Earthquake response and recovery

K.K.Charan Reddy, Rahul Chaudhary, Jaggumantri Surya, M.Pavan

Abstract

Earthquake prediction is one of the most focused areas in disaster management, especially with the emergence of more advanced machine learning. In a world highly based on data collection, visualization, and modeling, massive seismic data can be crucial to enabling accurate, reliable, and precise earthquake prediction. This paper introduces a decision tree-based earthquake prediction model that uses random forest, Decision Tree and gradient boost. Four critical phases are involved in this system: data collection, model training and predictive analytics application. The features of historical seismic data are analyzed to determine the patterns that support prediction including magnitude, depth, and location. However, there are still open challenges, such as the quality of the data and integration of diverse data sources. The areas mentioned thus far represent avenues or locations for further development and improvement, making systems responsive to combating earthquake-related disasters.

keywords: Earthquake Prediction, Machine Learning, Decision Tree, Random Forest, Gradient Boost, Predictive Analytics, Seismic Data, Historical Seismic Data, Data Integration.

Introduction

Earthquakes are among the most destructive natural disasters, leading to tremendous economic and human losses. Earthquake prediction is still a major challenge because of the unpredictable nature of seismic processes. Machine learning (ML) has now become a potent tool in interpreting intricate patterns of data and making predictions in many fields, including earthquake prediction. Classical methods for earthquake prediction involve statistical seismology, ground deformation monitoring, early warning systems, geodetic observations, and physics-based models. These methods are informative, but they may lack real-time adaptability and precision owing to the randomness of seismic occurrences. In this study, we utilize ML methods to process large volumes of seismic data and identify concealed patterns that can indicate future earthquakes.The data employed in this research is obtained from seismic monitoring organizations like the United States Geological Survey (USGS), the Japan Meteorological Agency, and the website ds.iris, providing high-quality and trustworthy data for training our models.

Literature Review

1. **Improving earthquake prediction accuracy in Los Angeles with machine learning:** - *Cemil Emre Yavas, Lei Chen, Christopher Kadlec and Yiming Ji*
   This study innovates earthquake prediction in Los Angeles, California, by utilizing advanced machine learning and neural network techniques. We carefully built a rich feature set to optimize predictive performance. By combining past work and merging new predictive features, we established a strong subset to estimate the greatest possible earthquake magnitude. Our best contribution is establishing a feature set that, when used in combination with the Random Forest machine learning algorithm, attains a high level of precision

in forecasting the maximum earthquake category in the subsequent 30 days. Out of sixteen machine learning algorithms tested, Random Forest was the most efficient. Our results highlight the revolutionary potential of neural networks and machine learning to improve earthquake prediction accuracy, providing valuable improvements in seismic risk preparation and management in Los Angeles.[5]

2. **Earthquake magnitude prediction in Hindukush region using machine learning techniques** - *K. M. Asim1, F. Martı´nez-A ´ lvarez2, A. Basit3, T. Iqbal*
focuses on using machine learning (ML) models to predict earthquake magnitudes in the seismically active Hindukush region. Given the area's frequent earthquakes due to tectonic activity, accurate magnitude prediction is crucial for disaster preparedness and risk mitigation. The study explores different ML techniques, including regression models and deep learning approaches, to analyze historical earthquake data. The researchers preprocess seismic datasets, extracting relevant features such as location, depth, and time of past earthquakes. Various ML algorithms, including decision trees, support vector machines, and neural networks, are applied to predict the magnitude of future seismic events. The results indicate that ML models can provide reasonable accuracy in predicting earthquake magnitudes, though challenges remain due to the complex and nonlinear nature of seismic activity. The study emphasizes the potential of ML-based approaches in earthquake forecasting, contributing to early warning systems and disaster management efforts. The research highlights the importance of continuous data collection and model improvement for better predictive performance in seismically vulnerable regions like Hindukush.[1]

3. **Major earthquake event prediction using various machine learning algorithms** - *Roxane Mallouhy, Chady Abou Jaoude, Christophe Guyeux, Abdallah Makhoul*
Aim to improve earthquake forecasting by leveraging advanced computational techniques. The study examines different ML models to analyze seismic data and identify patterns that could indicate an impending earthquake. These models include decision trees, support vector machines (SVM), artificial neural networks (ANN), and deep learning approaches. By training these algorithms on historical earthquake records and relevant geological features, the researchers assess their predictive accuracy and effectiveness. One of the key challenges in earthquake prediction is the complexity and randomness of seismic activity. The paper addresses this by comparing the performance of multiple ML techniques to determine which provides the most reliable forecasts. The study also discusses feature selection, data preprocessing, and evaluation metrics for assessing model accuracy. The results indicate that no model can ever predict earthquakes with absolute certainty, but ML methods greatly improve the capacity to identify possible seismic activity. The study adds to current research in earthquake prediction, providing insights that could be used to reduce disaster risks and enhance early warning systems.[3]

4. **Earthquake magnitude prediction in Turkey: a comparative study of deep learning methods,**

ARIMA and singular spectrum analysis - *Hatice Öncel Çekim1, Hatice Nur Karakavak1, Gamze Özel1, Senem Tekin*
This research paper examines various ways of predicting earthquake magnitudes in Turkey, a country with high seismic activity. The authors compare deep learning models, the ARIMA (AutoRegressive Integrated Moving Average) model, and Singular Spectrum Analysis (SSA) in order to identify which method makes the best predictions. Deep learning techniques, such as Long Short-Term Memory (LSTM) networks, are compared with conventional statistical techniques such as ARIMA, which is widely applied for time series forecasting. Singular Spectrum Analysis, a data-driven method for pattern identification in time series data, is also compared. The research uses past earthquake records from Turkey to train and validate the models for making future magnitude predictions. Comparison of the prediction accuracy is made by using performance metrics like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Results indicate that deep learning models, especially LSTM, perform better than ARIMA and SSA in recognizing intricate seismic patterns, the study found that the LSTM model delivered the most precise forecasts for the subsequent 36 months. Nevertheless, ARIMA is still a good baseline for short-term forecasts. The study brings out the promise of AI models for earthquake forecasting to improve disaster readiness. The study calls for greater enhancements by combining additional geological and environmental indicators to make predictions more accurate.[4]

5. **Big data analytics in prevention, preparedness response and recovery in crisis and disaster management** - *D Emmanouil, D Nikolaos*
Emmanouil and Nikolaos, in their 2015 research, discuss the role of big data analytics in revolutionizing the four key stages of crisis and disaster management: prevention, preparedness, response, and recovery. They highlight that the quality of decisions made during disasters is contingent upon the quality of data and information received. Through the extensive analysis of data received during the occurrence of disasters, decision-makers are able to make accurate decisions within constrained and ad-hoc time periods. The writers call for disaster-related data to be collected efficiently and effectively beforehand. In this proactive manner, historical and anticipated data—like rainfall, temperature, drought, and streamflow indicators—would be analyzed to forecast impending disaster occurrences. Through predictive analysis, its possible injuries, health effects, property damage, loss of lives, socio-economic disruptions, and environmental degradation are diminished by creating overall mitigation, control, and prevention programs beforehand. The research highlights the need to incorporate big data analytics into disaster management systems to move from reactive responses to proactive planning, eventually leading to increased resilience and minimizing the negative impacts of disasters.[2]

Methods and Algorithms Used

Decision Trees: A tree-based algorithm that could classify seismic data based on features like magnitude, depth, and location. Such can provide interpretable decision paths for identifying the crucial factors that lead to earthquake occurrences. Handles both numerical and categorical data efficiently and a base model for comparison.

Random Forest Model: An ensemble learning technique that builds multiple decision trees and combines their outputs to improve prediction stability and reduce overfitting. Randomly selects features and subsets of data to train individual trees, ensuring model robustness and generalization. Especially good for working with complex patterns in seismogram data.

Gradient Boost: The boost algorithm is able to enhance predictions by iteratively correcting the wrong predictions of prior models. Combines weak learners (shallow decision trees) to build a strong predictive model. Suitable for handling imbalanced datasets and delivering high-performance results with fine-grained tuning.

Support Vector Machine: Support Vector Machine (SVM) is a supervised learning algorithm of machine learning applied to classification and regression problems. Although it supports regression problems, SVM is specially designed for classification problems. it tries to find the best hyperplane in N-dimensional space in order to separate data points from various classes. The algorithm finds the maximum margin between the closest points of other classes.

KNearest Neighbour: K-Nearest Neighbors (KNN) is an easy method of classifying things based on what's around. Think of a streaming service that would like to predict whether a new user will cancel their subscription (churn) or not based on their age. They examines the ages of its current users and whether they churned or not. If the majority of the "K" nearest neighbors in age of new user canceled their subscription KNN will predict new user may also churn. The fundamental concept is users of similar ages are likely to behave similarly and KNN takes advantage of this proximity to make predictions.

Results

Model	Accuracies
Random Forest Regressor	93.27
Support Vector Machine	92.13
Decision Trees Regressor	90.73
K-Nearest Neighbour	91.68
XGBoost	93.03
Light GBM	93.36
Multi-Layer Perceptron (MLP)	92.71
Recurrent Neural Networks	91.15
LSTM	90.79
Gated Recurrent Unit	91.26
Transformer	86.53

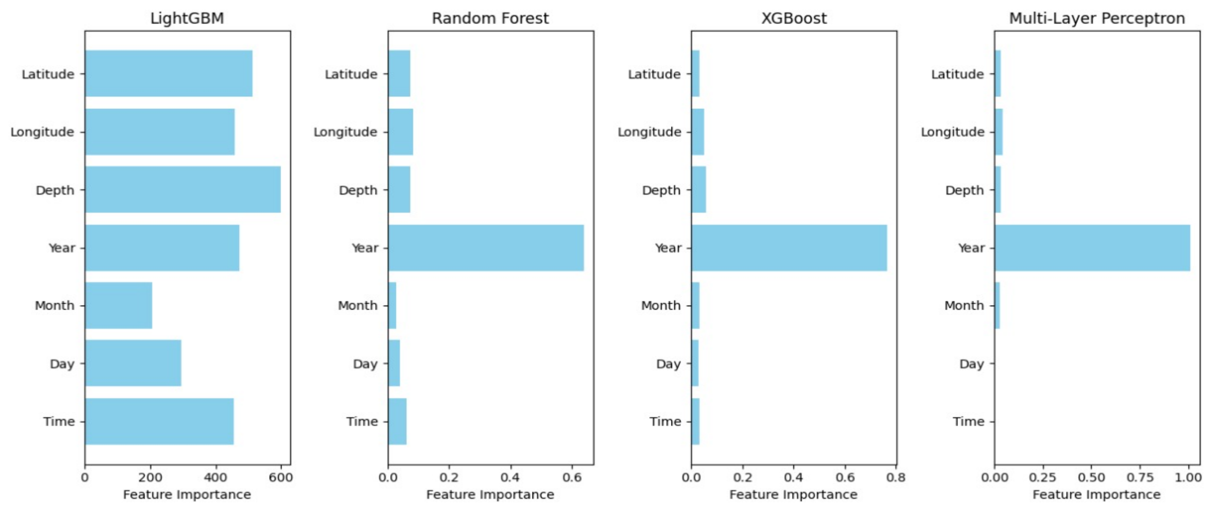


Figure 1: Graphical representation of model performance

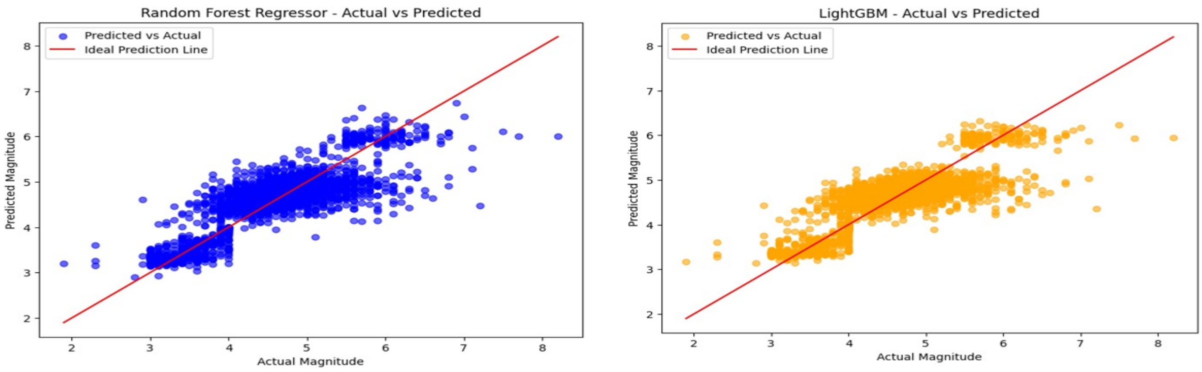


Figure 2: Graphical representation of model performance

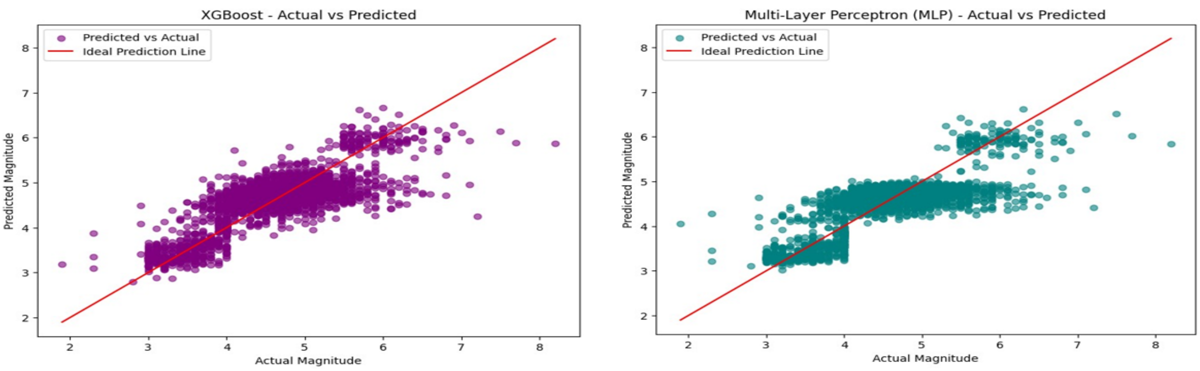
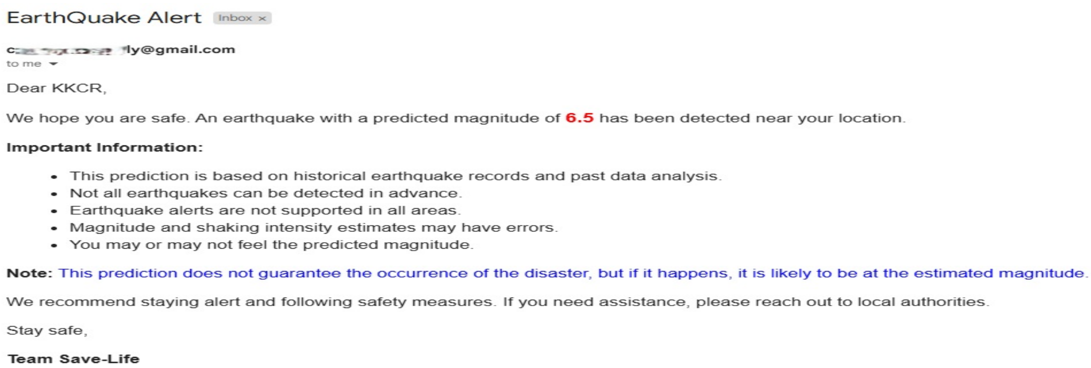


Figure 3: Graphical representation of model performance



EarthQuake Alert Inbox ×

c[redacted]ly@gmail.com
to me

Dear KKCR,

We hope you are safe. An earthquake with a predicted magnitude of **6.5** has been detected near your location.

Important Information:

- This prediction is based on historical earthquake records and past data analysis.
- Not all earthquakes can be detected in advance.
- Earthquake alerts are not supported in all areas.
- Magnitude and shaking intensity estimates may have errors.
- You may or may not feel the predicted magnitude.

Note: This prediction does not guarantee the occurrence of the disaster, but if it happens, it is likely to be at the estimated magnitude.

We recommend staying alert and following safety measures. If you need assistance, please reach out to local authorities.

Stay safe,

Team Save-Life

Figure 4: Pictorial representation of Alert Email

Conclusion

The present project is the solution to the challenge of predicting and managing disasters through advanced machine learning models, including Decision Trees, Random Forest, Gradient Boosting, and a Hybrid Model. All these techniques will improve the precision and reliability of the predictions made, which means an early warning system to the vulnerable communities to mitigate the impact of the natural disasters. Through handling historical and real-time data processing, the system affords critical inputs for decision making, thus enhancing timely responses towards disasters. Robustness as well as the possibility to adapt to every type of disaster and region renders the hybrid approach suitable. This project showcases the integration of machine learning into disaster management for the transformation of aged systems into proactive data-driven solutions. Further development combined with diverse datasets can provide this solution as an imperative approach for governments and organizations in preventing losses during disasters and building resilience toward catastrophes in the future

References

[1] KM Asim, Francisco Martínez-Álvarez, Abdul Basit, and Turab Iqbal. Earthquake magnitude prediction in hindukush region using machine learning techniques. *Natural Hazards*, 85:471–486, 2017.

[2] Dontas Emmanouil and Doukas Nikolaos. Big data analytics in prevention, preparedness, response and recovery in crisis and disaster management. In *The 18th International Conference on Circuits, Systems, Communications and Computers (CSCC 2015), Recent Advances in Computer Engineering Series*, volume 32, pages 476–482, 2015.

[3] Roxane Mallouhy, Chady Abou Jaoude, Christophe Guyeux, and Abdallah Makhoul. Major earthquake event prediction using various machine learning algorithms. In *2019 international conference on information and communication technologies for disaster management (ICT-DM)*, pages 1–7. IEEE, 2019.

[4] Hatice Öncel Çekim, Hatice Nur Karakavak, Gamze Özel, and Senem Tekin. Earthquake magnitude prediction in turkey: a comparative study of deep learning methods, arima and singular spectrum analysis. *Environmental Earth Sciences*, 82(16):387, 2023.

[5] Cemil Emre Yavas, Lei Chen, Christopher Kadlec, and Yiming Ji. Improving earthquake prediction accuracy in los angeles with machine learning. *Scientific Reports*, 14(1):24440, 2024.

Additional References

NOAA:
https://www.climate.gov/maps-data/dataset/past-weather-zip-code-data table
OpenWeather Map:
https://openweathermap.org/current
Interactive Earthquake Browser:
https://ds.iris.edu/ieb/index.html
Humanitarian Data Exchange(HDX):
https://data.humdata.org/dataset/reliefweb-disasters-list
EM-DAT:
https://www.emdat.be/
FEMA:
https://www.fema.gov/about/openfema/data-sets
Global Disaster Alert and Coordination System (GDACS):
https://www.gdacs.org/