Karaelmas Science and Engineering Journal
Journal home page: https://dergipark.org.tr/tr/pub/karaelmasfen
DOI: 10.7212/karaelmasfen.1494349

Research Article
Received / Geliş tarihi : 02.06.2024
Accepted / Kabul tarihi : 09.08.2024

Forecasting Earthquake Impact Scenarios in Istanbul with Machine Learning Algorithms

İstanbul'daki Deprem Etki Senaryolarının Makine Öğrenmesi Algoritmaları ile Tahmin Edilmesi

Remzi Gürfidan[1], Mehmet Ali Yalçınkaya[2*]

[1]Isparta University of Applied Sciences, Yalvaç Technical Sciences Vocational School, Department of Computer Technologies, Isparta, Türkiye

[2]Kırşehir Ahi Evran University, Faculty of Engineering and Architecture, Department of Computer Engineering, Kırşehir, Türkiye

Abstract

This study employs machine learning algorithms to forecast the impacts of a potential magnitude 7.5 earthquake in Istanbul, focusing on casualty rates, hospitalization needs, and temporary shelter requirements. Using a dataset compiled from the Istanbul Metropolitan Municipality Open Data Portal and the Turkish Statistical Institute, the research assesses Gradient Boosting, AdaBoost, Random Forest, and ExtraTrees algorithms. Gradient Boosting emerged as the most effective model, exhibiting high accuracy and low prediction errors in determining disaster impacts. This approach underscores the critical role of advanced analytics in enhancing urban disaster preparedness and management, providing valuable insights for policymaking and infrastructure development in earthquake-prone areas.

Keywords: Earthquake impact forecasting, disaster preparedness, machine learning, urban risk management.

Öz

Bu çalışmada, İstanbul'da olası bir 7.5 büyüklüğündeki depremin etkilerini, özellikle de can kaybı sayısı, hastaneye ihtiyaç duyacak kişi sayısı ve geçici barınma ihtiyacı duyacak kişi sayısını tahmin etmek için makine öğrenmesi algoritmaları kullanılmaktadır. İstanbul Büyükşehir Belediyesi Açık Veri Portalı ve Türkiye İstatistik Kurumu'ndan derlenen bir veri seti kullanılarak Gradyan Artırma (Gradient Boosting), Uyarlanabilir Artırma (AdaBoost), Rastgele Orman (Random Forest) ve Ekstra Ağaçlar (ExtraTrees) algoritmaları değerlendirilmiştir. Gradient Boosting modeli, yüksek doğruluk ve düşük tahmin hataları ile en etkili model olarak öne çıkmıştır. Bu yaklaşım, gelişmiş analitiklerin kentsel afet hazırlığı ve yönetimini geliştirme konusundaki kritik rolünü vurgulamakta ve depreme eğilimli bölgelerdeki alınacak önlemler ve altyapı gelişimi için değerli öngörüler sağlamaktadır.

Anahtar Kelimeler: Deprem etki tahmini, afet hazırlığı, makine öğrenmesi, kentsel risk yönetimi.

1. Introduction

Earthquakes rank among the most devastating natural disasters globally, causing significant destruction and loss of life. Turkey, due to its geological location, is situated in one of the most active earthquake zones in the world. The North Anatolian Fault Line (NAF) runs through the northern part of the country, causing significant seismic activity. Historically, this fault line has been responsible for most of the devastating earthquakes in Turkey. Notably, the major tremor on August 17, 1999, known as the Marmara Earthquake, occurred along this fault line. This earthquake caused extensive destruction, particularly in Kocaeli and its surroundings, leading to over 17,000 deaths and injuring tens of thousands, according to official reports (Erdik 2001). The scale of the destruction was recorded as one of the most severe disasters in Turkey's modern history.

The economic impact of the Marmara Earthquake was also severe. The direct damage and subsequent economic losses amounted to over 25 billion dollars, constituting approximately 2-3% of Turkey's gross domestic product (GDP) (Ambraseys 2001). Workplaces were destroyed,

*Corresponding author: mehmetyalcinkaya@ahievran.edu.tr

Remzi Gürfidan orcid.org/0000-0002-4899-2219
Mehmet Ali Yalçınkaya orcid.org/0000-0002-7320-5643

infrastructure was damaged, and many industrial facilities had to cease operations. This situation created long-term negative effects on both local and national economies. Moreover, with thousands left homeless, the state faced significant responsibilities for housing and reconstruction. The economic recovery took years and was significantly supported by international aid. This heavy burden led to major changes in Turkey's urban transformation and disaster management policies, promoting the development of stronger building standards and effective disaster management strategies (Stein 2000).

The earthquakes centered in Kahramanmaraş last year, particularly the main shock of magnitude 7.8 on February 6, 2023, caused serious destruction in Southern and Central Turkey. This disaster resulted in the loss of thousands of lives and left hundreds of thousands homeless. The provinces of Hatay, Kahramanmaraş, Gaziantep, Malatya, and Adıyaman saw the most severe damage, accounting for 81% of the estimated total damage. This calamity is recorded as the largest natural disaster Turkey has faced in the last 80 years. Furthermore, these provinces house approximately 7.4% of Turkey's population, significantly impacting the region's socio-economic structure (World Bank 2023). Economically, the cost of the earthquake to Turkey is estimated at 34.2 billion dollars in direct physical damages, equating to about 4% of the country's GDP in 2021 (Korkut et al. 2023).

Minimizing losses caused by earthquakes is possible through preventative measures and preparations. Earthquake scenario analysis allows for the anticipation of potential earthquake effects and the planning of necessary precautions based on this information. Particularly in major urban areas, the preparation of such scenarios can significantly reduce the loss of life and property (Jaiswal and Wald 2010).

Istanbul, as Turkey's largest metropolis, is distinguished by its dense population and cosmopolitan structure. The city harbors a serious risk of a major earthquake due to sporadic urbanization and a high proportion of structurally unsound buildings. The potential devastation of an earthquake in Istanbul is feared due to the concentration of the nation's population and industrial density in this region (Parsons 2000).

The literature contains various studies on estimating potential fatalities and other damages prior to earthquakes. In China, data from 84 earthquakes occurring between 1970 and 2017 were tested using methods such as elaboration likelihood model (ELM), artificial neural networks, support vector machine (SVM), and Gaussian curves. The study achieved an r squared (R^2) value of 96% (Xing et al. 2020). Another study analyzed 30 historical earthquakes in China, comparing the performance of the support vector regression (SVR) model with other machine learning (ML) methods. The comparisons concluded that the SVR model produced more successful outcomes than other models (Li et al. 2021). Additionally, another study attempted to estimate the number of potential fatalities in a severe earthquake using logistic regression on a dataset containing four attributes (age, gender, physical disability, and socioeconomic status). In one of the studies on earthquakes in Turkey, data collected from buildings damaged after three major earthquakes were used to predict structural damage and mitigate potential harms. This process demonstrated that the SVR model yielded better results than other models (Arslan et al. 2017). Corbi et al. (2019) discuss how machine learning can predict the timing and size of earthquakes by reconstructing complex system dynamics in subduction zones. Rouet-Leduc et al. (2017) show how machine learning can predict laboratory earthquakes based on acoustic signals previously thought to be noise, potentially improving earthquake forecasting. Ahamed and Daub (2019) present a machine learning model for predicting whether an earthquake rupture will propagate, using neural networks and random forest algorithms.

This study predicted the number of loss of lifes, the number of people needing hospital treatment, and the number of people requiring temporary shelter that a potential 7.5 magnitude earthquake could cause during night hours in Istanbul. The data used for this analysis were obtained from the Istanbul Metropolitan Municipality (IMM) Open Data Portal and the Turkish Statistical Institute (TUIK). Initially, the "Earthquake Scenario Analysis Results" and "Neighborhood-Based Building Counts" datasets from the Istanbul Metropolitan Municipality Open Data Portal were merged. Subsequently, population and area data for each neighborhood in Istanbul were obtained from TUIK and added to the dataset. This merger produced a comprehensive and original dataset. Various ML techniques and algorithms were used on the collected dataset to perform detailed analyses. These analyses are critical for better understanding the potential effects of earthquakes and planning necessary precautions in advance.

2. Material and Methods

This section describes the dataset used in the study, the algorithms applied, and the metrics employed to evaluate the data derived from these algorithms.

2.1. Dataset

In this study, a hybrid dataset was constructed. The attributes included in this dataset were obtained from various sources. There are 12 different attributes within this dataset. These attributes and their value ranges are shown in Table 1. The first attribute in the dataset, "Population Density," was calculated using the population and area information for each neighborhood in Istanbul, as published by the TUIK. The attributes related to the total number of buildings, the rate of buildings built before 1980, the rate for buildings constructed between 1980-2000, the rate for buildings constructed after 2000, and the rates for building intervals of 1-4 floors, 5-9 floors, and 9-19 floors were sourced from the "Neighborhood-Based Building Counts" dataset created by the IMM (IMM Open Data Portal 2017). The last four attributes (Number of Too Severely Damaged Buildings Rate, Number of Severely Damaged Buildings Rate, Number of Moderately Damaged Buildings Rate, Number of Lightly Damaged Buildings Rate) along with three dependent variables used for model predictions (Number of Casualties, Number of People in Need of Shelter, Number of People in Need of Hospital Treatment) were also derived from the "Earthquake Scenario Analysis Results" dataset produced by the IMM (IMM Open Data Portal 2021). This dataset contains a simulation of the potential outcomes of a 7.5 magnitude earthquake occurring at night in Istanbul. The rationale behind our study's focus on predicting the effects of a nighttime earthquake is precisely this. Following these processes, a unique dataset containing data from three different sources was established. For the hybrid dataset obtained by combining different datasets, the input values are similar for all datasets. Instead of forcing the algorithms to predict multiple outcomes, they are designed to predict each outcome independently. The absence of null values and outliers in the dataset eliminated the need for cleaning the dataset before the process. Techniques such as Principal Component Analysis (PCA) were not used for feature selection. The dataset used does not contain a large number of input parameters that would need to be eliminated.

The dataset underwent several preprocessing steps to ensure its quality and suitability for the analysis. These steps included:

- Data Cleaning: Removal of any duplicate records and handling of missing values. Missing values were addressed using mean imputation for continuous variables and mode imputation for categorical variables.
- Normalization: Scaling of numerical features to a standard range using min-max normalization, ensuring that all features contribute equally to the model training process.
- Outlier Detection: Identification and treatment of outliers to prevent them from skewing the model results.

Table 1. The features within the dataset and their value ranges.

Variables Type	Features	Min – Max Range
Independent Variables	Population Density	0.10992953-726.5714286
	Total number of buildings	95-8118
	Before 1980 Rate	0-0.958970793
	1980-2000_Between Rate	0.01122449-0.930743243
	2000_After Rate	0-0.986734694
	1-4 Floor Interval Rate	0.045566502-0.997938144
	5-9 Floor Interval Rate	0.002061856-0.951970443
	9-19 Floor Interval Rate	0-0.844036697
	Number of Too Severely Damaged Buildings Rate	0-0.095384615
	Number of Severely Damaged Buildings Rate	0.002132196-0.143076923
	Number of Moderately Damaged Buildings Rate	0.026448363-0.327443401
	Number of Lightly Damaged Buildings Rate	0.099461049-0.508196721
Dependent Variables	Number of Casualties	0-0.01563
	Number of People in Need of Shelter	0-0.460440986
	Number of People in Need of Hospital Treatment	0-0.071428571

2.2. Models Used

2.2.1. Gradient Boosting (GB)

GB is a method that sequentially enhances weak learners (typically decision trees) to form a strong predictive model. Each new learner focuses on minimizing the errors made by its predecessors. This process continues by assigning increasing weight to successive learners to reduce error terms. Mathematically, the model F(x) is shown at Equation 1. In this equation, (x) is the prediction of the t-th learner and λ is the learning rate (Friedman 2001).

$$F_{t+1}(x) = F_t(x) + \lambda h_t \quad (1)$$

2.2.2. Adaptive Boosting (Adaboost)

AdaBoost combines a series of weak classifiers to create a strong classifier. In each learning iteration, higher weight is given to observations that were incorrectly classified, thus the next classifier focuses on better predicting these observations. The fundamental update equation of AdaBoost is shown at Equation 2.

$$D_{t+1}(i) = \frac{D_t(i).\exp(-\alpha_t y_i h_t(x_i))}{Z_t} \quad (2)$$

In Equation 2, $D_t(i)$ is the weight of the i-th sample, α_t is the weight of the t-th learner, y_i is the actual class label, $h_t(x_i)$ is the predicted class, and Z_t is the normalization factor (Freund and Schapire 1997).

2.2.3. Random Forest (RF)

RF operates by aggregating the predictions of multiple decision trees and presenting the most frequent outcome as the final prediction. Each tree is constructed independently using randomly selected subsets of data. This method reduces both variance and bias, generally yielding more stable and reliable predictions. Mathematically, the RF prediction is calculated as follows:

$$f(x) = \frac{1}{b}\sum_{b=1}^{B} T_b(x) \quad (3)$$

In this formula, $T_b(x)$ represents the prediction of the b-th tree at point x (Breiman 2001).

2.2.4. Extremely Randomized Trees (ExtraTrees)

ExtraTrees method is an ensemble learning technique akin to RF, but it incorporates greater randomization at each node during splits. In this model, the optimal splitting point at each node is determined based on features chosen at random. This approach can further reduce the model's variance while decreasing computational time. ExtraTrees is particularly effective for high-dimensional datasets as it minimizes unnecessary computations and reduces the risk of overfitting. Mathematically, the ExtraTrees model can be expressed as shown at Equation 4.

$$f(x) = \frac{1}{b}\sum_{b=1}^{B} T_b(x, \theta_b) \quad (4)$$

In Equation 4, θ_b denotes the random parameters used in the construction of the b-th tree (Geurts et al. 2006)

2.2.5. Evaluation of the Models

Error metrics used to evaluate the success of ML algorithms are used to measure how well the model performs. These metrics help to assess how well a model's predictions match the true values and the generalization ability of the model.

The symbols P, r, m, c, and d commonly used in performance evaluation metrics are key terms frequently employed in modeling studies. P refers to the predicted values generated by the model, while r represents each observation or data point. m signifies the predicted values produced by the model, and c denotes the actual or observed values. Finally, d typically reflects the difference between the predicted and observed values, which is used to calculate error or deviation. These symbols form the foundation for measuring model performance and evaluating the accuracy of predictions.

Mean absolute error (MAE) is a metric that shows how close the predicted values are to the true values. This metric is calculated by Equation 5 (Hammid et al. 2018, Mishra et al. 2017, AlOmar et al. 2020).

$$MAE = \frac{1}{n}\sum_{r=1}^{n} \left| P_d^{r,m} - P_d^{r,c} \right| \quad (5)$$

Root means square error (RMSE) was chosen to compare the prediction errors of different trained models. The closer the RMSE value is to 0, the better the predictive ability of the model in terms of its absolute deviation. The RMSE value is calculated by Equation 6 (Hammid et al. 2018, Mishra et al. 2017, AlOmar et al. 2020, Willmott and Matsuura 2005).

$$RMSE = \sqrt{\frac{1}{n}\sum_{r=1}^{n} (P_d^{r,m} - P_d^{r,c})^2} \quad (6)$$

The coefficient of determination (R^2) is used to estimate model efficiency and is calculated by Equation 7 (Hammid et al. 2018).

$$R^2 = 1 - \frac{\sum_{r=1}^{n} (P_d^{r,m} - P_d^{r,c})^2}{\sum_{r=1}^{n} (P_d^{r,m} - P_d^{-r,m})^2} \quad (7)$$

MSE either assesses the quality of an estimator. The MSE metric is calculated by Equation 8.

$$MSE = \frac{1}{n}\sum_{r=1}^{n}(P_i - P_i^{'})^2 \quad (8)$$

Accuracy is one of the fundamental metrics used to evaluate the performance of a classification model. It represents the ratio of correctly classified samples to the total number of samples. Mathematically, accuracy is defined as follows:

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Samples} \quad (9)$$

This formula shows the percentage of correct predictions made by the model. For example, a model that makes 90 correct predictions out of 100 total samples would have an accuracy of 90% (James et al. 2013).

3. Results

3.1. Model Fine-Tunning Process and Results

Different algorithms are used to model data sets and provide predictions in the fields of ML and data analytics. Hyperparameters must be adjusted to boost these algorithms' efficiency and produce more accurate results. This work involves fine-tuning well-known algorithms including GB, AdaBoost, RF, and ExtraTrees using the datasets used for training. The hyperparameters were fine-tuned to maximize the performance of every algorithm. The most successful hyperparameter values obtained are presented in Table 2.

3.2. Selection Criteria for Hyperparameter Settings

In this study, the optimal hyperparameters obtained because of the experiments on the specified data sets were tested to improve the performance of each algorithm and improve the accuracy of the model. The process of hyperparameter tuning involved extensive experimentation to identify the optimal settings for each ML algorithm. The following criteria were used to select the final hyperparameters:

- Performance Metrics: MAE, RMSE, and R^2 were the primary metrics used to evaluate model performance.
- Cross-Validation: A k-fold cross-validation (with k=5) approach was employed to ensure that the models were robust and generalizable. This method helps in mitigating overfitting by training and validating the model on different subsets of the data.
- Grid Search: A grid search strategy was used to explore a predefined set of hyperparameters systematically. The grid search covered a range of values for each hyperparameter to identify the combination that yielded the best performance.

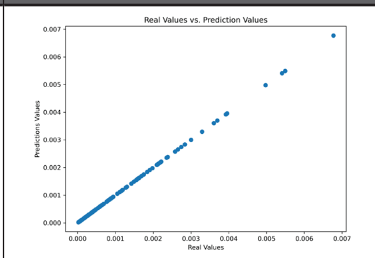
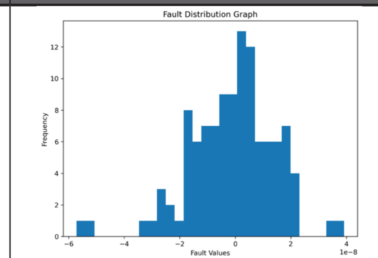
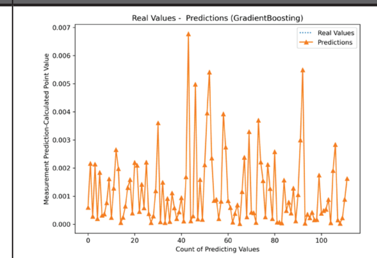
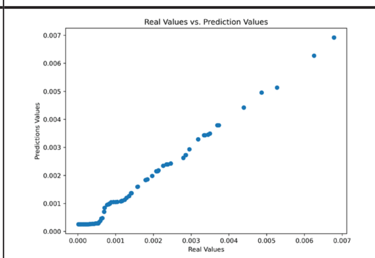
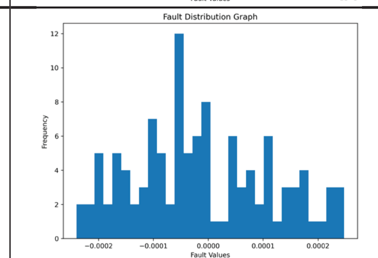
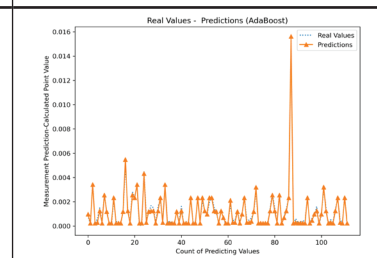
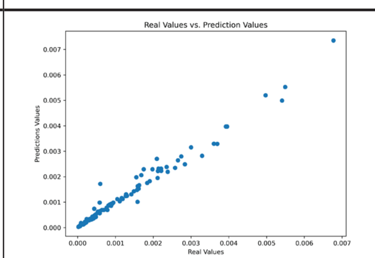
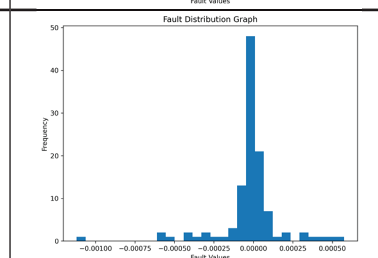
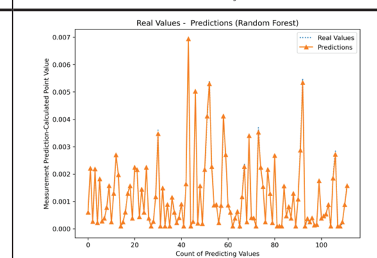
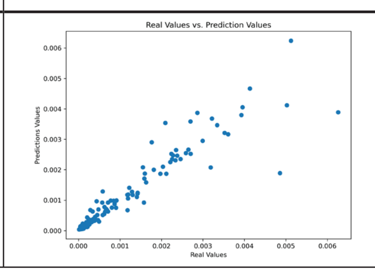
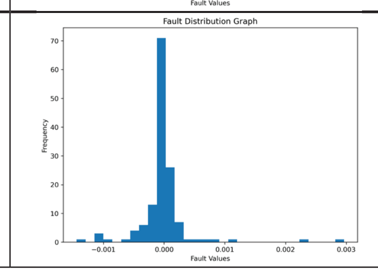
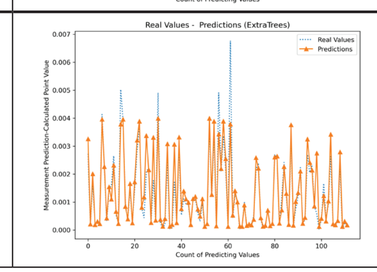
3.3. Trainings Conducted for Loss of Lifes Data and Results

Table 3 presents the training and test graphs, prediction error distributions, prediction and accuracy graphs, and validation graphs for the most and least successful

Table 2. The best values of the hyperparameter values of the ML models tried and obtained for the best result.

Algorithms	Trained Hypermeters	Best Hypermeters Values
GB	'n_estimators' : [50, 75, 100], 'max_depth': [2, 4, 8], 'min_samples_split' : [1,2,4], 'min_samples_leaf' : [1,2 ,8], 'learning_rate' : [0.05, 0.1 , 0.5, 1]	'learning_rate': 1, 'max_depth': 8, 'min_samples_leaf': 1, 'min_samples_split': 4, 'n_estimators': 50
ADABOOST	'n_estimators' : [30, 50, 75, 100,200], 'learning_rate' : [0.05, 0.1 , 0.5 , 1, 2]	'learning_rate': 0.1, 'n_estimators': 100
RF	'n_estimators' : [50,100,200], 'max_depth': [2, 4, 8, 16], 'min_samples_split' : [2,3,4], 'min_samples_leaf' : [2,3,4,8]	'max_depth': 16, 'min_samples_leaf': 2, 'min_samples_split': 3, 'n_estimators': 50'
EXTRATREES	'n_estimators': [50, 100, 200], 'max_depth': [None, 10, 20, 30], 'min_samples_split': [2, 5, 10], 'min_samples_leaf': [1, 2, 4]	'max_depth': 10, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100

Table 3. Graphs for estimating possible casualties.

ALG.	Real Values & Prediction	Fault Distribution	Predicting Values
GB			
ADABOOST			
RF			
EXTRATREES			

algorithms used in predicting the loss of life in the Istanbul earthquake scenario using ML. Table 3 shows the training and test graphs, prediction error distributions, prediction and accuracy graphs and validation graphs of the most successful and least successful algorithms of the models trained to predict the loss of life in the earthquake scenario for Istanbul with ML. In the training and test graph, a linear line graph is expected for successful training. In the test and training graphs of the ExtraTrees Regressor algorithm, there are some values that break the linear line. However, in the GB Regressor algorithm, it is seen that this graph forms a linear line graph. In the error distributions graph, it is expected to see distributions that do not differ as much as possible. In the ExtraTrees Regressor algorithm, the error distribution graph of the GB Regressor algorithm concentrates almost all of them at the 0 point with a tolerance of 1e-8, while the false predictions of 0.003 are seen at different frequencies. This clearly demonstrates the prediction success. In the prediction and accuracy graph, the blue lines show the actual values, and the orange lines show the values predicted by the model. What we want to see in this graph is the overlap of these two different colored lines. In the ExtraTrees Regressor algorithm, the blue and orange lines overlap, but not perfectly. On the other hand, in the GB Regressor algorithm, only orange colors are seen in the graph with two different colors. This proves the high accuracy of the prediction.

According to Table 4, the performance of the models trained using four different algorithms (GB, AdaBoost, RF and ExtraTrees) is evaluated. MSE, MAE, RMSE, R^2 and Accuracy metrics were commonly used. The GB model performed the best, achieving very low error rates in MSE, MAE and RMSE values. The r^2 value of 0.99 indicates a very high explanation rate. Moreover, the accuracy rate is also very high at 99.99%. Although the AdaBoost model has slightly higher error rates in the other metrics, it still has a very high R^2 value and an acceptable accuracy rate. The RF and ExtraTrees models have higher error rates and lower R^2 values. These models were less successful compared to GB and AdaBoost.

3.4. Trainings Conducted for the Number of Shelter Needs and Results

As can be seen from Table 5, the GB Regressor algorithm is quite successful in the Real Values and Prediction graph and in the predicting graph. In addition, although it looks like a scattered column graph in the Error distribution graph, the tolerance multiplier value is quite low compared to other algorithms.

Table 4. Metrics for estimating possible casualties.

Algorithm	MSE	MAE	RMSE	R^2	Accuracy
GB	2.366194013919057e-16	1.1559507427903762e-08	1.538243808347382e-08	0.99	99.99
AdaBoost	1.561541483737388e-08	0.0001048052560926842	0.00012496165346766937	0.99	86.54
RF	3.99632097624034e-08	0.00010438015990870665	0.00019990800324750234	0.97	86.41
ExtraTrees	1.9283873995273218e-07	0.0001937100020273874	0.0004391340796985952	0.87	71.82

Table 5. Graphs for estimating the number of people who will potentially need shelter.

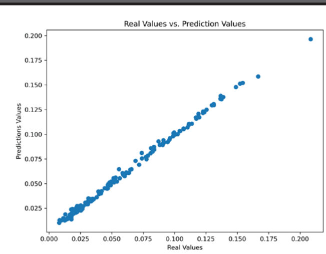
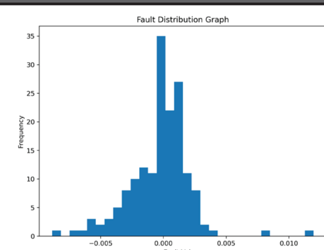
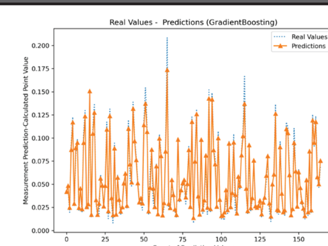
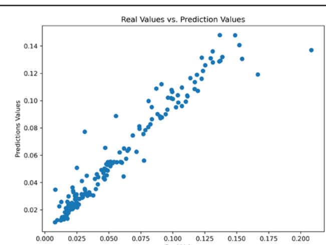
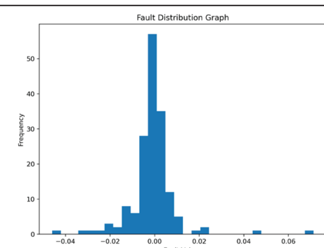
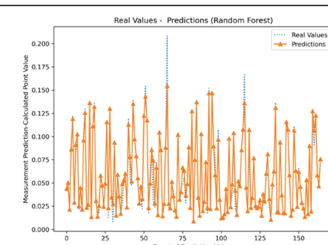
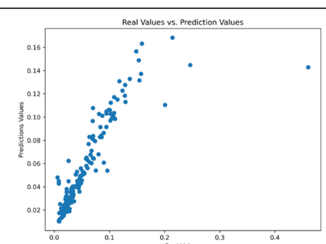
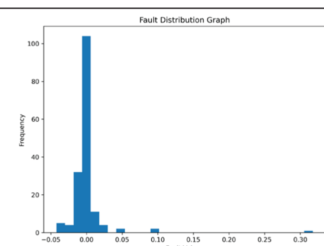
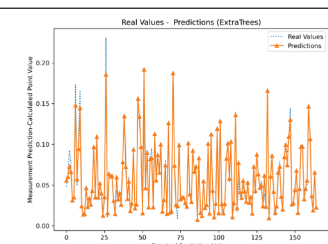
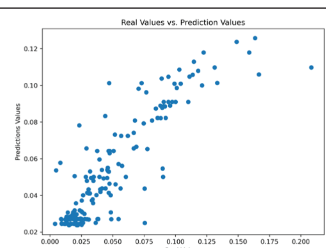
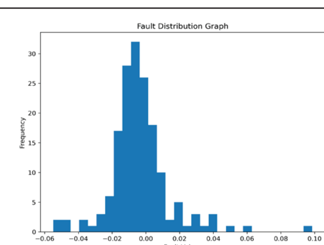
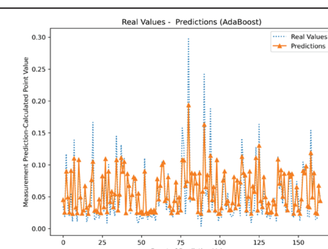
ALG.	Real Values & Prediction	Fault Distribution	Predicting Values
GB			
RF			
EXTRATREES			
ADABOOST			

Table 6 evaluates the performance of models trained with four different algorithms (GB, RF, ExtraTrees and AdaBoost). Again, various metrics such as MSE, MAE, RMSE, R^2 and Accuracy are used. The GB model has very low error rates and shows high performance in other metrics. In particular, the R^2 value of 0.99 represents a very high explanation rate. The accuracy rate is also quite high at 95.03%. RF and ExtraTrees models show increased error rates and decreased R^2 values. They perform slightly lower compared to the GB model. The AdaBoost model has higher error rates and lower R^2 values compared to the other algorithms, and its accuracy is significantly lower than the other models.

3.5. Trainings Conducted for the Number of Treatment Needs in the Hospital and Results

As can be seen from Table 7, the ExtraTrees Regressor algorithm is quite successful in the Real Values and

Table 6. Metrics for estimating the number of people who will potentially need shelter.

Algorithm	MSE	MAE	RMSE	R^2	Accuracy
GB	5.907810132908763e-06	0.0016719927767314552	0.0024305987190214604	0.99	95.03
RF	0.00011401742245214174	0.005899589280678199	0.010677894101935164	0.93	84.83
ExtraTrees	0.0008459893248166345	0.00980782252598891	0.029085895633736886	0.70	74.49
AdaBoost	0.00035198198721687386	0.012938087377026323	0.01876118299086904	0.75	50.22

Table 7: Graphs for estimating the number of people who will potentially require hospital treatment

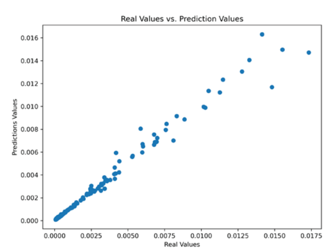
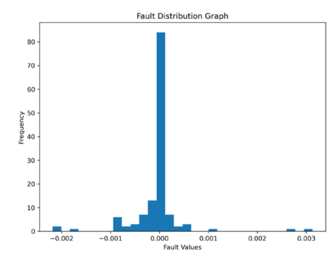
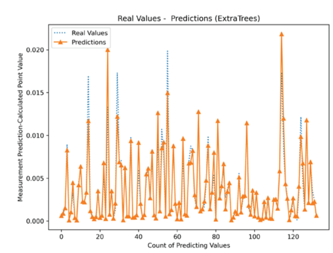
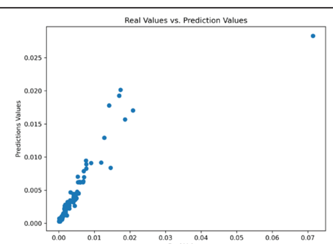
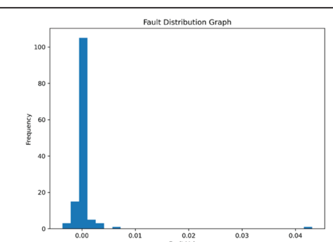
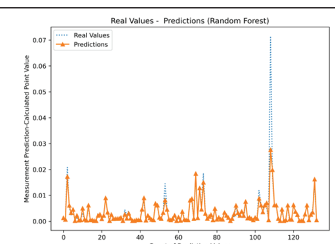
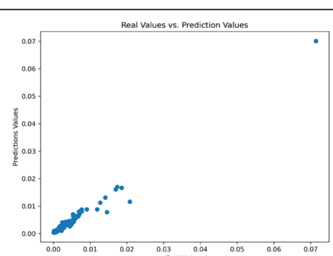
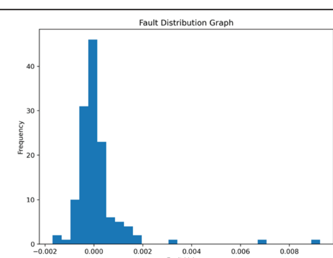
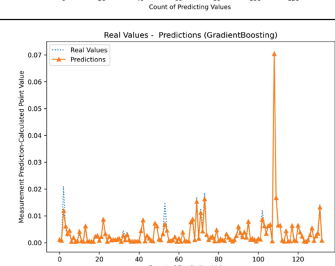
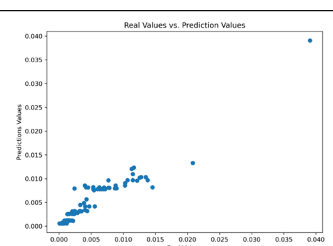
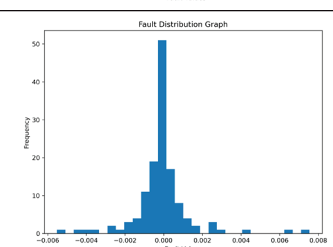
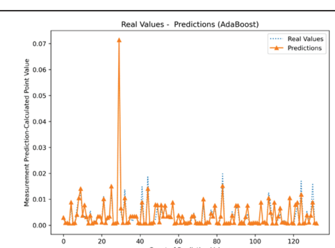
ALG.	Real Values & Prediction	Fault Distribution	Predicting Values
EXTRATREES			
RF			
GB			
ADABOOST			

Table 8. Metrics for estimating the number of people who will potentially require hospital treatment.

Algorithm	MSE	MAE	RMSE	R^2	Accuracy
ExtraTrees	2.917018737514533e-07	0.00022766724876766196	0.0005400943193104823	0.97	92.57
RF	1.4933443438588122e-05	0.0008048125853343117	0.0038643813785117173	0.70	72.05
GB	1.3869019221264514e-06	0.0005328583098645869	0.0011776680016568554	0.97	55.25
AdaBoost	2.221970996035838e-06	0.0008394849115527844	0.0014906277187936086	0.91	51.80

Prediction Values graph. In addition, although it looks like a scattered column graph in the error distribution graph, the tolerance multiplier value is quite low compared to other algorithms.

Table 8 shows the performance of the models trained with four different algorithms (ExtraTrees, RF, GB and AdaBoost). Again, various metrics such as MSE, MAE, RMSE, R^2 and Accuracy are used. The ExtraTrees model has very low error rates and a high R^2 value. The RMSE value is also quite low. This indicates that the model explains the dataset well and makes accurate predictions. The accuracy rate is also high at 92.57%. The RF model has higher error rates and lower R^2 value compared to the other models. However, its accuracy is still acceptable (72.05%). The GB and AdaBoost models perform moderately in terms of error rates and R^2 values. However, their accuracy rates are lower than the other models (55.25% and 51.80%).

According to the performance metrics given in the tables, the GB algorithm is superior to the other algorithms. This superiority is especially evident in metrics such as MSE, MAE, RMSE, R^2 and Accuracy.

One of the main reasons behind the success of the GB algorithm is that this algorithm has a structure that iteratively corrects errors. By correcting the errors at each step, GB improves model performance and can produce more accurate and generalized results. This iterative learning process allows the model to better learn the complexities in the dataset, resulting in lower error rates and higher accuracy rates. In terms of practical applications, a model with low MSE and RMSE values can make more accurate predictions, which can reduce the cost of error. A high R^2 value indicates that the model is better able to explain the variability of the data set, which increases model reliability. In line with this analysis, it is clear why the GB algorithm is more successful especially in our dataset. Therefore, it is recommended to prefer the GB algorithm for similar problems.

4. Discussion and Conclusion

Istanbul, as Turkey's largest metropolis, is notable for its dense population and cosmopolitan structure, yet it faces significant earthquake risks due to unplanned urbanization and a high proportion of structurally unsound buildings. This study utilized various ML techniques to predict potential casualties, the number of people requiring hospital treatment, and the number of people needing temporary shelter in the event of a potential magnitude 7.5 earthquake occurring during nighttime in Istanbul. The dataset used in this analysis was a composite set enriched with data obtained from the Istanbul Metropolitan Municipality Open Data Portal and the TUIK.

The dataset includes diverse features such as neighborhood-based building counts, distribution of buildings according to their construction years, and structural damage rates. These features are critical for understanding the potential outcomes of earthquake scenarios and planning necessary measures during model training.

The modeling process employed algorithms like GB, AdaBoost, RF, and ExtraTrees, evaluating each algorithm's performance using metrics such as MAE, RMSE, and the R^2. Results indicate that the GB algorithm outperformed others by achieving lower error rates and a high determination coefficient, signifying its ability to model the dataset with high accuracy and close proximity to true values.

When compared to the literature, similar studies also demonstrate the potential of ML approaches to successfully predict earthquake outcomes. While these studies provide significant insights, our study contributes uniquely by utilizing a hybrid dataset specific to Istanbul, which integrates neighborhood-based building counts, construction year distributions, and structural damage rates. Unlike the aforementioned studies that primarily focus on broad regional datasets or specific technical methods, our research emphasizes the application of multiple ML algorithms (GB, AdaBoost, RF, and ExtraTrees) on a localized and context-specific dataset. This allows for a more nuanced understanding of earthquake impacts in an urban setting like Istanbul. Furthermore, our approach highlights the practical implications for urban disaster preparedness and management, providing actionable insights for policymakers. Compared to previous studies, our work stands out in the following ways:

- Localized Dataset: The use of a hybrid dataset specific to Istanbul, incorporating detailed local attributes not commonly found in broader datasets.
- Algorithm Comparison: A comprehensive comparison of multiple ML algorithms on the same dataset, providing a clear evaluation of their relative performance.
- Practical Implications: Direct applicability to urban disaster management and preparedness in Istanbul, offering specific recommendations based on the findings.

These aspects of our study offer a valuable contribution to the existing body of literature, addressing the need for localized analysis and practical applications in disaster management.

This study utilizes a dataset specific to the Istanbul province. Future studies aim to conduct similar analyses in other regions of Turkey or in other earthquake-prone areas around the world. This will help assess how the proposed models perform under different geographical and demographic conditions. Additionally, it is intended to test new and emerging ML and deep learning techniques on the current topic, alongside the existing algorithms.

In conclusion, this study contributes to understanding the possible damages a major earthquake could cause in Istanbul and assists in planning to mitigate these impacts. The findings provide valuable insights for policymakers in enhancing urban transformation and disaster management strategies. Future research could extend similar modeling to different regions, further improving earthquake risk management and emergency preparedness.

5. Acknowledgment

The authors have not received any financial support for the research, authorship, or publication of this study. No conflict of interest or common interest has been declared by the authors.

Author contribution: The authors contributed equally to the study.

Ethics committee approval: This study does not require ethics committee permission or any special permission.

6. References

AlOmar, MK., Hameed, MM., AlSaadi MA. 2020. Multi hours ahead prediction of surface ozone gas concentration. Robust artificial intelligence approach Atmospheric Pollution Research, 11(9): 1572–1587. Doi: 10.1016/j.apr.2020.06.024

Ahamed, S., Daub, E. 2019. Machine learning approach to earthquake rupture dynamics. ArXiv, Doi:10.48550/arXiv.1906.06250

Ambraseys, N. 2001. The earthquake of 10 July 1894 in the Gulf of Izmit (Turkey) and its relation to the earthquake of 17 August 1999. Journal of Seismology, 5(1): 117-128. Doi: 10.1023/A:1009871605267

Arslan, G., Fawzy, D., Coskun, A. 2017. On the prediction of structural reactions to big earthquakes in Turkey. PressAcademia Procedia. 5(1): 335-340. Doi: 10.17261/Pressacademia.2017.608

Breiman, L., 2001. Random forests. Machine learning. 45(1): 5–32. Doi: 10.1023/A:1010933404324

Corbi, F., Sandri, L., Bedford, J., Funiciello, F., Brizzi, S., Rosenau, M., Lallemand, S. 2019. Machine learning can predict the timing and size of analog earthquakes. Geophysical Research Letters, 46: 1303 - 1311. Doi:10.1029/2018GL081251.

Erdik, M. 2001. Report on 1999 Kocaeli and Düzce (Turkey) earthquakes. Structural Control For Civil And Infrastructure Engineering. p. 149-186. Paris. France.

Freund, Y., Schapire, RE. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences. 55(1): 119-139. Doi: 10.1006/jcss.1997.1504

Friedman, JH. 2001. Greedy function approximation: A gradient boosting machine. Annals of Statistics. 29(5):1189-1232. Doi: 10.1214/aos/1013203451

Geurts, P., Ernst, D., Wehenkel, L. 2006. Extremely randomized trees. Machine Learning. 63(1): 3–42. Doi: 10.1007/s10994-006-6226-1.

Hammid, AT., Sulaiman, MH, Abdalla, AN. 2018. Prediction of small hydropower plant power production in Himreen Lake dam (HLD) using artificial neural network. Alexandria Engineering Journal. 57(1): 211–221. Doi: 10.1016/j.aej.2016.12.011.

IMM Open Data Portal 2017. Neighborhood based building numbers. https://data.ibb.gov.tr/dataset/mahalle-bazli-bina-analiz-verisi/resource/cef193d5-0bd2-4e8d-8a69-275c50288875

IMM Open Data Portal 2021. Earthquake scenario analysis results. https://data.ibb.gov.tr/dataset/deprem-senaryosu-analiz-sonuclari/resource/9c3ac492-de4b-4245-b418-7ad3df67a193

James, G., Witten, D., Hastie, T., Tibshirani, R. 2013. An introduction to statistical learning. Springer. New York, USA, 607 pp.

Jaiswal, K., Wald, DJ. 2010. An empirical model for global earthquake fatality estimation. Earthquake Spectra, 26(4): 1017-1037. Doi: 10.1193/1.3480331

Korkut, U., Basbugoglu, T., Sahin, O. 2023. Turkey's 2023 elections: another victory for Erdogan. Political Insight, 14(3): 16-19. Doi: 10.1177/20419058231198579

Li, B., Gong, A., Zeng, T., Bao, W., Xu, C., Huang, Z. 2021. A zoning earthquake casualty prediction model based on machine learning. Remote Sensing, 14(1): 30. Doi: 10.3390/rs14010030

Mishra, G., Sehgal, D., Valadi, JK. 2017. Quantitative structure activity relationship study of the antihepatitis peptides employing random forests and extra-trees regressors. Bioinformation, 13(3): 60, Doi: 10.6026/97320630013060.

Parsons, T. 2000. Heightened odds of large earthquakes near Istanbul: An interaction-based probability calculation. Science, 288(5466): 661-665, Doi: 10.1126/science.288.5466.661

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C., Johnson, P. 2017. Machine learning predicts laboratory earthquakes. Geophysical Research Letters. 44: 9276 - 9282. Doi:10.1002/2017GL074677.

Stein, RS. 2000. Earthquake conversations. Scientific American, 282(1): 68-75. Doi:10.1038/scientificamerican012003-DipeiQnMaPMOL1RMjOPhi

Willmott, CJ., Matsuura, K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research. 30(1): 79–82, Doi: 10.3354/cr030079.

World Bank 2023. Earthquake damage in Türkiye estimated to exceed $34 billion: World bank disaster assessment report. https://www.worldbank.org

Xing, H., Junyi, S., Jin, H. 2020. The casualty prediction of earthquake disaster based on Extreme Learning Machine method. Natural Hazards, 102(3): 873-886, Doi: 10.1007/s11069-020-03937-6