For more context:

Early warning of Earthquake using machine learning

Nikhil Reddy, Yuyan Chen, Santhoshi Ravichandran, Ashwini Rajaram
Mentors : Laurent Barcelo, Reda Snaiki, Alex Hernandez

Main Contribution(s):

In this study, our primary focus was on leveraging the INSTANCE (Italian seismic dataset)[1] to predict the onset and occurrence of earthquakes. Notably, current early warning systems are only able to provide alerts 3-5 seconds after the onset of an earthquake[2]. Our primary objective was to predict earthquake onsets with a lead time of up to 120 seconds, enabling timely emergency responses. To achieve this, we employed seismic data from earthquake events and noise data as distinct classes, utilizing both Random Forest (RF) and Neural Networks for pattern recognition. Our RF model achieved results with an 87% test accuracy and an F1-score of 0.84.

Problems Encountered:

A significant challenge we encountered was the limited availability of datasets with information about seismic waveforms and data trends. Most existing datasets primarily provide information about earthquake location and magnitude, lacking waveform information crucial in predicting and forecasting earthquakes. Due to these dataset limitations, we were constrained to utilize 120-second waveforms. However, employing longer-duration time-series data, spanning several hours and originating from the same location, holds the potential to enhance our ability to recognize patterns associated with earthquake onset. Unfortunately, our project faced resource and time limitations that prevented us from conducting extensive data preprocessing and fine-tuning of our models. Additionally, it became apparent that a profound understanding of earthquake domain knowledge is imperative for the development of effective forecasting models and early warning systems. As a result of the randomness involved in earthquakes, accurate early prediction happens to be a complex task.

Key ways in which centralized entities can contribute to our efforts include:

1. **Access to Comprehensive Datasets:** Providing access to large-scale, long-term datasets that combine seismic waveform data with other critical earthquake-related features, such as epicenter, magnitude, duration, and the

occurrence of tsunami and aftershocks. This comprehensive dataset would significantly enhance our prediction and estimation capabilities.

2. **Interdisciplinary Collaboration:** Collaborative efforts involving seismologists, computer scientists, governmental bodies, and emergency management organizations can harness diverse expertise and resources. This multidisciplinary approach can contribute to the creation of a robust and efficient early warning system.

Key Takeaways:

Our key takeaways include the importance of using the right dataset for training machine learning models. Beyond early warning systems, machine learning applications extend to earthquake magnitude estimation, epicenter identification, and post-earthquake tsunami alerts.The aim to further investigate the dataset, collaborate with domain experts, and combine data to build a comprehensive prediction dataset for various earthquake mitigation strategies. Collaborative work between various domains, such as seismology and computer science, can lead to impactful results.