function [20].

- **CNN+LSTM hybrid**

Combining CNN for spatial feature extraction with LSTM for temporal dependencies: CNN branch: Two Conv1D layers with ReLU activation followed by a flattening layer: Two LSTM layers with dropout. The final layers were concatenated and fed into dense layers for prediction [21]. The hybrid CNN+LSTM model consists of two main branches. The CNN branch begins with two consecutive 1D convolutional layers, with 32 and 64 filters respectively, both using a kernel size of 2 and ReLU activation function. These layers are followed by a flattening operation. The LSTM branch includes two stacked LSTM layers, each with 50 units; the first returns sequences while the second outputs the final hidden state. The outputs of both branches are concatenated and passed through a dense layer with 50 units and ReLU activation, followed by a dropout layer with a rate of 0.2 to prevent overfitting. Finally, a dense layer with a single unit is used for the regression output.

- **Transformer + Gaussian processes**

This hybrid model leverages Transformers for long-term dependencies and Gaussian Processes for uncertainty quantification. Multi-head attention was used to capture dependencies. Gaussian Processes were implemented for probabilistic predictions with a kernel combining RBF and constant kernels [22]. The Transformer + Gaussian Processes (GP) model combines deep learning with probabilistic modeling to enhance prediction and uncertainty estimation. The Transformer module is composed of a custom encoder block featuring two multi-head attention layers (with 2 attention heads), feed-forward layers with 32 units, and a dropout rate of 0.1 to reduce overfitting. After processing the time-series input, the Transformer's output is passed through a fully connected dense layer with 50 units and an output layer with 1 unit. The model is trained for 50 epochs with the Adam optimizer and a batch size of 32.

Once trained, the predicted outputs from the Transformer model are used as input to a Gaussian Process (GP) regression model, which is configured with a radial basis function (RBF) kernel scaled by a constant kernel. The GP model is optimized with 10 restarts to avoid local minima. The GP outputs both the predicted values and standard deviation estimates, enabling the construction of 95% confidence intervals (±1.96σ), which are visualized in the result figures to represent prediction uncertainty.

Implementation Details and Reproducibility

All experiments were conducted using Python programming language. The main software libraries and their versions used in this study are as follows: TensorFlow 2.12.0, scikit-learn 1.3.0, pandas 1.5.3, numpy 1.23.5, matplotlib 3.7.1, and stats models 0.13.5. To ensure the reproducibility of the models, all key hyperparameters were reported. The batch size was set to 32 and all models were trained for 50 epochs. The learning rate for the Transformer model was set to 0.001 using the Adam optimizer. The dropout rate was fixed at 0.2 across deep learning models to reduce overfitting. The Transformer encoder used 2 attention heads and a feed-forward layer with 32 units. In the CNN+LSTM model, two convolutional layers with 32 and 64 filters and a kernel size of 2 were used, followed by two LSTM layers each with 50 units. For the Gaussian Process regressor, an RBF kernel multiplied by a constant kernel was employed with 10 restarts and alpha set to 1e-2. These implementation details ensure the reproducibility and transparency of the proposed approach.

Evaluation metrics

RMSE (Root Mean Square Error) and R^2 (Determination Coefficient) metrics were used to evaluate the performance of the models. The $RMSE$ is the square root of the mean square of the differences between the model estimates and actual values. The following is expressed in Equation (2). The low values indicate that the model's accuracy is high [23].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2} \tag{2}$$

R^2 represents the ratio of variance in the dependent variable, as explained by the model. It takes values between 0 and 1. A value close to 1 indicates that the model explains the data perfectly, and a value close to 0 indicates that the model loses its explanatory power. The formulation is given in Equation (3) [24].

$$R^2 = 1 - \frac{\sum(y_i - \hat{y_i})^2}{\sum(y_i - \bar{y})^2} \tag{3}$$

These two metrics are complementary to each other. The $RMSE$ indicates the prediction errors, while R^2 shows the explanatory power of the model.

Results

From the comparison of the models presented in Table 2, the strengths and limitations of the models are established based on RMSE and values. For instance, the Random Forest (RF) model has an RMSE of 0.072 and a R^2 of 0.30, which is known to be efficient and fast but has a poor ability to capture temporal characteristics of time series data. This limitation makes it less suitable for the more challenging temporal modeling problems. The ARIMA model, however, is better than RF with an RMSE of 0.065 and an R^2 of 0.42. It can deal with linearity and short-time series trends. However, it cannot simulate complex or nonlinear temporal dependencies like in dynamic systems, such as earthquake magnitude forecasting. The LSTM model has a better temporal modeling capability R^2 of 0.51, meaning it can capture sequential dependencies. However, its RMSE of 0.097 indicates room for improvement, particularly in handling complex spatial-temporal relationships. The hybrid CNN+LSTM model combines the best convolutional neural networks for extracting features from spatial data and recurrent neural networks for data sequences dependent on time. This approach leads to an RMSE of 0.090 and R^2 of 0.58, which proves that the model is efficient in learning the spatial-temporal evolution and performs better than the baseline models such as RF and ARIMA. The Transformer+ Gaussian Processes (GP) model is the best among the models compared with RMSE of 0.063 and an R^2 of 0.62. This hybrid approach combines Transformers' long-term dependency modeling capabilities with the probabilistic insights of Gaussian Processes, providing robust predictions and uncertainty estimation. This makes it the most reliable model for earthquake magnitude predictions, especially in applications requiring accuracy and confidence quantification. In conclusion, while each model has its strengths, hybrid approaches like CNN+LSTM and Transformer + GP emerge superiorly due to their ability to integrate spatial, temporal, and probabilistic aspects into the prediction framework (Table 2).

Table 2. Model comparisons.

Model	$RMSE$	R^2	Key Notes
Random Forest (RF)	0.072	0.30	Simple, fast, lacks temporal insight.
ARIMA	0.065	0.42	Linear dependencies only.
LSTM	0.097	0.51	Strong temporal modeling.
CNN+LSTM	0.090	0.58	Effective spatial-temporal mix.
Transformer +GP	0.063	0.62	Robust uncertainty estimation.

The Random Forest model struggles to capture the temporal trends and provides static future predictions. LSTM shows an improved ability to follow the general trends of the data. It predicts slightly better future values compared to Random Forest. Captures sequential dependencies but may still struggle with complex temporal shifts. ARIMA predicts declining trends for future earthquake magnitudes. It performs better for linear and short-term patterns. This hybrid model captures spatial and temporal dependencies, making more dynamic and accurate predictions. It shows better generalization and flexibility compared to standalone LSTM or Random Forest models. The transformer + Gaussian hybrid model offers the most accurate predictions and incorporates a 95% confidence interval. Predictions are robust, with uncertainty quantified effectively. Combines long-term dependency modeling of Transformers with probabilistic insights from Gaussian Processes, making it the most reliable approach. Transformers + Gaussian Processes excels in accuracy and reliability. Hybrid models consistently outperform standalone methods by leveraging complementary strengths. Hybrid models like CNN+LSTM and Transformer + Gaussian can significantly enhance earthquake prediction systems. In Figure 3, the Random Forest model demonstrates a general ability to capture the trend of earthquake magnitudes. However, it underestimates certain peak values and lacks the ability to model temporal dependencies, which limits its effectiveness in sequential forecasting. In Figure 4, the LSTM model successfully captures temporal dependencies in the historical data. The future predictions appear smoother, indicating the model's tendency toward conservative estimation, though overall trends are reasonably followed. In Figure 5, the ARIMA model shows reasonable performance in modeling linear components of the time series. However, the predicted future values are relatively flat and fail to reflect nonlinear or abrupt changes, which reduces forecasting accuracy in dynamic settings. In Figure 6, the CNN+LSTM hybrid model captures both spatial and temporal features more effectively. The predicted future magnitudes exhibit better alignment with recent patterns, though small deviations remain around highly variable points. In Figure 7, the Transformer + Gaussian Processes model offers both accurate predictions and uncertainty quantification. The future estimates fall within narrow confidence intervals, highlighting the model's robustness and the reliability of its predictive uncertainty.

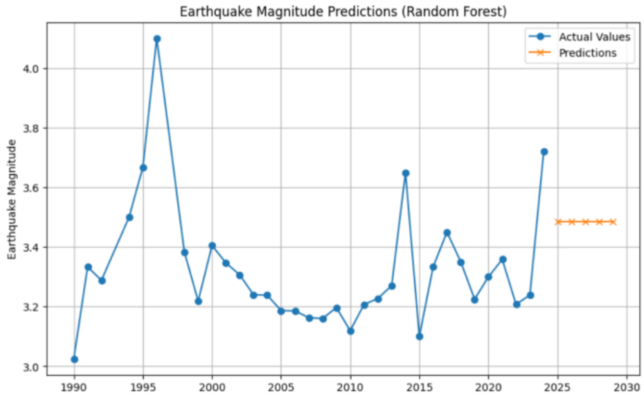

Figure 3. Earthquake magnitude estimates with Random Forest.

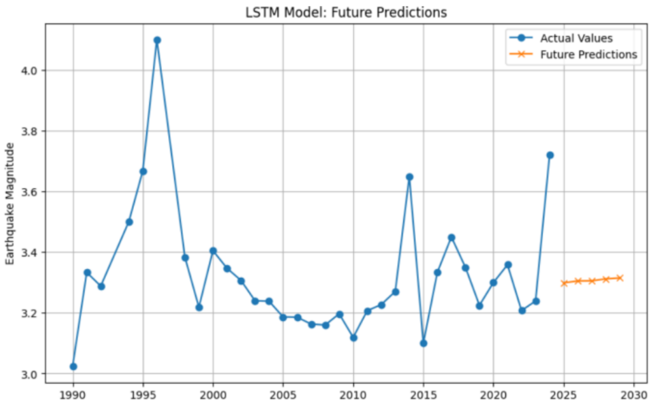

Figure 4. Earthquake magnitude estimates with LSTM.

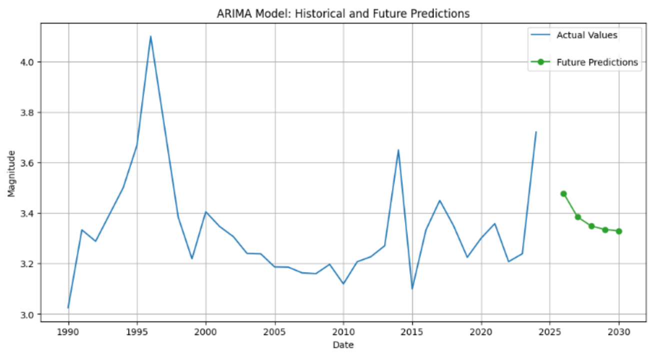

Figure 5. Earthquake magnitude estimates with ARIMA.

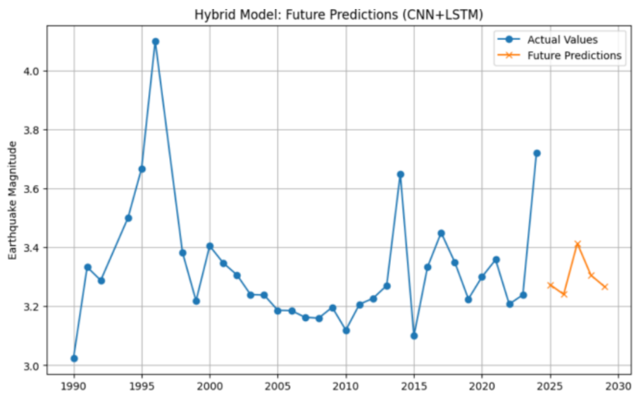

Figure 6. Earthquake magnitude estimates with CNN+LSTM.

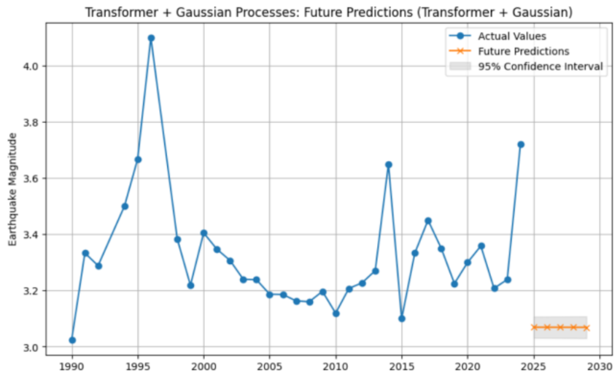

Figure 7. Earthquake magnitude estimates with Transformer + Gaussian.

Discussion

The performance results of the evaluated models provide important information on earthquake magnitude estimation. Traditional models such as Random Forest ($RMSE$: 0.072, R^2: 0.30) and ARIMA ($RMSE$: 0.065, R^2: 0.42) have fundamental limitations in capturing the complex temporal dynamics inherent in seismic data. Although LSTM showed improved temporal modeling (R^2: 0.51), its relatively high $RMSE$ (0.097) points to difficulties in handling the complex spatiotemporal relationships of earthquake events. In particular, the superior performance of hybrid models such as CNN+LSTM ($RMSE$: 0.090, R^2: 0.58) and Transformer + GP ($RMSE$: 0.063, R^2: 0.62) emphasizes the importance of architectures that can simultaneously handle both spatial features and temporal dependencies while providing uncertainty quantification.

These findings highlight a clear trend toward hybrid approaches that integrate the complementary strengths of different modeling paradigms. Recent studies also emphasize the increasing use of deep learning approaches in seismic forecasting, particularly in the Turkish context. For instance, Kas (2023) [25] applied LSTM, GRU, and BiLSTM models to forecast earthquake timing using seismic data from Turkey spanning 1900–2018, highlighting the comparative strengths of different neural architectures. Li et al. (2024) [26] focused on real-time monitoring by utilizing CNN-based models for detecting seismic phases and estimating magnitudes in the 2023 Kahramanmaraş aftershock sequence. Shah et al. (2024) [27] employed Random Forest models for damage prediction based on multiple features such as magnitude, building stability, and population density in Turkish earthquakes. These recent studies support the continued relevance and expansion of AI-based models in earthquake prediction and risk assessment, and are consistent with the hybrid modeling direction taken in this study.

The ability of the Transformer + GP model to combine long-range dependency modeling with probabilistic uncertainty estimation represents a significant advance for practical earthquake prediction applications. This is particularly valuable for early warning systems in which the confidence interval is as important as the point estimate. The performance gap between traditional statistical methods and hybrid deep learning architectures underscores the need for models specifically designed to handle the nonlinear, spatiotemporal properties of seismic data. In addition to prediction accuracy, computational efficiency is a crucial factor, especially in real-time earthquake monitoring and early warning systems. Classical models such as Random Forest and ARIMA are relatively lightweight and can be deployed in time-sensitive environments. Deep learning models like LSTM and CNN+LSTM offer improved accuracy but come with higher computational costs due to their sequential and multi-layered nature. The Transformer + GP model, while the most accurate and informative, requires significant computational resources during both training and inference. Therefore, the choice of model should balance prediction quality with time and resource constraints depending on the specific application.

Conclusion

This study demonstrates how hybrid models, particularly the hybrid deep learning models, can be used to predict earthquake magnitude. The Transformer + GP model was the most efficient, and the best performance was achieved by combining deep learning methods with probabilistic modeling. Some models, such as Transformer + GP, are well suited for seismic early warning systems, mainly due to their predictive capabilities and confidence intervals. The integration of uncertainty quantification offers better decision-making capabilities. The results indicate that hybrid models outperform standard-alone machine learning methods. The CNN+LSTM model efficiently extracted spatial and temporal features, while the Transformer + GP model excelled in quantifying uncertainty, which is crucial for earthquake prediction. Hybrid models can be used in early warning systems to estimate the likelihood and magnitude of future earthquakes. Gaussian Processes add value by offering confidence intervals for decision-making under uncertainty. The data set's size and quality can affect the model's performance. ARIMA as a model is simple and, therefore, cannot handle some complex models. In future studies, comparing model performance across regional and global seismic datasets could help assess the generalizability and robustness of the proposed approaches under different geological conditions.

Future work will require more data and other attributes, such as geographical location, to improve forecast accuracy.

References

[1] Z. Y. İlerisoy, F. Gökşen, and A. Soyluk, "Deprem Kaynaklı İkincil Afetler ve Türkiye Örneklemi / Secondary Disasters Caused by Earthquakes and Turkey Sample," 2022. [Online]. Available: https://www.researchgate.net/publication/357635207

[2] A. C. Yoloğlu and F. Zorlu, "Regional Agglomeration and Disparities in Relation to Earthquake Risks in Türkiye," Journal of City and Regional Planning, vol. 06, no. 01, pp. 78–96, 2024.

[3] E. Demirelli, H. İ. Solak, and İ. Tirkyakioğlu, "Makine öğrenmesi algoritmaları ile deprem katalogları kullanılarak deprem tahmini," Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi, Sep. 2023, doi: 10.17714/gumusfenbil.1268504.

[4] N. Yelboğa, "Kahramanmaraş Depremi Özelinde Travmatik Yas Ve Sosyal Hizmetin Yas Danişmanliği Müdahalesi," Tobider - International Journal of Social Sciences, vol. 7, no. 1, Mar. 2023, doi: 10.30830/tobider.sayi.13.6.

[5] T. Bhandarkar, V. K, N. Satish, S. Sridhar, R. Sivakumar, and S. Ghosh, "Earthquake trend prediction using long short-term memory RNN," International Journal of Electrical and Computer Engineering (IJECE), vol. 9, no. 2, p. 1304, Apr. 2019, doi: 10.11591/ijece.v9i2.pp1304-1312.

[6] A. Doğan and E. Demir, "Earthquake Prediction in Turkey with Structural Recurrent Neural Networks," in 2020 28th Signal Processing and Communications Applications Conference (SIU), IEEE, Oct. 2020, pp. 1–4. doi: 10.1109/SIU49456.2020.9302400.

[7] R. Li, X. Lu, S. Li, H. Yang, J. Qiu, and L. Zhang, "DLEP: A Deep Learning Model for Earthquake Prediction," in 2020 International Joint Conference on Neural Networks (IJCNN), IEEE, Jul. 2020, pp. 1–8. doi: 10.1109/IJCNN48605.2020.9207621.

[8] A. Berhich, F.-Z. Belouadha, and M. I. Kabbaj, "LSTM-based Models for Earthquake Prediction," in Proceedings of the 3rd International Conference on Networking, Information Systems & Security, New York, NY, USA: ACM, Mar. 2020, pp. 1–7. doi: 10.1145/3386723.3387865.

[9] M. Karcı and İ. Şahin, "Derin Öğrenme Yöntemleri Kullanılarak Deprem Tahmini Gerçekleştirilmesi," ARTIFICIAL INTELLIGENCE STUDIES, vol. 5, no. 1, pp. 23–34, 2022, doi: 10.30855/AIS.2022.05.01.03.

[10] B. Bhargava and S. Pasari, "Earthquake Prediction Using Deep Neural Networks," in 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE, Mar. 2022, pp. 476–479. doi: 10.1109/ICACCS54159.2022.9785011.

[11] A. Doğan, "Earthquake prediction for the northwest of Turkey with machine learning methods," Yerbilimleri/ Earth Sciences, vol. 44, no. 2, pp. 166–178, Aug. 2023, doi: 10.17824/yerbilimleri.1325321.

[12] P. Kavianpour, M. Kavianpour, E. Jahani, and A. Ramezani, "A CNN-BiLSTM model with attention mechanism for earthquake prediction," J Supercomput, vol. 79, no. 17, pp. 19194–19226, Nov. 2023, doi: 10.1007/s11227-023-05369-y.

[13] N. S. M. Ridzwan and S. H. Md. Yusoff, "Machine learning for earthquake prediction: a review (2017–2021)," Earth Sci Inform, vol. 16, no. 2, pp. 1133–1149, Jun. 2023, doi: 10.1007/s12145-023-00991-z.

[14] S. Gücek and İ. Zorluer, "Bir Boyutlu Analiz Yöntemiyle Sahaya Özel Sıvılaşma Risk Haritalarının Oluşturulması: Afyonkarahisar Örneği," Afyon Kocatepe University Journal of Sciences and Engineering, vol. 21, no. 4, pp. 908–921, Aug. 2021, doi: 10.35414/akufemubid.930999.

[15] S. Baruah et al., "Moment Magnitude (M W) and Local Magnitude (M L) Relationship for Earthquakes in Northeast India," Pure Appl Geophys, vol. 169, no. 11, pp. 1977–1988, Nov. 2012, doi: 10.1007/s00024-012-0465-9.

[16] T. C. Hanks and H. Kanamori, "A moment magnitude scale," J Geophys Res Solid Earth, vol. 84, no. B5, pp. 2348–2350, May 1979, doi: 10.1029/JB084iB05p02348.

[17] E. M. Scordilis, "Empirical Global Relations Converting M S and m b to Moment Magnitude," J Seismol, vol. 10, no. 2, pp. 225–236, Apr. 2006, doi: 10.1007/s10950-006-9012-4.

[18] J.Brownlee, "Long Short-term Memory Networks with Pyt," 2017.

[19] G. Tunnicliffe Wilson, "Time Series Analysis: Forecasting and Control,5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1," J Time Ser Anal, vol. 37, p. n/a-n/a, Mar. 2016, doi: 10.1111/jtsa.12194.

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

[21] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, 'CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System', IEEE Access, vol. 10, pp. 99837–99849, 2022.

[22] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning. The MIT Press, 2005. doi: 10.7551/mitpress/3206.001.0001.

[23] T. Chai and R. R. Draxler, "Root means square error (RMSE) or mean absolute error (MAE)?," Feb. 28, 2014. doi: 10.5194/gmdd-7-1525-2014.

[24] L. Bouzid, M. A. Yallese, S. Belhadi, and A. Haddad, "Modelling and optimization of machining parameters during hardened steel AISID3 turning using RSM, ANN and DFA techniques: Comparative study," Journal of Mechanical Engineering and Sciences, vol. 14, no. 2, pp. 6835–6847, Jun. 2020, doi: 10.15282/JMES.14.2.2020.23.0535.

[25] Ö. Kas, "Turkey Earthquake Prediction with Deep Learning Algorithms," MSc Thesis, National College of Ireland, 2023. [Online]. Available: https://norma.ncirl.ie/7197/1/omerkas.pdf

[26] W. Li, M. Chakraborty, J. Köhler, C. Quinteros-Cartaya, G. Rümpker, and N. Srivastava, "Earthquake monitoring using deep learning with a case study of the Kahramanmaraş Turkey earthquake aftershock sequence," Solid Earth, vol. 15, pp. 197–210, 2024. doi: 10.5194/se-15-197-2024.

[27] S. Shah, A. Lin, S. Lin, and J. Patel, "Turkey's Earthquakes: Damage Prediction and Feature Significance Using a Multivariate Analysis," arXiv preprint, arXiv:2411.08903, 2024. [Online]. Available: https://arxiv.org/abs/2411.08903