# **Predicting Earthquakes by Analyzing Seismic Signals**

**Problem statement**

Seismic waves are a measurement of energy in the form of vibrations that travel through the earth, which are used by earth scientists as the primary tool for understanding and predicting earthquakes. Early predictions are extremely useful in protecting people’s lives and preventing destruction of property by timely evacuation and other preemptive actions.

Recently, there have been efforts to use machine learning methods on seismic signals for earthquake prediction. The Los Alamos National Laboratory (LANL) has posted on Kaggle a dataset collected from earthquake simulations, and the challenge is to develop machine learning models for this task. In this project, we will address this challenge.

Link to competition and dataset: <https://www.kaggle.com/c/LANL-Earthquake-Prediction>

**Dataset**

The training dataset has 2 columns (x, y), representing a single, continuous seismic signal. x is the signal value, and y is the time-to-next-earthquake. The test data consists of segments of seismic signals (x values), and for each signal point in a segment, a time-to-next-earthquake (y) must be predicted by a machine learning model.

The dataset is large with over 600 million data points in the training signal (9 GB in csv format) and working with it will prove to be challenging (loading, visualizing, training models etc.).

**Project milestones and expected results**

* Dataset exploration and visualization. Some research on domain. Result: a good understanding of the problem and methods to work with the large dataset.
* Preprocessing and feature engineering (if necessary). Result: enhanced dataset.
* Decide on potential models (eg: time series models such as recurrent neural networks). and experiment with them. Result: shortlisted feasible models.
* Further experimentation and validation of the models and results (tuning, variations, combinations).
* Presenting our findings and writing the final report.

**Group Members**

|  |  |
| --- | --- |
| Samer El Kababji | 250936038 |
| Sunanda Gamage | 250992128 |
| Bryan Southwell | 251058589 |