Earthquake Classification using Machine Learning

Ganyu Teng: ganyut@stanford.edu

Motivation

This project aims to using machine learning methods (SVM, NN, 1D CNN) that classify the natural and induced earthquakes. Such technique is useful especially for places that are tectonic active and have ongoing oil or geothermal production (i.e., Sichuan Basin, China; Geysers, CA). In such cases, it is important to classify whether an observed earthquake is induced by human activities to have a better hazard management.

Problem

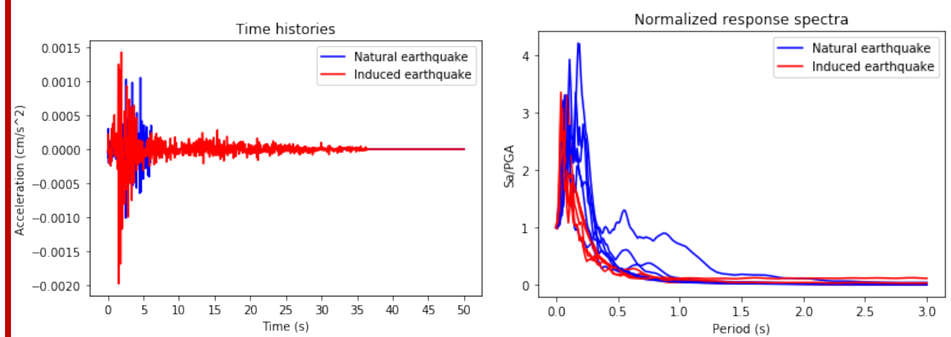
Given a response spectra or seismogram of an earthquake, is the earthquake natural or induced?

Data and Processing

We collected seismograms and response spectra for induced and natural earthquakes. The seismogram is the time history describes the shaking along the time. The response spectrum (the maximum response of a single-degree of freedom object) is obtained from the earthquake time history and widely used in the structural engineering.

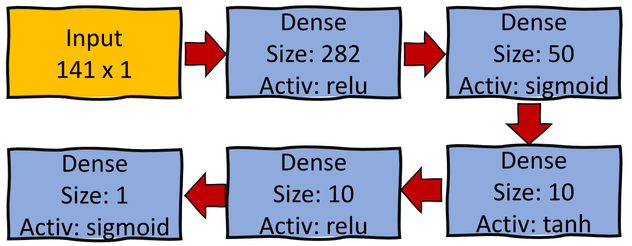
- Induced earthquakes (34,785 seismograms):
  - Time: 2013/01/01 – 2019/01/01
  - Latitude [35.0°, 38.0°]; Longitude [-100.0°, -96.0°]
  - Magnitude: [3.0, 3.5]
- Natural earthquakes (3,868 seismograms): :
  - Time: 2010/01/01 – 2019/01/01
  - Latitude [36.0°, 39.0°]; Longitude [-124.0°, -100.0°]
  - Magnitude: [3.0, 3.5]

10% of the data was used as test set, 18% was validation set.

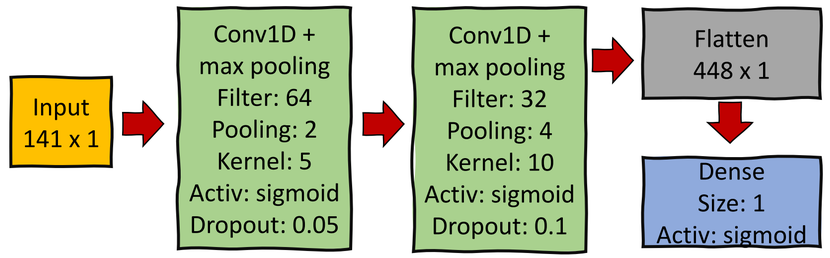


Approach

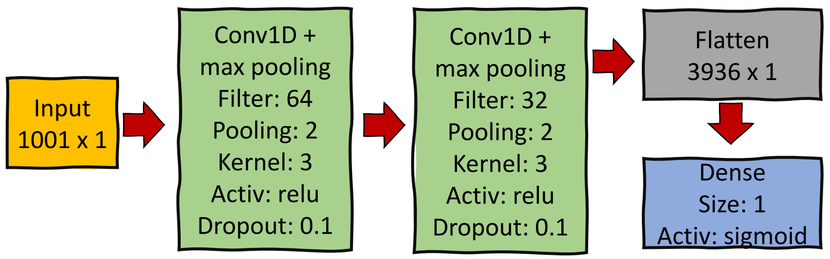
- SVM on response spectra (after hypertuning):
  - Kernel: Polynomial; Degree: 5; C: 1.0; Gamma: 0.01
- NN on response spectra (after hypertuning):



- 1D CNN on response spectra (CNN Sa, after hypertuning):

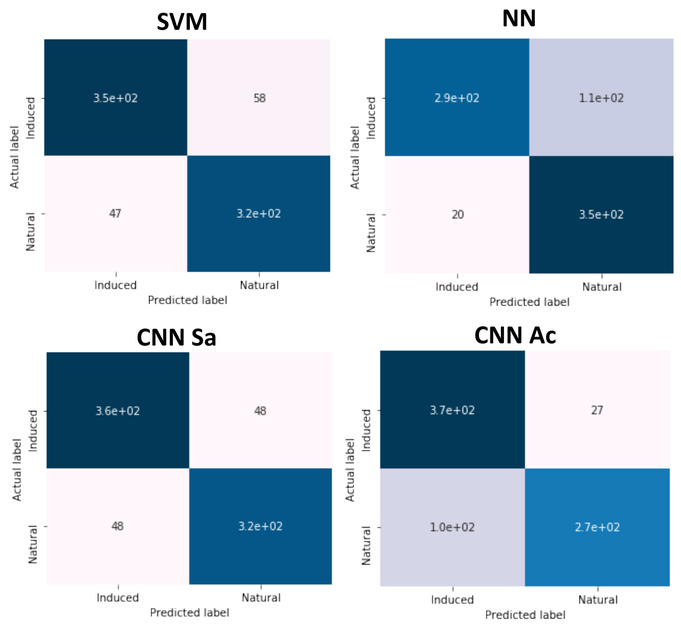


- 1D CNN on seismograms (1D CNN Ac, no hypertuning):



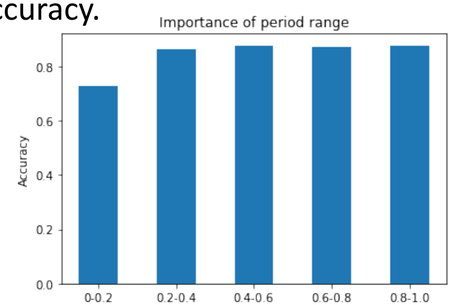
Results

Model	Training accuracy	Test accuracy
SVM	88.8%	86.4%
NN	88.6%	83.1%
CNN Sa	89.1%	87.6%
CNN Ac	84.6%	82.9%



Discussion

We study which period range is most important by randomizing data of that range. We found out that the shorter periods (<0.2s) are more important than other periods. The values at longer periods barely contribute to the accuracy.



Future works

- Consider spectrograms which include information about frequency
- Study and classify different types of induced earthquakes (e.g. hydraulic fracturing and water disposal)