SETSCI Conference Proceedings
18, 125-128, 2024

6th International Symposium on Innovations in Scientific Areas
7-8 June 2024, Ankara – Türkiye
Section: Engineering
https://doi.org/10.36287/setsci.18.1.00125

Earthquake Prediction Methodology Based On Neural Network Model

Azra Dilek Erhan[1*], Ikhwan Kim[1]

[1]*Landscape Architecture, Istanbul Technical University, Istanbul, Türkiye*
(erhan22@itu.edu.tr)
(iikimss3@gmail.com)

Abstract – Forecasting earthquakes is challenging due to the difficulty in detecting underground phenomena. However, statistically, 30% of large earthquake foreshocks occur before the main shock, which indicates that, theoretically, it is possible to forecast them. However, systematically collecting historical earthquake data has been challenging because of the lack of computational data collection and procedures. Addressing these limitations in historical data collection, we have compiled a database of earthquake events in Türkiye from 2004 to 2023, encompassing 187,399 cases with magnitudes ranging from 2.0 ML to 7.0 ML. This database includes parameters such as latitude, longitude, date, and magnitude. With the database, we built a neural network model capable of predicting earthquakes over 2.0 ML across Turkey up to one month in advance. To evaluate the model, we made it to predict earthquakes in January 2024 and matched the result to the actual events. The pattern between the two was 98% the same. Additionally, the model forecasted an earthquake over 4.0 ML in Istanbul between February and March, which happened on February 19th. Though this prediction model can predict one month's future and various evaluations must be made throughout the time, this paper carries the value of potential.

Keywords – *Earthquake Prediction, Neural Network Model, Artificial Intelligence, Machine Learning, Türkiye*

I. INTRODUCTION

Earthquakes are known for their immense destructive capacity. Türkiye, a country prone to seismic activity, has experienced numerous destructive earthquakes, including the Kahramanmaraş earthquake on February 6th, 2004. This 7.7 magnitude quake resulted in the tragic loss of over 50,000 lives [1]. Consequently, the ability to forecast earthquakes has become crucial.

Predicting earthquakes presents challenges, mainly because of tectonic activity's complex and dynamic nature [2]. Yet, statistically, 30% of large earthquake foreshocks occur before the main shock [3], suggesting that it is theoretically possible to predict them. Recent advances in artificial intelligence (AI) and machine learning have opened new frontiers, enabling researchers to address former challenges. The neural network model represents a significant methodology in earthquake prediction. This research explores how neural networks can help us better predict earthquakes by using past seismic data to estimate the possibility of future earthquakes.

An earthquake forecast estimates the probability of an earthquake occurrence, considering the size, frequency, location, and time of earthquakes. The ability to predict earthquakes has been limited by several factors, including detecting subtle underground phenomena [4] that precede major seismic events. Seismic data has been limited due to constraints on equipment capabilities. However, seismic activities are calculated and recorded in an earthquake catalog, which is fed to algorithms with machine-learning models for earthquake predictions. With the aid of these records, upcoming earthquakes can be predicted despite the difficulties and limitations related to detecting seismic events and complex features using time-stamped sequences of previous earthquakes [3].

II. MATERIALS AND METHOD

Efforts by Turcotte and Donald and Tehseen et al. [5] have focused on analyzing past earthquake patterns to predict future events, and researchers have debated this topic. Geller et al. [6] claim it is impossible to forecast when an earthquake will occur, although Anderson [7] has considered the possibility of establishing statistical techniques for applying the Bayesian model to forecast both short- and long-term earthquakes. Varotsos et al. [8] presented an approach for experimental earthquake forecasting, whereas Wang [9] proposed a neural network model utilizing single multi-layer perception to estimate seismic activity. However, obstacles like the complexity of seismic data records and the constantly changing capabilities of detection equipment have frequently hindered these attempts. New models have started to appear due to complex computing methods and the spread of artificial intelligence. Reyes et al. [10], Wang et al. [11], Manral et al. [12], Handan Çam, and Osman Duman [13] have made strides in applying machine learning algorithms to seismic data. Nonetheless, these models have struggled with issues of data purity and the objective evaluation of their predictions. For instance, the apparent increase in minor seismic events in datasets may reflect not an actual increase in such events but rather improvements in seismic detection technology. Researchers collected data from different timelines in different intervals and studied different numbers of cases. Even though their time interval was wider and the data structurally proper, the purity of the data they used is questioned. Wang et al. collected case data from 1970 to 2021, and Manral et al. used data collected since 1965. According to their data, the frequency of earthquakes under 2.0 ML has steadily increased annually. The frequency of weak earthquakes may have increased over time; however, it is more logical to assume that the ability to detect minor earthquakes has developed over time due to technological advances. Despite the promising attempts of former prediction methodologies, they were limited in providing accurate forecasts and lacked objective evaluation phases in their analyses. This suggests that currently, there are no statistical or physical models capable of accurately predicting large earthquakes. Nonetheless, applying machine learning and statistical methods to earthquake prediction presents a promising method for developing more precise and dependable outcomes.

This research has taken an alternative approach by building a neural network model that has been specially trained on credible and meticulously selected data. Using a machine learning model, specifically a Multilayered Perceptron (MLP) neural network. The model focuses on the historical earthquake data from Turkiye, collected from 2004 to 2024, coinciding with the standardization of seismic data collection by the AFAD [14]. The dataset consists of coordinates (longitude and latitude) with a magnitude range from 2.0 ML to 7.0 ML, which is recorded as a time series.

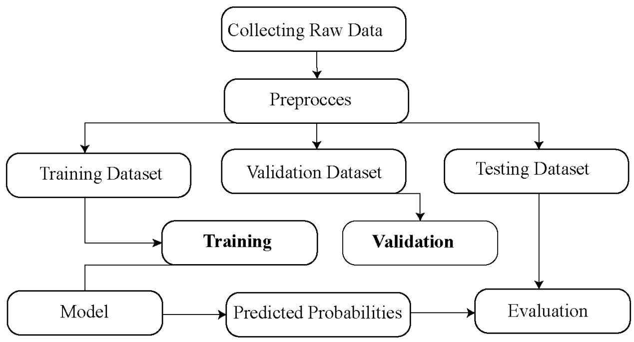
Table 1 Collected data format

Date	Longitude	Latitude	Magnitude
2023-12-30 22:59:34	38.7836	38.2900	2.4
2023-12-30 20:43:24	29.1583	40.2431	3.4
2023-12-30 20:30:52	36.1314	38.2206	2.4
2023-12-30 16:46:06	36.2194	37.8253	2.5

The selected dataset provides consistent and systematically recorded seismic events that have been detected with the same system since 2004, ensuring the data's quality. Although large earthquakes often recur at lengthy intervals (hundreds to thousands of years) [15], making it challenging to identify patterns over a prolonged time frame due to consistency in data collection methods, the time period selected provides a valid foundation for training the neural network.

A neural network is a machine-learning model replicating the human brain's structure and functions. It comprises a network of interconnected nodes, or neurons, that analyze information and generate hypotheses as a whole. The first step is preparing the raw data for neural network models to work and conduct efficiently. Data preparation involves selecting the relevant features, which serve as the primary input variables for the model. To ensure the integrity of the findings, it is essential to confirm the accuracy and reliability of the data and remove any duplicates or outliers from the dataset. The training data quality influences the machine-learning model's accuracy and performance. Also, processes such as data cleaning, formatting, and transformation are essential steps for not just machine learning algorithms' performance but also data analysis and the detection of relations among related features due to choosing the right machine-learning model.

Table 2 Structure of the process



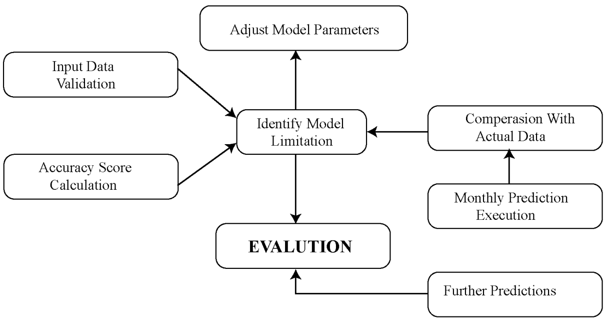
Additionally, we created a binary target variable to determine whether the magnitude reached or exceeded a magnitude of 4.0 ML or above. The binary classification approach simplifies the problem of predicting significant earthquakes. The methodology enables us to discover relations among features, sub-datasets, and time servers by clustering, grouping, pattern selection, and visualization. Given the varying scales of the geospatial data, feature scaling was necessary to normalize the data and ensure effective convergence during neural network training. A standard scaler was used to remove the mean and scale the features to unit variance to improve neural network performance and stability. The neural network model is suitable for binary classification since it balances computing efficiency and model complexity [16]. The model is trained with a maximum of 500 iterations, which allows sufficient opportunity for the network's weights to converge to an optimal state. Fixed random state ensures that the results are reproducible.

New data consisting of locations and a target date ("2024-04-01") were prepared to demonstrate the model's application. The coordinates for these locations correspond to major cities in Türkiye, each represented by its longitude and latitude. Data undergoes the same scaling transformation as the training data to ensure the model receives input in the expected format. The trained model predicts the probability of each location experiencing an earthquake of magnitude 4.0 ML or above on the specified target date. For the purpose of the study, we focus on the probability associated with the positive class (earthquake magnitude ≥ 4.0). The probabilities are then reported for each city, providing insights into potential earthquake risks in different regions of Türkiye. This predictive capability highlights the model's practical application in earthquake risk management and preparedness planning.

III. RESULTS & DISCUSSION

Because of how machine learning algorithms are built, evaluating them involves a few distinctive approaches. One of the standards evaluating methods is the accuracy score [17]; the accuracy score of a machine learning algorithm is a measure to evaluate the performance of a classification model. It represents the proportion of correct predictions made by the model compared to the total number of predictions made. The accuracy score provides information about the quality of the input data and the reliability of the pattern.

Table 3 Evaluation process



This evaluation technique showed that the model could forecast earthquakes accurately (98%). Although accuracy is a simple and obvious metric, it has limitations, especially when dealing with imbalanced class distribution datasets. Due to these limitations, the algorithm made monthly predictions as target dates between February and March (one-month interval) 2024 and successfully predicted a significant earthquake in İstanbul, demonstrating its practical efficacy. Along with Istanbul, predictions for three other cities, İzmir, Ankara, and Antalya, have been made with probabilities of 97.67%, 51.80%, and 82.69%. These cities were selected because they are major urban areas with high population densities. To ensure the accuracy of the evaluation phase, both high-risk and low-risk areas were included in the selection. In all cities except Ankara, earthquakes occurred within the predicted time interval and magnitude. Due to the evaluation being developed, we predicted ten randomly selected cities besides these major cities, aiming to keep the evaluation's validity as high as possible for April 2024.

Table 4 Prediction & Real Events (February-March)

Prediction			→	Actual event		
Cities	Earthquakes	Possibility(%)		Cities	Time	Strenght
Antalya	over 4.0	82,69		Antalya	2024-03-10	4.7
	over 6.0	0.28				
Izmir	over 4.0	98.91		Izmir	2024-02-24	4.2
	over 6.0	0.46				
Istanbul	over 4.0	97.67		Istanbul	2024-02-19	4.0
	over 6.0	0.31				
Ankara	over 4.0	51.80		Ankara	-	-
	over 6.0	0.15				

According to Table 5, April's prediction indicated that among ten cities, three—İstanbul, İzmir, and Afyonkarahisar—were predicted to experience earthquakes with a magnitude above 4.0. However, an earthquake occurred only in İzmir. In this evaluation phase, the algorithm continues to test and train itself using cumulative data.

Table 5 Prediction & Real Events (April)

Prediction			→	Actual event		
Cities	Posibility(%)			Cities	Date	Strength
	Over 4.0	Over 6.0				
Ankara	46.86	0.42		Ankara	-	-
İstanbul	96.80	1.15		İstanbul	-	-
İzmir	98.83	3.54		İzmir	19/04/2024	4.5
Antalya	74.20	0.92		Antalya	-	-
Kayseri	5.73	0.08		Kayseri	-	-
Bayburt	1.17	0.00		Bayburt	-	-
Batman	1.22	0.00		Batman	-	-
Samsun	3.80	0.05		Samsun	-	-
Ağrı	3.04	0.00		Ağrı	-	-
Adana	4.53	0.07		Adana	-	-
Afyonkarahisar	83.14	0.76		Afyonkarahisar	-	-

IV. CONCLUSION

In conclusion, this study developed a neural network model to predict significant future earthquakes in Türkiye, specifically targeting events above 4.0 ML magnitude, using artificial intelligence's capabilities through a Multilayer Perceptron (MLP) classifier. A statistical approach and historical earthquake data with characteristics like locations, magnitude, and time of occurrence were used to build the methodology. This work employs real-world events and contemporary computer tools to improve forecast accuracy. The primary achievement of this research is establishing a model that predicts earthquake probabilities for specific geographic locations with a high degree of accuracy. According to our evaluations, 98% similarity has been detected between the model's predictions for January 2024. Also, actual occurrences show the potential of neural network models in earthquake predictions and risk assessments.

Furthermore, the successful prediction of the earthquake in Istanbul (February 2024) provides a practical example. While the model shows optimistic results, it has its limitations. The current model predicts only up to one month in advance due to input data limitations, which prevents long-term planning risk management. April results indicate a need for further advancement in the algorithm's long-term evaluation phases. Additionally, the model's reliance on data from a uniform detection system means that its performance depends on the quality and amount of the input data. Future research could explore additional variables and more complex network architecture to extend the model forecasting timeline. Additionally, incorporating more diverse data sources and testing models on different geographic regions can help develop the model. Implementing a real-time data system could help increase the models' prediction accuracy. Ultimately, the successful neural network model represents a significant step toward reducing this disaster's destructive nature. We can lower risks by combining machine learning techniques with conventional statistical methods. Further research and developments are essential to realize the potential of AI in this critical area.

References

[1] Siat Yıldırım, *An Observation Study On The Psychosocial Effect of The 6 February Kahramanmaraş Earthquake And The Importance Of Sustainable Response To Earthquake Victims*, 2023.

[2] Amirul Hoque, Jyoti Raj , Dr. Ashim Saha. "*Approaches of Earthquake Magnitude Prediction using machine learning techniques.*". International Conference on Computational Intelligence & IoT (ICCIIoT) (2018): 1-5.

[3] Wyss, Max. "*Cannot earthquakes be predicted?.*" Science 278.5337 (1997): 487-490.

[4] Braile and Brehm 1998; Knopoff 2000; Kirschvink 2000; Ellsworth et al. (1999)

[5] Turcotte, Donald L. "*Earthquake prediction.*" Annual review of earth and planetary sciences 19.1 (1991): 263-281.

[6] Geller, Robert J., et al. "*Earthquakes cannot be predicted*." Science 275.5306 (1997): 1616-1616.

[7] Anderson, J.G., 1981. A simple way to look at a Bayesian model for the statistics of earthquake prediction. Bulletin of the Seismological Society of America, 71(6), pp.1929-1931.

[8] Varotsos P. and Alexopoulos K., 1984. Physical properties of the variations of the electric field of the earth preceding earthquakes, I. Tectonophysics, 110(1-2), pp.73-98

[9] Wang W., Cao X. and Song X., 2001. *Estimation of the Earthquakes in Chinese Mainland by using artificial neural networks*. Chinese Journal of Earthquakes, 3(21), pp.10-14.

[10] Reyes, Jorge, Antonio Morales-Esteban, and Francisco Martínez-Álvarez. "*Neural networks to predict earthquakes in Chile*." Applied Soft Computing 13.2 (2013): 1314-1328.

[11] Wang, Xi, et al. "*Small Earthquakes Can Help Predict Large Earthquakes: A Machine Learning Perspective.*" Applied Sciences 13.11 (2023): 6424.

[12] Manral, Gaurav Singh, and Alka Chaudhary. "*Prediction of Earthquake Using Machine Learning Algorithms*." 2023 4th International Conference on Intelligent Engineering and Management (ICIEM). IEEE, 2023.

[13] Handan Çam, Osman Duman, "Yapay Sinir Ağı Yöntemiyle Deprem Tahmini: Türkiye Batı Anadolu Fay Hattı Uygulaması",2016

[14] https://www.afad.gov.tr

[15] Liu, M.; Stein, S. Mid-continental earthquakes: Spatiotemporal occurrences, causes, and hazards. *Earth-Sci. Rev.* **2016**, *162*, 364–386.

[16] G. P. Zhang, *"Neural networks for classification: a survey.",* Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 30, pp. 451-462, 2000.

[17] Isaac Martín De Diego, Ana R. Redondo, Rubén R. Fernández, Jorge Navarro, Javier M. Moguerza. "General Performance Score for classification problems". 2022