National Scale Earthquake Susceptibility Mapping Utilizing Explainable Artificial Intelligence in The Nepal Himalaya

Suchita Shrestha[1*] and Tae-Seob Kang[2]

[1]Department of Mines and Geology, Ministry of Industry, Commerce and Supplies, Government of Nepal, Lainchour, Kathmandu, Nepal (suchitashrestha@gmail.com)

[2]Division of Earth Environmental System Science, Pukyong National University, Busan 48513, Republic of Korea

(*Corresponding E-mail: suchitashrestha@gmail.com)

Abstract: Nepal faces high earthquake risk due to its active tectonics and rapid development on vulnerable terrain. This study presents a national-scale earthquake susceptibility assessment using explainable artificial intelligence. Historical earthquake records were combined with key geophysical and geomorphic factors, including fault proximity, fault density, tectonic zones, topography, and seismic event density. Random Forest and Extreme Gradient Boosting models were developed to estimate spatial earthquake probability, and explainable AI techniques were applied to interpret variable importance and model behavior. The Random Forest model achieved higher accuracy and lower uncertainty compared to XGB. The resulting maps highlight clusters of elevated seismic probability along major fault corridors. The approach supports transparent, data-driven seismic risk evaluation suitable for planning and disaster preparedness in Nepal.

Keywords: ***Earthquake, Nepal Himalaya, Artificial intelligence.***

Introduction

Nepal is undergoing significant annual population growth, rapid urbanization, and extensive infrastructure development. Unfortunately, this rapid development, often without adherence to earthquake codes, has heightened the region's earthquake risk. The seismic activity observed in recent decades underscores the potential for future earthquakes to impact this densely populated area (Gautam and Chaulagain, 2016). This study focused on enhancing spatial probability mapping using explainable artificial intelligence (XAI).

Traditional seismic hazard zone (SHZ) methods typically rely on statistical analyses of historical earthquake data, which can be limited in scope. XAI offers a solution by integrating diverse geological, geomorphological, and geophysical data, thereby expanding the analytical framework. Unlike traditional methods like probabilistic seismic hazard analysis, which often involve complex mathematical models and algorithms, XAI employs machine learning (ML) algorithms designed to generate transparent and interpretable outcomes. This makes it more accessible for scientists, engineers, and policymakers to grasp the underlying factors influencing seismic hazard assessments.

Methodology

The earthquake catalog encompasses data from all historical earthquake events within the study area, regardless of their sources, magnitude scales, recording agencies, or event sizes. Geo-related covariates are important in earthquake studies because they can help identify areas that are likely to experience earthquakes. Previous studies on earthquake hazards and probabilities derived several relevant factors for earthquake probability assessments. Elevation (Elv), topographic position index (TPI), magnitude density (MagDen), depth density (DepDen), fault proximity (FaultProx), fault density (FaultDen), and “tectonic zone” covariates are crucial factors that contribute significantly to the assessment of earthquake probability.

The rise of AI has been meteoric in recent years, promising transformative changes across various industries. However, the increasing reliance on black-box machine learning (ML) models for critical predictions has raised concerns about the opacity of their decision-making processes. This lack of transparency has sparked a demand among stakeholders for more understandable AI systems.

Explainable AI (XAI) aims to address this challenge by developing ML techniques that are not only accurate but also interpretable. When applying XAI techniques alongside modeling approaches such as Random Forest (RF) and Extreme Gradient Boosting (XGB), the objective is to create models that retain high accuracy while providing clear explanations of their reasoning.

Results and Discussion

We utilized a set of seven geo-related covariates to develop two predictive models: Random Forest (RF) and Extreme Gradient Boosting (XGB). Among these covariates, "tectonic" was a categorical factor. The values of all seven covariates were numerically encoded in a spreadsheet, with earthquake and non-earthquake cases designated as the target outcomes.

Numeric values representing earthquake probability were assigned to distinct datasets. These forecasted values, confined within the 0 to 1 range, can be interpreted as probabilities. Figure 1a illustrates the distribution of earthquake probability values obtained using the RF model, while Figure 1b depicts the earthquake probabilities generated by the XGB model across the Nepalese Himalayas.

In the overall process, 1,710 earthquake and non-earthquake instances were analyzed. The RF model accurately classified 801 out of 855 earthquake events and 800 out of 855 non-earthquake events. In comparison, the XGB model correctly identified 765 earthquake events and 741 non-earthquake events. The misclassification indices, determined using a confusion matrix, indicate an uncertainty of 6.37% for the RF model and 11.93% for the XGB model.

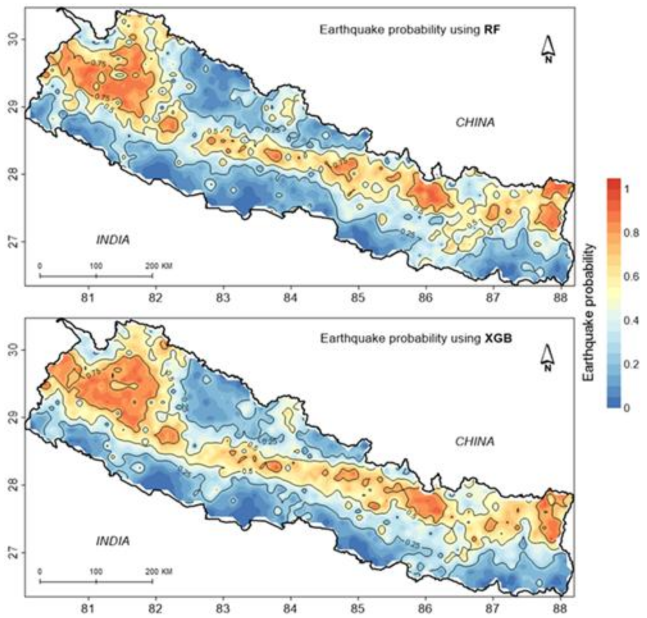

Figure 1, Earthquake probability map of Nepal generated by the a) random forest (RF) and b) extreme gradient boosting (XGB) models.

Conclusions

The RF model exhibited excellent predictive performance with an AUC of 0.79, while the XGB model demonstrated an AUC of 0.76, indicating a strong goodness-of-fit. For future studies, we recommend exploring the impacts of additional geo-related covariates and incorporating findings from precursor studies to enhance the accuracy of earthquake probability assessments.

References

Gautam, D., and Chaulagain, H. (2016). Structural performance and associated lessons to be learned from world earthquakes in Nepal after the 25 April 2015 (Mw 7.8) Gorkha earthquake. Engineering Failure Analysis, 68, 222–243. https://doi.org/10.1016/j.engfailanal.2016.06.002