

EVALUATING BASELINE MACHINE LEARNING MODELS FOR EARTHQUAKE MAGNITUDE PREDICTION IN TÜRKİYE

Ayhan AYDIN^{1*}

¹Ondokuz Mayıs University, Faculty of Engineering, 55200, Samsun, Türkiye

Abstract: Due to its tectonic framework, Türkiye faces a constant and severe threat from devastating earthquakes. To mitigate seismic risks, predictive modeling utilizing Artificial Intelligence (AI) methodologies has become increasingly vital. In this study, baseline machine learning models—including Decision Tree, Artificial Neural Networks (ANN), Bayesian Linear Regression, and Linear Regression—were evaluated on the Microsoft Azure Machine Learning Studio platform using historical seismic catalogs (magnitude ≥ 3.5) from the Kandilli Observatory. The performance results demonstrated that models relying solely on historical spatial and temporal parameters yielded insufficient predictive accuracy, with maximum R-squared (R^2) scores reaching only 0.249. These baseline findings confirm that classical catalog-based predictions fail to capture the chaotic, non-linear dynamics of seismicity. The study concludes that to achieve reliable forecasting and life-saving early warning mechanisms, future models must transition from simple historical regressions to multidimensional Deep Learning architectures (such as LSTM or XGBoost) that integrate real-time physical constraints, including GNSS strain data and active fault mechanics.

Keywords: Earthquake prediction, Machine learning, Deep learning, Big data, Seismotectonics, Azure

*Corresponding author: Ondokuz Mayıs University, Institute of Graduate, Department of Smart Systems Engineering, Samsun, Türkiye

E-mail: ayhan.aydin@omu.edu.tr (A. AYDIN)

Ayhan AYDIN



<https://orcid.org/0000-0001-9127-0951>

Received: April 10, 2025

Accepted: June 05, 2025

Published: June 15, 2026

Cite as: Aydın, A. (2026). Evaluating baseline machine learning models for earthquake magnitude prediction in Türkiye. *Black Sea Journal of Artificial Intelligence*, 2(1), 1-5.

1. Introduction

Throughout history, humanity has continuously struggled with natural disasters, with earthquakes being among the most destructive. Positioned within the Alpide-Himalayan orogenic belt, the Anatolian Plate is continuously squeezed between the Eurasian, African, and Arabian plates, creating highly active tectonic systems such as the North Anatolian Fault Zone (NAFZ) and the East Anatolian Fault Zone (EAFZ). These active fault networks have triggered devastating events, such as the 1939 Erzincan, 1999 Gölcük-Düzce, 2020 Samos (Mw 6.9), 2020 Elazığ (Mw 6.8), and the catastrophic 2023 Kahramanmaraş doublet (Mw 7.7 and 7.6). As populations and industrial centers expand near these tectonically volatile zones, the need for robust earthquake forecasting models has become paramount to prevent severe socioeconomic and human losses (Gencer et al., 2025; Khurshudov and Arslan, 2022).

Historically, classical seismology and statistics relied heavily on empirical models or straightforward statistical approaches, such as ARIMA, to identify linear trends in seismic catalogs. Early prediction attempts even observed animal behavior or electromagnetic anomalies. However, the mechanical processes underlying earthquakes—strain accumulation, friction, and viscoelastic relaxation—are highly non-linear, chaotic, and spatio-temporally heterogeneous. Consequently, traditional methods often

fail to capture the complex dynamics of fault ruptures.

1.1. Literature Review and Analysis

In recent years, the exponential growth of Big Data analytics and Artificial Intelligence (AI) has driven a paradigm shift in seismology, transitioning from deterministic constraints to flexible, probabilistic machine learning (ML) models (Khurshudov and Arslan, 2022). AI algorithms can uncover hidden, non-linear patterns across massive, multidimensional datasets that human analysts or classical statistics cannot process.

A critical analysis of the current literature reveals three distinct evolutionary phases in AI-based earthquake prediction in Türkiye:

1.2. Catalog-Based Spatial and Temporal Forecasting

Initial ML studies predominantly utilized historical catalogs. For instance, Khurshudov and Arslan (2026) analyzed over 66,000 seismic events in Türkiye using Random Forest (RF) and Gated Recurrent Unit (GRU) models, demonstrating that AI could effectively simulate small-to-medium earthquake magnitudes based solely on historical coordinates and magnitudes. Similarly, Doğan (2023) utilized algorithms such as XGBoost, Support Vector Machines (SVM), and Bayesian Ridge to successfully predict the spatial focal coordinates (latitude/longitude) and depths of future $M \geq 6.0$ earthquakes in Northwestern Türkiye (Bülteni et al., 2023). However, models relying solely on catalogs suffer



from a fundamental limitation: they record past energy releases but fail to account for the physical, real-time stress currently accumulating in the crust.

1.3. Physics-Informed Geodetic Integration: To address this limitation, researchers began integrating physical constraints into ML datasets. A landmark framework by Demirelli et al. (2023) demonstrated that integrating active geological fault geometries (fault length,

proximity) and GNSS-derived geodetic strain rates with historical catalogs drastically reduced prediction errors. By feeding strain-rate data into the Random Forest and XGBoost algorithms, they achieved an outstanding Mean Squared Error (MSE) of 0.09, demonstrating that combining crustal physics with big data yields vastly superior forecasts. Samiloglu et al. (2025) further validated this by using.



Figure 1. Devastation caused by the 7.7 and 7.6 magnitude twin earthquakes centered in Kahramanmaraş, Türkiye, on February 6, 2023 (*Kahramanmaraş Depreminin Kendine Has Özellikleri | EkoAvrasya Vakfı, 2026*).

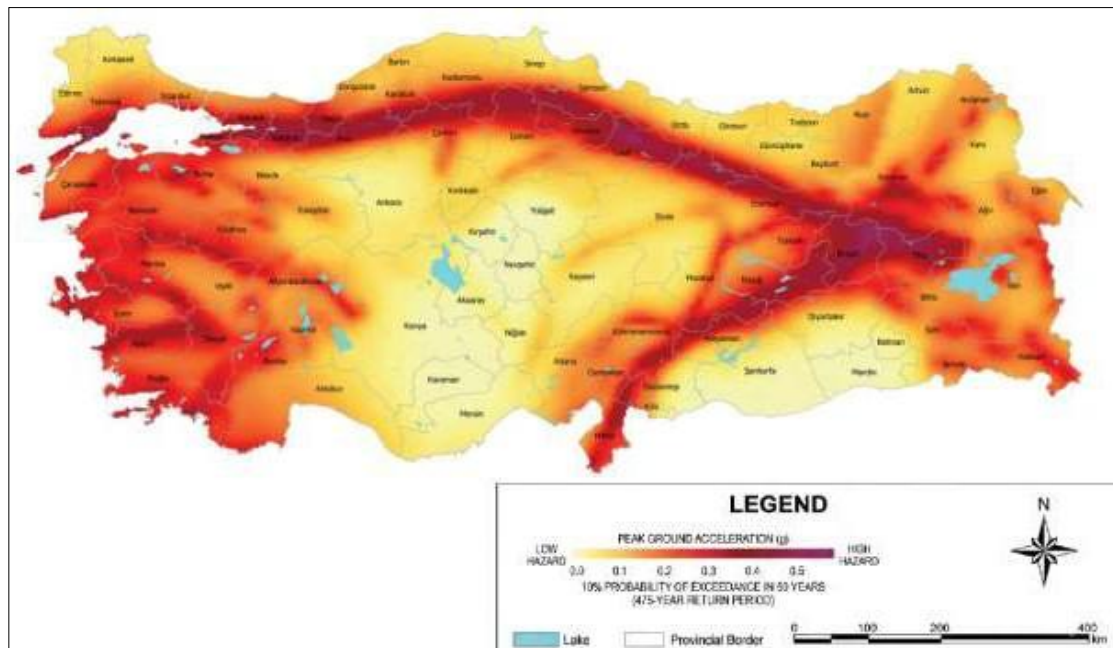


Figure 2. Seismic Hazard Map of Türkiye, illustrating peak ground acceleration (PGA) values with a 10% probability of exceedance in 50 years (475-year return period) (Sahin, 2019).

1.4. The Shift to Deep Learning and Hybrid Architectures

Most recently, researchers have adopted deep learning to capture both temporal memory and spatial relationships. Kaftan (2025) compared five models (LSTM, ANFIS, Decision Tree, Random Forest, CNN) in Western Türkiye and found that Long Short-Term Memory (LSTM) networks outperformed others due to their ability to remember long-term seismic dependencies, achieving an RMSE of 0.1391. Gencer and Cizmeci (2025) introduced even more advanced hybrid architectures, such as CNN+LSTM and Transformer models integrated with Gaussian Processes, to quantify prediction uncertainty, which is highly critical for early warning systems.

Interpretation: The literature strongly indicates that basic regression models are no longer sufficient. The future of seismic forecasting lies in multidimensional, hybrid Deep Learning models that synthesize real-time GNSS strain data, fault mechanics, and historical catalogs (CLAWS Journal Previous Issue, Claws Journal Issue - Centre for Land Warfare Studies (CLAWS), 2026.). The primary objective of the current study is to construct an earthquake magnitude prediction framework utilizing Azure Machine Learning Studio, while critically acknowledging the necessity of transitioning towards these advanced, multidimensional inputs.

2. Materials and Methods

2.1. Computational Platform and Architecture

Initial model development was conducted using Microsoft Azure Machine Learning Studio, a cloud-based platform

that provides a drag-and-drop interface for rapid algorithm deployment and performance evaluation. For advanced deep learning tasks (such as LSTM or CNNs), frameworks utilizing Python environments, including libraries like Pandas (for data manipulation), GeoPandas (for geodetic mapping), and SciKit-Learn, are established as the industry standard for handling high-dimensional seismic computations (Gencer and Cizmeci, 2025).

2.2. Multidimensional Dataset Construction

Based on modern literature best practices, a robust prediction model requires a multi-tiered dataset rather than just a simple catalog:

Historical Seismic Catalogs: The foundational data were acquired from the Boğaziçi University Kandilli Observatory and Earthquake Research Institute (KOERI) and verified against AFAD catalogs. The dataset includes events in Türkiye with magnitudes from 1900 to 2024. Core features include Date, Time, Latitude, Longitude, Depth, and Magnitude.

Geological Fault Parameters: To inject physical constraints, active fault parameters from MTA (General Directorate of Mineral Research and Exploration) must be integrated. These include the total length of the fault, the shortest distance from the earthquake epicenter to the fault line, and the fault mechanism (e.g., normal, strike-slip).

Geodetic GNSS Strain Data: Relative station movements derived from Global Navigation Satellite Systems (e.g., TUSAGA-Aktif) are incorporated to calculate crustal strain and shear stress using Hooke's Law. Incorporating GNSS time series enables algorithms to monitor active tectonic loading and stress drops.

Table 1. Data preview: Recent seismic activity in Türkiye and surrounding regions

Occurrence Date	Occurrence Time	Latitude	Longitude	Depth (km)	Mw	Location
2024-03-31 00:00:00	2024-07- 07T09:33:13.57	403.095	270.207	6.5	3.7	DISBUDAK-LAPSEKI (CANAKKALE) [South East 1.6 km]
2024-03-30 00:00:00	2024-07- 07T05:09:38.65	356.662	35.241	5.3	3.6	MEDITERRANEAN SEA
2024-03-30 00:00:00	2024-07- 07T00:25:03.23	384.867	39.35	5.0	4.7	GUNEYKOY-SIVRICE (ELAZIG) [South East 2.9 km]
2024-03-29 00:00:00	2024-07- 07T20:06:18.97	380.632	369.053	5.3	3.6	ERICEK-GOKSUN (KAHRAMANMARAS) [South East 3.0 km]
2024-03-25 00:00:00	2024-07-	353.425	276.872	16.5	3.6	MEDITERRANEAN SEA
2024-03-25 00:00:00	2024-07- 07T10:10:45.87	378.607	36.267	4.2	3.9	TOPALLAR- SAIMBEYLI (ADANA)

2.3. Data Preprocessing Pipeline

Data preprocessing is an absolute prerequisite for optimizing machine learning model convergence:

Data Cleaning: Rows containing a magnitude of 0 or incomplete geographic coordinates were removed.

Declustering (Noise Reduction): Seismic swarms and

aftershocks heavily bias machine learning models. Algorithms were applied to filter out aftershocks within specific spatial and temporal windows (e.g., removing subsequent events within 15 km and 2 months following a mainshock) to isolate independent tectonic ruptures.

Data Splitting and Normalization: The cleaned dataset was

partitioned into 70% for training and 30% for testing. Min-Max Normalization was subsequently applied to scale all inputs to a common range, preventing features with large numerical ranges (such as fault length) from dominating the learning process.

2.4. Machine Learning and Deep Learning Algorithms

Various algorithms were evaluated to understand their capacity to capture non-linear seismic behavior:

Linear and Bayesian Linear Regression: Foundational statistical techniques that model relationships between independent features and magnitude. Bayesian regression is particularly useful as it incorporates uncertainty in parameters, though linear models generally fail to capture the chaotic nature of seismicity (Pattern Recognition and Machine Learning | Springer Nature Link, 2006).

Decision Trees (DT) and Random Forest (RF): While single Decision Trees are prone to overfitting, Random Forest is an ensemble method that combines multiple trees to improve generalization. RF has proven highly effective when fed multidimensional spatial data, achieving MSEs as low as 0.09 when combining catalogs with fault and GNSS data (Breiman, 2001).

Extreme Gradient Boosting (XGBoost): A highly optimized, sequential ensemble algorithm that aggressively minimizes errors from previous trees. It excels in handling complex spatial boundaries and fault geometries (Chen and Guestrin, 2016).

Long Short-Term Memory (LSTM): A specialized Recurrent Neural Network (RNN) equipped with "forget, input, and output gates" designed to remember long-term temporal dependencies. LSTM is superior for analyzing GNSS time series, as it can track the slow accumulation of stress over decades and correlate it with sudden drops in seismic magnitude (Hochreiter and Schmidhuber, 1997).

2.5. Performance Evaluation Metrics

Model efficacy was evaluated using several standard regression metrics:

MAE (Mean Absolute Error): Measures the average magnitude of errors without considering their direction. A

value closer to 0 indicates high prediction accuracy.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

y_i = Actual values

$\hat{y}_i$ = Predicted values

n = Number of observations

MSE and RMSE (Mean Squared Error and Root Mean Squared Error): Because these metrics square the errors before averaging, they strongly penalize large deviations. In seismology, a low RMSE is critical because missing a high-magnitude earthquake carries catastrophic real-world consequences.

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

R-squared (R^2): Determines the proportion of variance in earthquake magnitude that can be explained by the input features. A score close to 1 implies that the model successfully learned the underlying physical patterns, whereas a low score indicates that the model's predictions are unreliable.

$$R^2 = \frac{SSR}{SST} \quad (3)$$

3. Results

Using the historical catalog of earthquakes ≥ 3.5 magnitude, prediction mechanisms were tested across the four chosen algorithms in Azure. The performance results are summarized below in Table 2. While the MAE and RMSE values appear reasonably low in isolation, they are severely low across all models (ranging from 20% to 25%). Bayesian Linear Regression achieved the best relative result, but due to the low scores, the baseline models failed to capture sufficient variance, rendering these predictions unreliable.

Table 2. Performance evaluation metrics of machine learning models for earthquake magnitude prediction

Model	MAE	RMSE	R^2
Artificial Neural Networks	0.528119	0.891551	0.200725
Bayesian Linear Regression	0.565008	0.882595	0.216703
Linear Regression	0.537651	0.864171	0.249064
	0.538475	0.876844	0.226878

3.1. Comparison with Advanced Literature Findings

The limitations observed in our baseline results are echoed in the advanced literature, which proves that relying strictly on historical catalogs is insufficient. For instance, Demirelli et al. (2023) demonstrated that when historical catalogs are augmented with GNSS strain rates and active fault line geometries, ensemble methods like Random Forest and XGBoost achieve an outstanding MSE of 0.09. Furthermore, Deep Learning methods drastically outperform traditional algorithms for time-series seismic

data. Kaftan (2025) utilized a Long Short-Term Memory (LSTM) network for Western Türkiye, achieving highly accurate predictions with an RMSE of 0.1391 and an MAE of 0.1046. Similarly, Samiloglu et al. (2025) integrated GNSS-derived stress accumulation data with an LSTM architecture along the North Anatolian Fault Zone, successfully modeling the non-linear build-up and drop of physical stress that correlates directly with earthquakes.

4. Discussion and Conclusion

In this study, an earthquake magnitude prediction framework was evaluated using historical data and Big Data methods on Azure Machine Learning Studio. The experimental results demonstrated that baseline algorithms—Decision Trees, ANNs, Bayesian Linear Regression, and Linear Regression—produced inadequate scores (max 0.249) when trained solely on basic historical parameters (coordinates, depth, time).

The poor performance of the baseline model emphasizes a crucial scientific reality: earthquake forecasting cannot be reduced to simple historical coordinate regression. Earthquakes are driven by chaotic, non-linear mechanisms involving active crustal deformation and stress accumulation (Kaftan, 2025). Our model lacked critical physical input parameters, which severely limited its predictive power.

The integration of current literature reveals the necessary path forward. To accurately predict complex seismic events, prediction models must transition to Multidimensional Deep Learning Architectures. By incorporating real-time GNSS geodetic strain measurements and proximity to geological faults into algorithms such as LSTM, XGBoost, or Random Forest, models can physically track active tectonic loading of the crust.

For future work, it is essential to optimize the dataset by consulting domain experts to incorporate physical constraints (e.g., fault types, continuous GNSS time series, and geodetic stress drops) and to deploy advanced recurrent neural networks (such as LSTM) that can capture long-term seismic dependencies. Transitioning to these physics-informed AI models will yield highly reliable earthquake early-warning mechanisms, ultimately saving lives and mitigating disaster risks.

Author Contributions

The percentages of the authors' contributions are presented below. All authors reviewed and approved the final version of the manuscript.

	A.A.
C	60
D	50
S	40
DCP	50
DAI	30
L	
W	40
CR	50
SR	
PM	
FA	50

C= concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management, FA= funding acquisition.

Conflict of Interest

The authors declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because of there was no study on animals or humans.

Data availability

For verification and permission requests, please contact Ayhan AYDIN (<https://github.com/AyhanAydinPhD>). In accordance with the Regulation on the Protection of Data, researchers are prohibited from sharing their data with third parties. The data may be used in scientific studies provided that permission is obtained from the relevant institutions. The authors have shared the code generated in this study after the project was completed. These can be requested via Ayhan AYDIN's GitHub page or by email.

Declaration of generative AI and AI-assisted technologies

In the manuscript preparation process, the authors used Gemini (Google) and ChatGPT (OpenAI) to improve the clarity and readability of the English text. This involved using specific prompts such as: "Please improve the academic tone and grammatical flow of the following section while preserving the technical meaning." After using these tools, the authors carefully reviewed and edited the content as needed and take full responsibility for the integrity of the published article.

References

- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324/METRICS>
- Bülteni, M., Doğan, A., Üniversitesi, H., Osb, B., Meslek, T. B., Okulu, Y., & Ankara, T. (2023). Makine öğrenimi yöntemleri kullanılarak Türkiye'nin kuzeybatısı için deprem tahmini. *Yerbilimleri*, 44(2), 166–178. <https://doi.org/10.17824/YERBILIMLERI.1325321>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785;CSUBTYPE:STRING:CONFERENCE>
- CLAWS Journal Previous Issue, *Claws Journal Issue - Centre for Land Warfare Studies (CLAWS)*. (n.d.). Retrieved June 28, 2026, from <https://archive.claws.co.in/journal/journal-claws-journal/index.html>
- Gencer, K., & Cizmeci, İ. H. (2025). Enhanced Earthquake Magnitude Prediction Using Hybrid Machine Learning and Deep Learning Models. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 16(2), 369–376. <https://doi.org/10.24012/DUMF.1663473>
- Gencer, K., Fakültesi, İ. C.-D. Ü. M., & 2025, undefined. (2025). Enhanced Earthquake Magnitude Prediction Using Hybrid Machine Learning and Deep Learning Models. *Dergipark.Org.TrK Gencer, İH CizmeciDicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 2025•dergipark.Org.Tr, 16(2), 369–376*. <https://doi.org/10.24012/dumf.1663473>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/NECO.1997.9.8.1735>

- Kahramanmaraş Depreminin Kendine Has Özellikleri / EkoAvrasya Vakfı.* (2026). Retrieved June 28, 2026, from <https://ekoavrasya.net/kahramanmaraş-depreminin-kendine-has-ozellikleri/>
- Khurshudov, E., & Arslan, A. (2022). *From Historical to AI-Based Approaches in Earthquake Forecasting: The Case of Türkiye.* <https://doi.org/10.36937/cebel.2026.11114>
- Pattern Recognition and Machine Learning / Springer Nature Link.* (n.d.). Retrieved June 28, 2026, from <https://link.springer.com/in/book/9780387310732>
- Sahin, S. (2019). Türkiye’de Afet Yönetimi ve 2023 Hedefleri. *Türk Deprem Araştırma Dergisi*, 1(2), 180–196. <https://doi.org/10.46464/TDAD.600455>
- Şamiloğlu, F. B., Işık, V., Aktuğ, B., & Şamiloğlu, A. T. (2025). Bolu İli Çevresinde Geçmiş Deprem Kataloğu ve GNSS Zaman Serilerinin Derin Öğrenme ile Deprem Tahmininde Kullanımı. *Turkish Journal of Earthquake Research*, 7(3), 571–589. <https://doi.org/10.46464/TDAD.1706258>
- Sciences, I. K.-A., & 2025, undefined. (n.d.). Machine Learning Applications for Earthquake Magnitude Prediction in Western Turkey. *Mdpi.ComI KaftanApplied Sciences*, 2025•*mdpi.Com*. Retrieved June 28, 2026, from <https://www.mdpi.com/2076-3417/15/20/10909>