Building a New Earthquake Early Warning System Using Machine Learning

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

Main Contribution(s):

Current early-warning alerts are sent around three to four seconds after an earthquake starts [1]. In this project, we leveraged the INSTANCE (Italian seismic dataset)[2] to predict the occurrence of earthquakes 10 seconds prior. It includes 3-channel waveform recordings for both earthquake and noise. Each waveform has a length of 120 seconds, sampled at 100 Hz. For our use case, we only use the first 10 seconds of each recording since for most earthquake recordings, the P-wave onset, or the start of the earthquake, happened after the first 20 seconds. We trained Random Forest (RF) and Neural Network models, and our RF achieved 87% test accuracy and an F1-score of 0.84.

Challenges faced:

- Limited availability of datasets with information about seismic waveforms and data trends. Most existing datasets primarily provide information about earthquake location and magnitude, lacking vital waveform information.
- Due to these dataset limitations, we were constrained to utilize 120-second waveforms. Employing longer-duration time-series data, spanning several hours and originating from the same location, enhances our ability to recognize patterns associated with earthquake onset.

How centralized entities can help:

- Provide access to more comprehensive large-scale long-term datasets, potentially combining waveform data with other earthquake-related features including epicenter, magnitude, duration, occurrence of tsunami and aftershocks for better prediction and estimation.
- Centralized entities can bring collaborations among seismologists and emergency management organizations whose expertise can help computer scientists better understand the mechanism and physics of earthquake and earthquake rescue, so that they can together build a more comprehensive early warning system.

Key Takeaways:
Having a solid early warning system can save lives, reduce injuries, and better coordinate emergency response organizations. Yet, it is a very challenging task to predict earthquake occurrence. Having a comprehensive dataset is crucial for effective ML in earthquake

predictions. By combining data and expertise from fields like seismology and computer science, we can enhance earthquake mitigation strategies.

References:

[1] Allen, Richard M., et al. "The status of earthquake early warning around the world: An introductory overview." *Seismological Research Letters* 80.5 (2009): 682-693.

[2] Michelini, Alberto, et al. "INSTANCE–the Italian seismic dataset for machine learning." *Earth System Science Data* 13.12 (2021): 5509-5544.