Journal of Artificial Intelligence and Data Science (JAIDA)
e-ISSN: 2791-8335
https://dergipark.org.tr/pub/jaida
Vol. 4, No. 2, pp. 59-67, 2024
Research Article
Received: 22/10/2024 Accepted: 13/12/2024

Earthquake Probability Prediction with Decision Tree Algorithm: The Example of Izmir, Türkiye

İsmahan ERMİŞ [1*], İsa CÜREBAL[2]

Abstract

This study investigates earthquake records in the Izmir province of western Türkiye, focusing on seismic activity prediction through the application of decision tree models. Utilizing earthquake data from 1900 to 2024, including magnitude, depth, latitude, and longitude variables, the aim is to estimate future seismic events in a region known for its significant earthquake risks. The decision tree model, a machine learning approach, was trained with 80% of the dataset and tested on the remaining 20%. Performance was assessed using metrics such as precision, recall, F1 score, and overall accuracy, with the model achieving an accuracy rate of 92%. However, its ability to predict larger earthquakes was hindered due to the limited availability of data for higher-magnitude events. A chi-square test demonstrated a statistically significant relationship between earthquake depth and magnitude. Additionally, a risk analysis map was created using Geographic Information Systems (GIS), highlighting fault lines and areas prone to frequent seismic activity. The study concludes that while the decision tree model is effective for predicting smaller earthquakes, the accuracy for larger events could be improved with more comprehensive data. These findings underscore the importance of targeted earthquake preparedness in Izmir, particularly in coastal areas susceptible to both seismic events and secondary hazards like tsunamis.

Keywords: Artificial Intelligence; Decision Trees; Earthquake; Izmir.

1. Introduction

This study aims to analyze the earthquake records of Izmir, one of the most important provinces in western Türkiye, and its immediate surroundings (between 37° 45' and 39° 15' north latitude, 26° 15' and 28° 20' east longitude) in the paleo-seismology based on magnitude, depth, latitude and longitude variables. This earthquake data analysis aimed to estimate the magnitudes and locations of earthquakes that may occur in the future. The study's outputs are expected to be beneficial for Izmir, which is in a position to have serious risks in terms of seismic activity in Türkiye.

Various methods have been used to predict and analyze earthquakes. Other machine learning algorithms such as Decision Trees (DTs), Support Vector Machines (SVM), Random Forests (RF) and Neural Networks (NN) have also been used to improve earthquake prediction models [31]. For example, Support Vector Machines (SVM) [18], Random Forests Algorithm (RFA) [17], K-Nearest Neighbors (KNN) [22], Long Short-Term Memory (LSTM) [7] and Decision Trees (DTs) [2, 3, 6, 8, 12, 24, 29, 33, 36, 37, 39]. In this study, the decision trees method, which is assumed to provide accurate and reliable earthquake predictions, was preferred [30].

The decision trees method, a popular machine learning algorithm, is widely used in various studies to predict earthquake magnitudes and assess the impact of seismic events [16, 26]. The decision tree algorithm, especially the C4.5 variant, shows promise in characterizing factors that predict earthquakes and developing decision rules based on environmental variables and seismic data [4, 19]. Decision trees can identify important predictors by analyzing historical earthquake data and rank attributes using information theory criteria such as entropy and frequency of occurrence [4].

Located in the Coastal Aegean section of the Aegean Region in western Türkiye, Izmir is the third largest city in the country in terms of both population and economic importance. Izmir and the surrounding region is seismically active and has experienced damaging earthquakes in the past important human and economic effects. The tectonic structure of Western Anatolia, characterized by complex fault systems, contributes to the high seismic hazard in the region [13, 15, 27, 34]. The study focuses on the seismic activity in Izmir and examines whether decision trees can effectively handle complex seismic patterns in the region.

*Corresponding author

İsmahan ERMİŞ; Balıkesir University, Institute of Social Sciences, Division of Geography, Balıkesir, Türkiye; e-mail: ermisismahan@hotmail.com;
0009-0007-7899-645X

İsa CÜREBAL; Balıkesir University, Faculty of Art and Science, Department of Geography, Balıkesir, Türkiye; e-mail: curebal@balikesir.edu.tr;
0000-0002-3449-1595

The research questions of this study are as follows;

MAIN QUESTION: When the past earthquake data are analyzed, can the characteristics and distribution of earthquakes that may occur in the future in Izmir be predicted?

SUBQUESTION 1: How accurately can future earthquakes be predicted with the decision tree model?

SUBQUESTION 2: What are the advantages and limitations of the decision tree model?

SUBQUESTION 3: Is there a statistical relationship or similarity between the magnitudes and depths of earthquakes in Izmir province?

SUBQUESTION 4: What is the distribution of earthquakes that occurred in Izmir province in the last century and in which areas are these earthquakes concentrated?

2. Methodology

In this study, a dataset of earthquake records provided by Boğaziçi University Kandilli Observatory and Earthquake Research Institute (KRDAE) was used. This dataset, which covers the earthquakes that occurred between 1900 and 2024 (01.10.2024), includes the year, depth, latitude, longitude and magnitude information of the earthquakes.

The study focuses on earthquakes with magnitudes (xM) between 3.0 and 8.0, and the data set is organized and visualized before applying the decision tree algorithm.

The decision tree model was implemented in Colab environment and 80% of the data was allocated for training and 20% for testing (Figure 1). Model fitting was applied on the separated test - training set and the results of this model were obtained. When latitude, longitude and depth information was given to the model as the outputs of the results, the model was trained to give as a result output the magnitude of potential future earthquakes based on the given information.

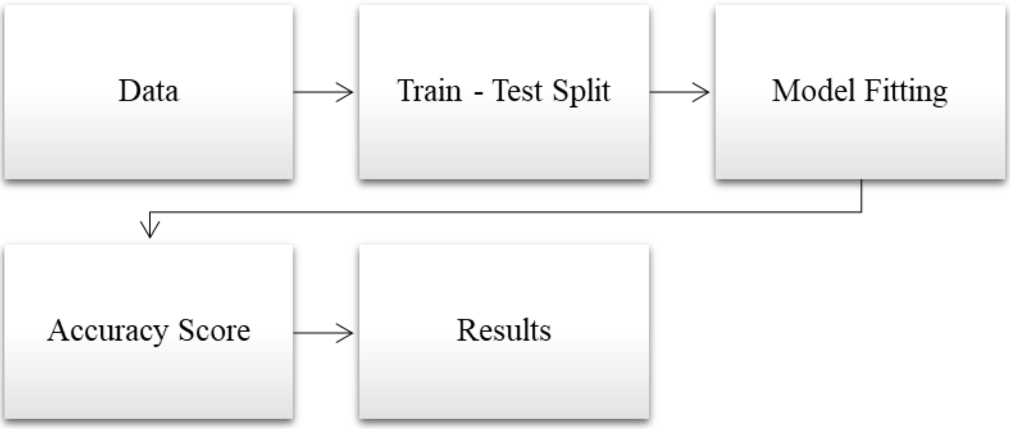

Figure 1. *Flow Diagram of Methodology*

The decision tree algorithm is first applied to the seismic dataset to generate binary predictions. The process starts by selecting a root node from the original dataset. In each iteration, the algorithm calculates the information gain for all available attributes. The attribute with the highest information gain is selected and the dataset is divided into subsets based on this attribute. The algorithm then processes each subset iteratively, considering only the remaining attributes that were not used in the previous splits. This iterative process continues until the decision tree is fully formed (Figure 2). The path from the root to the leaf node represents the values of the input variables, while each leaf node corresponds to a predicted value of the target variable. Decision trees are highly effective as classifiers that can capture complex variations in data [5]. This algorithm was used to estimate the magnitude of earthquakes based on their latitude, longitude and depth.

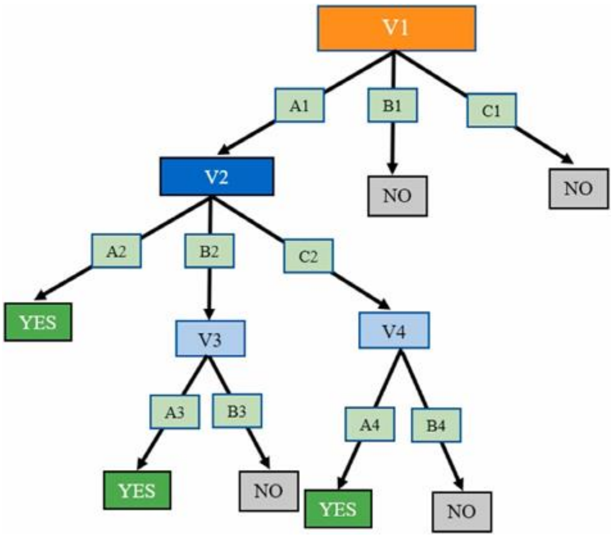

Figure 2. *Decision Trees (DTs) Flowchart [21, 25]*

We started by first analyzing the prediction performance of the method used for the prediction results of the model with a self-consistency test on the training set. Then, in order to evaluate the overall performance of the method more comprehensively, accuracy, recall and F1 score were calculated among the metrics used for performance analysis.

$$Precision\ (P) = \frac{TP}{\mathrm{TP} + \mathrm{FP}} \tag{1}$$

$$\mathrm{Recall\ (R)} = \frac{TP}{\mathrm{TP} + \mathrm{FN}} \tag{2}$$

$$\mathrm{F1 - Score} = \frac{2PR}{\mathrm{P} + \mathrm{R}} \tag{3}$$

$$\mathrm{Accuracy} = \frac{TP + TN}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \tag{4}$$

In the above equations TP, FN, TN and FP are true positives, false negatives, true negatives and false positives respectively. Precision is the proportion of predicted hotspots that are true hotspots. Recall is the proportion of true positive hotspots that are predicted hotspots. F1 score is a measure to balance recall and precision. Accuracy is calculated as the ratio of correctly predicted areas in the model to the total dataset [9, 32].

After the implementation phase of the model was completely completed, the Confusion Matrix was used to visualize the distributions of the data and how many of these distributions were correctly predicted. Confusion Matrix is a tool that shows the morphs that are most likely to be confused with each other and improves performance prediction in ensemble detection tasks. It is also a technique used to measure the success rate of detection approaches [1 ,38]. This method was applied to visually demonstrate the success rate. As a result of all the calculations of these processes, it is aimed to see what the advantages and limitations of the model are.

Chi-Square test was applied to reach a conclusion on whether there is a statistical relationship or similarity between the magnitudes and depths of earthquakes in Izmir region. Chi-square test is a statistical analysis method used to examine the relationship between categorical variables. Based on the research hypothesis, it determines whether there is a difference in the proportions of risk factors between groups by evaluating the differences between rows and columns [14, 28, 35]. The relationship between depth and magnitude was analyzed and visualized with the applied test.

Geographic Information Systems (GIS) is software that models geospatial environments, enhances analysis functions, and provides human-centered geographic information for better understanding and communication [23]. GIS supports a wide range of spatial queries and plays an important role in future location model development and application in fields such as Geography, Civil Engineering, and Computer Science [10]. The data obtained as a result of queries can be presented in visualization methods such as points, lines or areas. The data is effectively integrated, managed and analyzed with geographic information from maps, images and text. These features provide a powerful tool for solving geospatial problems [11 ,20]. Since the data obtained in the study were in the form of point data, a general risk analysis map was created using these data. An earthquake map of the relevant study area was prepared, and thanks to this map, the distribution of earthquakes in and around Izmir province and where the magnitude ranges are more intense were visualized.

The answers to the questions asked in this study were obtained with Excel 2016, Google Colab and ArcMap 10.8 software. The Accuracy, Precision and F1-Score of the model, the results of the model were evaluated by examining the training and test datasets with the aim of successfully evaluating the prediction performance by processing the data within the framework of certain rules. GIS and data visualization techniques have made significant contributions to the analysis process of risky areas. The statistical relationship between the magnitudes, depths and occurrence regions of earthquakes in the Izmir region was visualized in tables.

3. Findings

Under this heading, the research questions given in the introduction of the study have been comprehensively answered. In order to make the findings obtained throughout the research process more understandable, they were visualized with different graphical methods and detailed using various tables. Thus, it is ensured that the results of the research are presented more clearly and clearly both quantitatively and qualitatively.

MAIN QUESTION: When the past earthquake data are analyzed, can the characteristics and distribution of earthquakes that may occur in the future in Izmir be predicted?

Yes, the data is predictable. As a result of the statistical analysis of the data, the classification success report of the model (Table 1) was created. Accordingly, the probabilities of predicting the magnitudes of earthquakes are objectified with Precision, Recall and F1-Score values.

SUBQUESTION 1: How accurately can future earthquakes be predicted with the decision tree model?

Table 1 shows that while high precision, recall and f1-score values were obtained especially for small earthquakes, these metrics decreased significantly for large earthquakes. The imbalance of the data, especially the limited number of high magnitude earthquakes, was found to negatively affect the success of the model in predicting these earthquakes.

Table 1. *Classification Report Table*

Class	Magnitud (xM)	Precision	Recall	F1-Score	Support
0	3.0 – 3.9	0.93	1.00	0.96	3014
1	4.0 – 4.9	0.24	0.03	0.05	189
2	5.0 – 5.9	0.14	0.03	0.05	35
3	6.0 and upper	0.00	0.00	0.00	3

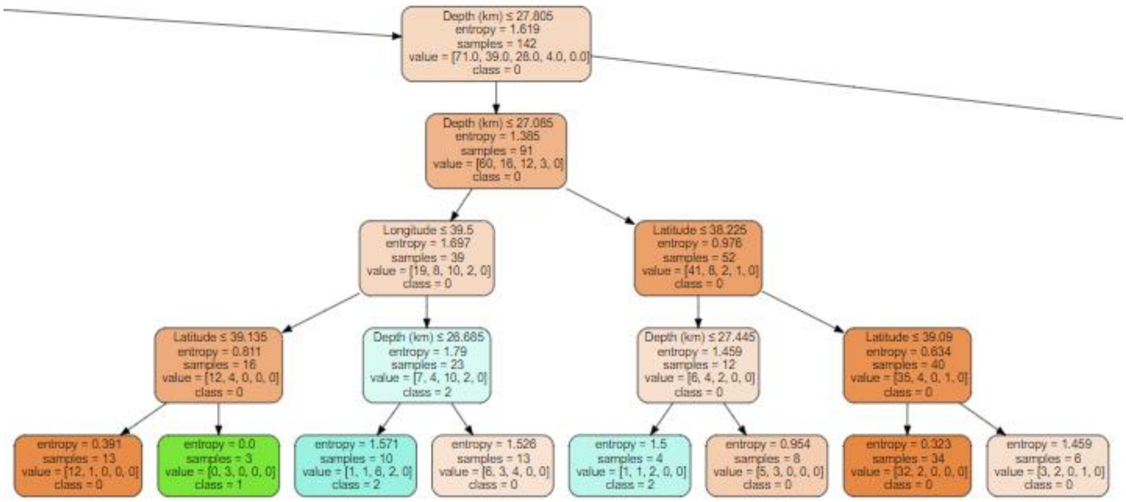

Figure 3. *Decision Trees' (DTs) One Tree Example*

The decision tree model in Figure 3 has one branch. In the output of this model, the answers given to the this question are given in Table 2 and Figure 3. In the table, it is determined at which level earthquake magnitudes can be trained and at which level they can be estimated. In Figure 3, the data at the root node is divided into two based on the 'Depth' feature according to the limit of 27.805 km. The entropy value at this node is calculated as 1.619, which indicates that the dataset is quite diverse and complex. The value of "Samples" is 142 and there are a total of 142 data samples in this node. "Value" is [71, 39, 28, 4, 0], which means that 71 of the 142 samples belong to class 0, 39 to class 2 and 28 to class 3. The "Class" value is assigned as 0, indicating that class 0 has the most instances. Each "Value" value describes in detail to which classes the instances in the node belong. In this context, the decision tree performs the final classification by decomposing the data step by step. At leaf nodes, the entropy value approaches zero or becomes zero, indicating that the datasets have become completely homogeneous and there are only examples belonging to a single class.

Table 2. *Model Performance Metrics*

1. Training Performance For Regression Tree	1. Test Performance For Regression Tree
0.93	0.92
Accuracy	**Weighted Avg.**
0.92	0.90

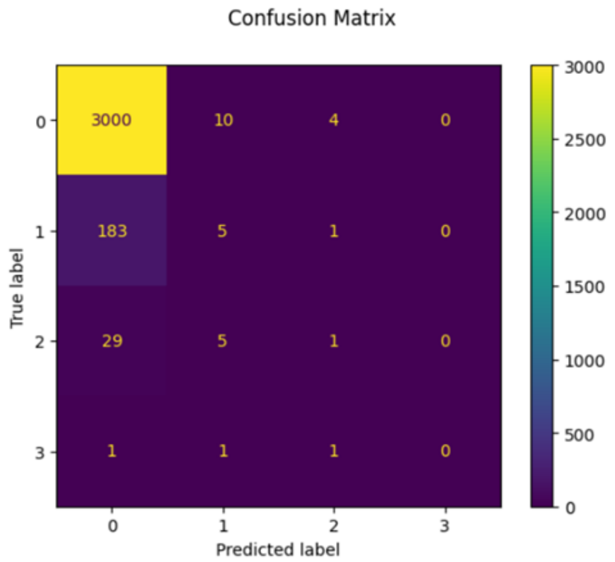

Figure 4. *Confusion Matrix*

SUBQUESTION 2: What are the advantages and limitations of the decision tree model?

The results obtained from the model (Figure 4) show that the model is inadequate in predicting earthquakes with larger magnitudes. In fact, there are only 35 cases of earthquakes with magnitude 5.0 and only 3 cases of earthquakes with magnitude 6.0 and above. The small data set limited the model's ability to predict earthquakes of this magnitude. In such cases, the model was unable to accurately predict rare events (such as large earthquakes). However, when the overall model performance is analyzed in Table 3, the model provided a good prediction result of 92%.

Table 3. *Model Performance Table*

Accuracy	Precision	Recall	F1-Score
0.92	0.88	0.92	0.89

SUBQUESTION 3: Is there a statistical relationship or similarity between the magnitudes and depths of earthquakes in Izmir province?

In order to visually see the relationship between depth and magnitude, the table in Figure 5 was created. In the relevant table, a high density is observed especially in the 0-20 km depth range. It was revealed that many earthquakes occurred at shallow depths (close to the surface). In terms of magnitude, it is generally observed that values between 3-5 are more frequently observed, but in earthquakes deeper than 40 km, the magnitude shows a wider distribution and can reach values of 5 and above.

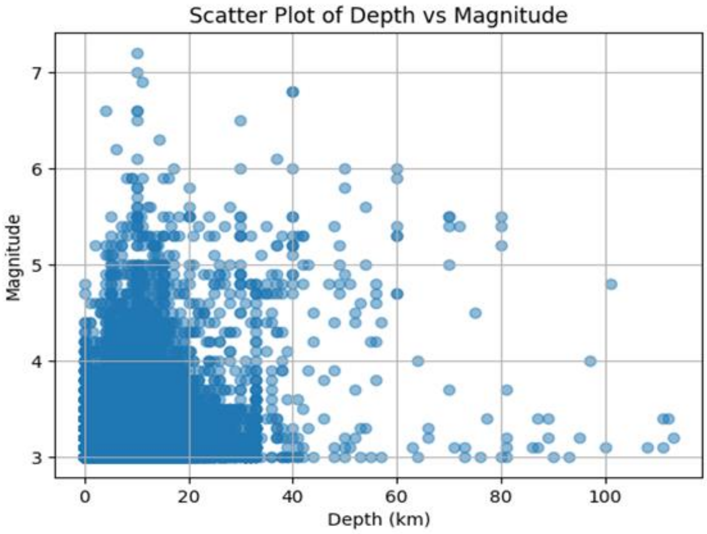

Figure 5. *Scatter Plot of Depth vs Magnitude (xM)*

As a result of the chi-square test performed to reveal the statistical relationship between depth and magnitude, the χ^2 (chi-square) value was 3924.06 and the p-value was 0.0. This result (Table 4) shows that there is a statistically significant relationship between depth and magnitude categories. The p-value is well below 0.05 (in this case very close to 0), indicating that there is a significant non-random relationship between depth and magnitude.

Table 4. *Chi-Square Result*

Chi-Square Test	P-Score
3924.06	0.0

SUBQUESTION 4: What is the distribution of earthquakes that occurred in Izmir province in the last century and in which areas are these earthquakes concentrated?

The distribution and concentration areas of earthquakes of 5.0 and above occurring in Izmir province were visualized using ArcGIS software. As a result of this visualization, an intensity map (Figure 6) showing the distribution of earthquakes and their hectic areas on specific fault segments was obtained.

Izmir province is a region with a high earthquake risk due to active fault lines on both land and seafloor (Figure 6). Frequent earthquakes, especially along active fault lines on the seafloor (north of Lesvos and south of Lesvos and north of Samos), pose a great danger to Izmir and the surrounding settlements. On the map, these earthquakes are particularly concentrated in coastal areas and the faults on the seafloor, which are very close to Izmir, significantly increase the seismic risk of the city. Izmir's proximity to the seashore also brings up secondary hazards such as earthquake-induced sea flooding. Therefore, coastal areas stand out as the riskiest areas due to their proximity to faults in the sea. Looking at the earthquakes that have occurred in Izmir, it is observed that earthquakes with a magnitude of 5.1 and above occur more frequently on the MTF (Mitilini Fault), IZF (Izmir Fault), GLFZ (Gülbahçe Fault Zone) and SMF (Samos Fault Zone) compared to other faults.

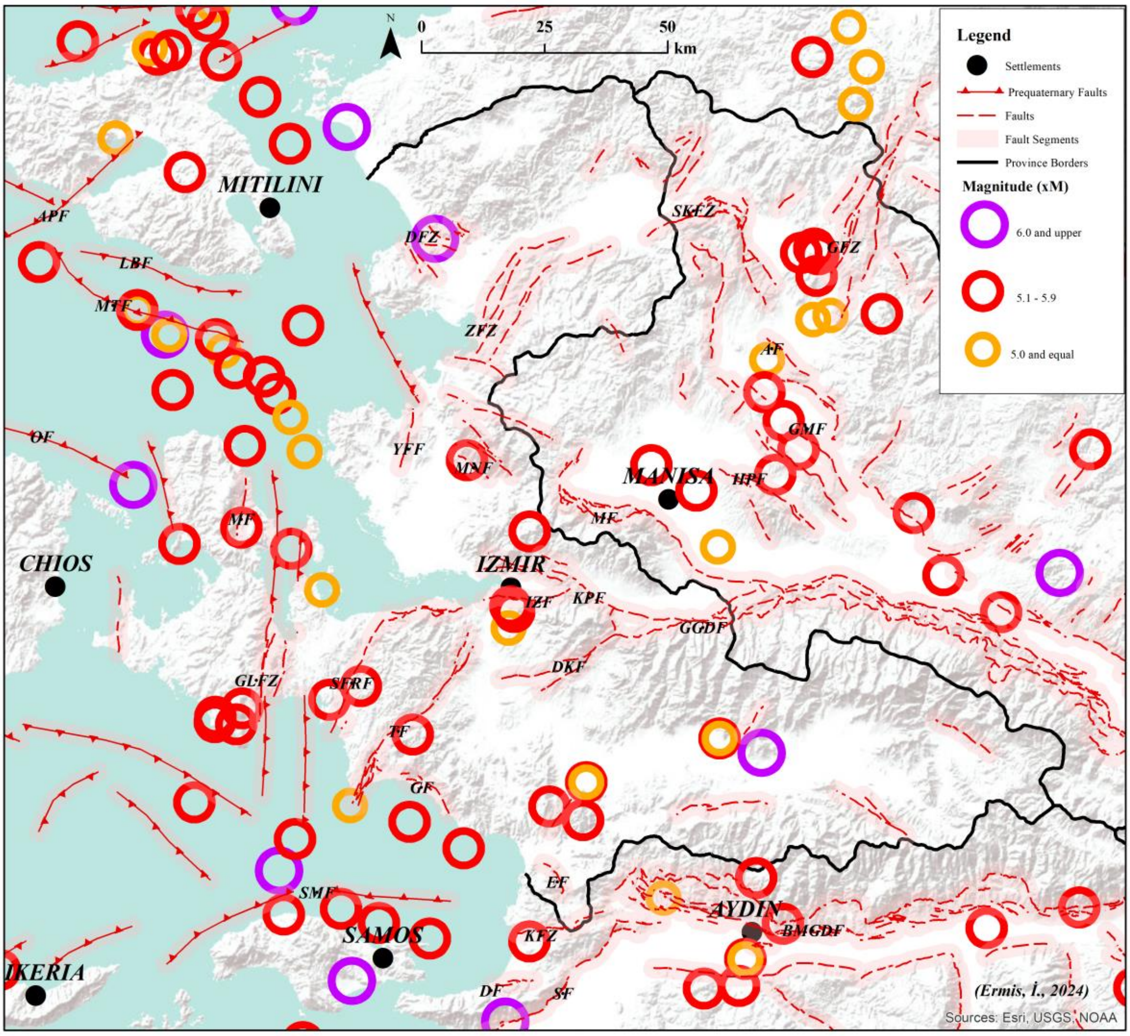

Figure 6. *Distribution Map of Past Earthquakes (< 5.0) and Main Fault Zones: Information on fault lines was taken from the website of the General Directorate of MTA (https://www.mta.gov.tr/) and the faults on the seabed were taken from the EMODNet website (https://emodnet.ec.europa.eu/). While preparing the map, only earthquakes with magnitude 5.0 and above were selected because too much data causes image pollution.*

The faults in Figure 6 and their descriptions are as follows: APF (Aghia Paraskevi Fault); LBF (Lesvos Basin Fault); MTF (Mitilini Fault); OF (Oinousses Fault); MF (Manisa Fault); GLFZ (Gülbahçe Fault Zone); SFRF (Seferihisar Fault); TF (Tuzla Fault); GF(Gümüldür Fault); SMF(Samos Fault); DF(Davutlar Fault); SF(Söke Fault); KFZ(Kuşadası Fault Zone); EF(Ephesus Fault); BMGDF(Büyük Menderes Graben Detachment Fault); DKF(Dagkızılca Fault); IZF(Izmir Fault); KPF(Kemalpasa Fault); GGDF(Gediz Graben Detachment Fault); MNF(Menemen Fault Zone); YFF(Yenifoça Fault); ZFZ(Zeytindag Fault Zone); DFZ(Dikili Fault Zone); SKFZ(Soma-Kırkağaç Fault Zone); GFZ(Gelenbe Fault Zone); AF(Akhisar Fault); GMF(Gölmarmara Fault); HPF(Halitpaşa Fault).

4. Conclusions and Suggestions

The findings of the study clearly show that Izmir is a region with high seismic risk. The analyses revealed that although we can successfully predict small earthquakes, the prediction accuracy for large earthquakes is limited. This is due to the lack of data on large earthquakes.

Forecasts using the decision tree model showed that the overall accuracy of the model was as high as 92%. However, it was observed that the model failed especially for large earthquakes. The limited data set for large earthquakes negatively affected the performance of the model. Especially for earthquakes of 5.0 and above, the precision, recall and f1-score values were almost zero. This shows that the model is inadequate in predicting such large events. According to Somodevilla et al., 2012 in the related literature review, it was mentioned that it is indeed possible to determine the seismic depth based on its magnitude.

According to the results of this study, the relationship between depth and magnitude was found to be statistically significant in the decision tree model. According to Asim et al., 2017, the Random Forest algorithm, which is included in the decision tree algorithm, gives 77% accuracy, while the decision trees algorithm, which is a detailed study of RFA in this study, gives 92% accuracy rate. However, according to Mignan, A., & Broccardo, M. (2020), the argument that it may not be fine-tuned when there is a basic machine learning classifier (Support Vector Machine, Decision Trees, Naïve Bayes, etc.) in this study due to the uneven distribution of the data confirmed this view. However, according to Mignan, A., & Broccardo, M. (2020), when it is a basic machine learning classifier (support vector machine, decision trees, naive Bayes, etc.), it may not be fine-tuned. According to Ridzwan and Yusoff, 2023, decision tree is the most accurate predictor algorithm compared to other machine learning models, whereas in this study, although decision tree gave very high rates, it was an algorithm that suffered from overfitting and difficulty in classifying small numbers of data.

When the relationship between depth and magnitude is analyzed, it is revealed with a statistically significance result that as the depth increases, the magnitude of the earthquakes also increases. The results of the chi-square test show that there is a strong non-random relationship between depth and magnitude. In particular, earthquakes occurring at shallow depths (0-20 km) are generally smaller, while deeper earthquakes have larger magnitudes.

When the earthquake risks in the region are analyzed, it is seen that İzmir province and its surroundings have a high risk especially due to the active fault lines located on the coast. According to Polat et al., 2009, in addition to the study that the earthquakes experienced in Izmir are generally around the Gulf of Izmir and that these areas are in danger, the study conducted by considering the faults located in the sea has contributed to the study on this subject. The study conducted by Tepe et al., 2021 provided valuable information on the intensity distribution of past earthquakes. In the study, it was emphasized that the earthquakes occurring on the Izmir Fault were destructive. With the revision of the study, it has been determined that the Gülbahçe Fault Zone (GLFZ), Mitilini Fault (MTF) and Samos Fault (SF) also have destructive effects. In addition to this contribution, the distribution of earthquakes between 1900 and 2024 has been revised and analyzed. In addition, the study maintains its importance since there is no study that has been conducted by applying a decision tree model in magnitude estimation based on the distribution in the context of Izmir province. Faults in the seafloor (north of Lesvos, south of Lesvos, north of Samos) pose a significant threat to the settlements in and around Izmir. They pose not only earthquake risk but also secondary hazards such as tsunamis for the population living in coastal areas.

In conclusion, the model needs to be strengthened by using more data to predict the magnitude of future earthquakes in Izmir. The lack of data, especially for large earthquakes, limits the performance of the prediction models. The findings of the study show that Izmir province is a region that should be carefully monitored in terms of seismic risk and emphasize that earthquake measures should focus on coastal areas.

References

[1] D. Abdullah and E. D. Putra, "Comparasi edge detection roberts dan morfologi pada deteksi plat nomor kendaraan roda dua," J. Sci. Appl. Informatics, vol. 1, no. 3, pp. 66–69, 2018.

[2] S. Ahamed and E. G. Daub, "Machine learning approach to earthquake rupture dynamics," arXiv preprint arXiv:1906.06250, 2019.

[3] F. Ahmed et al., "Earthquake magnitude prediction using machine learning techniques," in 2024 IEEE Int. Conf. Interdiscip. Approaches Technol. Manag. Soc. Innov. (IATMSI), 2024, vol. 2, pp. 1–5.

[4] A. Ardakani and V. Kohestani, "Evaluation of liquefaction potential based on CPT results using C4.5 decision tree," J. AI Data Mining, vol. 3, pp. 85–92, 2015.

[5] K. M. Asim, A. Idris, F. Martínez-Álvarez, and T. Iqbal, "Short term earthquake prediction in Hindukush region using tree based ensemble learning," in 2016 Int. Conf. Front. Inf. Technol. (FIT), 2016, pp. 365–370.

[6] K. M. Asim, F. Martínez-Álvarez, A. Basit, and T. Iqbal, "Earthquake magnitude prediction in Hindukush region using machine learning techniques," Nat. Hazards, vol. 85, pp. 471–486, 2017.

[7] Y. Cai, M. L. Shyu, Y. X. Tu, Y. T. Teng, and X. X. Hu, "Anomaly detection of earthquake precursor data using long short-term memory networks," Appl. Geophys., vol. 16, pp. 257–266, 2019.

[8] M. Cassel, "Machine learning and the construction of a seismic attribute-seismic facies analysis data base," M.S. thesis, Univ. Oklahoma, 2018. [Online]. Available: https://shareok.org/

[9] R. Chen et al., "Rigorous assessment and integration of the sequence and structure based features to predict hot spots," BMC Bioinformatics, vol. 12, pp. 1–14, 2011.

[10] R. L. Church, "Geographical information systems and location science," Comput. Oper. Res., vol. 29, no. 6, pp. 541–562, 2002.

[11] İ. Cürebal and E. Özşahin, Harita Bilgisi (Bilgisayar Uygulamalı Tasarım ve Analiz). Bursa, Türkiye: Ekin Basın Yayın Dağıtım, 2022.

[12] K. Demertzis, K. Kostinakis, K. Morfidis, and L. Iliadis, "A comparative evaluation of machine learning algorithms for the prediction of R/C buildings' seismic damage," arXiv preprint arXiv:2203.13449, 2022.

[13] A. Doğru, E. Görgün, H. Ozener, and B. Aktuğ, "Geodetic and seismological investigation of crustal deformation near Izmir (Western Anatolia)," J. Asian Earth Sci., vol. 82, pp. 21–31, 2014.

[14] T. M. Franke, T. Ho, and C. A. Christie, "The chi-square test: Often used and more often misinterpreted," Am. J. Eval., vol. 33, no. 3, pp. 448–458, 2012.
[15] E. Gok and O. Polat, "An assessment of the microseismic activity and focal mechanisms of the Izmir (Smyrna) area from a new local network (IzmirNET)," Tectonophysics, vol. 635, pp. 154–164, 2014.
[16] S. Goswami, S. Chakraborty, S. Ghosh, A. Chakrabarti, and B. Chakraborty, "A review on application of data mining techniques to combat natural disasters," Ain Shams Eng. J., vol. 9, no. 3, pp. 365–378, 2018.
[17] J. Han, J. Kim, S. Park, S. Son, and M. Ryu, "Seismic vulnerability assessment and mapping of Gyeongju, South Korea using frequency ratio, decision tree, and random forest," Sustainability, vol. 12, no. 18, p. 7787, 2020.
[18] C. Jiang, X. Wei, X. Cui, and D. You, "Application of support vector machine to synthetic earthquake prediction," Earthq. Sci., vol. 22, pp. 315–320, 2009.
[19] A. Karbassi, B. Mohebi, S. Rezaee, and P. Lestuzzi, "Damage prediction for regular reinforced concrete buildings using the decision tree algorithm," Comput. Struct., vol. 130, pp. 46–56, 2014.
[20] G. Q. King, "Geography and GIS technology," J. Geogr., vol. 90, no. 2, pp. 66–72, 1991.
[21] S. B. Kotsiantis, "Decision trees: A recent overview," Artif. Intell. Rev., vol. 39, pp. 261–283, 2013.
[22] A. Li and L. Kang, "Knn-based modeling and its application in aftershock prediction," in 2009 Int. Asia Symp. Intell. Interact. Affect. Comput., 2009, pp. 83–86.
[23] G. Lü et al., "Reflections and speculations on the progress in geographic information systems (GIS): A geographic perspective," Int. J. Geogr. Inf. Sci., vol. 33, no. 2, pp. 346–367, 2019.
[24] A. Mignan and M. Broccardo, "Neural network applications in earthquake prediction (1994–2019): Meta-analytic and statistical insights on their limitations," Seismol. Res. Lett., vol. 91, no. 4, pp. 2330–2342, 2020.
[25] O. Onat and H. Tanyıldızı, "Machine learning-based estimation of the out-of-plane displacement of brick infill exposed to earthquake shaking," Eng. Appl. Artif. Intell., vol. 136, p. 109007, 2024.
[26] G. V. Otari and R. V. Kulkarni, "A review of application of data mining in earthquake prediction," Int. J. Comput. Sci. Inf. Technol., vol. 3, no. 2, pp. 3570–3574, 2012.
[27] O. Polat et al., "IzmirNet: A strong-motion network in metropolitan Izmir, Western Anatolia, Türkiye" Seismol. Res. Lett., vol. 80, pp. 831–838, 2009.
[28] R. Rana and R. Singhal, "Chi-square test and its application in hypothesis testing," J. Primary Care Specialties, vol. 1, no. 1, pp. 69–71, 2015.
[29] N. S. M. Ridzwan and S. H. M. Yusoff, "Machine learning for earthquake prediction: A review (2017–2021)," Earth Sci. Inform., vol. 16, no. 2, pp. 1133–1149, 2023.
[30] M. Senkaya, A. Silahtar, E. F. Erkan, and H. Karaaslan, "Prediction of local site influence on seismic vulnerability using machine learning: A study of the 6 February 2023 Türkiye earthquakes," Eng. Geol., p. 107605, 2024.
[31] I. Sikder and T. Munakata, "Application of rough set and decision tree for characterization of premonitory factors of low seismic activity," Expert Syst. Appl., vol. 36, pp. 102–110, 2009.
[32] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation," in Australas. Joint Conf. Artif. Intell., pp. 1015–1021, 2006.
[33] M. J. Somodevilla, A. B. Priego, E. Castillo, I. H. Pineda, D. Vilariño, and A. Nava, "Decision support system for seismic risks," *J. Comput. Sci. Technol.*, vol. 12, 2012.
[34] Ç. Tepe et al., "Updated historical earthquake catalog of İzmir region (western Anatolia) and its importance for the determination of seismogenic source," Turk. J. Earth Sci., vol. 30, no. 8, pp. 779–805, 2021.
[35] N. S. Turhan, "Karl Pearson's Chi-Square Tests," Educ. Res. Rev., vol. 16, no. 9, pp. 575–580, 2020.
[36] P. Xiong et al., "Towards advancing the earthquake forecasting by machine learning of satellite data," Sci. Total Environ., vol. 771, p. 145256, 2021.
[37] C. E. Yavas, L. Chen, C. Kadlec, and Y. Ji, "Predictive modeling of earthquakes in Los Angeles with machine learning and neural networks," IEEE Access, 2024.
[38] E. ZeeAbrahamsen and J. Haberman, "Correcting 'confusability regions' in face morphs," Behav. Res. Methods, vol. 50, pp. 1686–1693, 2018.
[39] J. Zhang, "Dive into decision trees and forests: A theoretical demonstration," arXiv preprint arXiv:2101.08656, 2021.