Predicting Earthquake Locations From Accelerometer Data

GA-DSI 2020
Patrick Dornian

Background - OpenEEW and SSN

The west coast of Mexico is a hotbed of seismic activity and an ideal testing ground for new seismic detection technology.

OpenEEW (Early Earthquake Warning) is an ongoing open source project based in Mexico that provides years worth of readings from an IoT network of accelerometer based seismic sensors.

The Mexican Servicio Sismológico Nacional (SSN) provides a comprehensive searchable database of detected seismic activity in the country with logs of location, time, and earthquake magnitude.

Cross referencing these two data sources lets us investigate a number of questions.

Problem Statement

Given five minutes of accelerometer data from all active devices in the OpenEEW network at the time of a known seismic event, can we predict the latitude and longitude of the earthquake epicentre with any degree of accuracy?

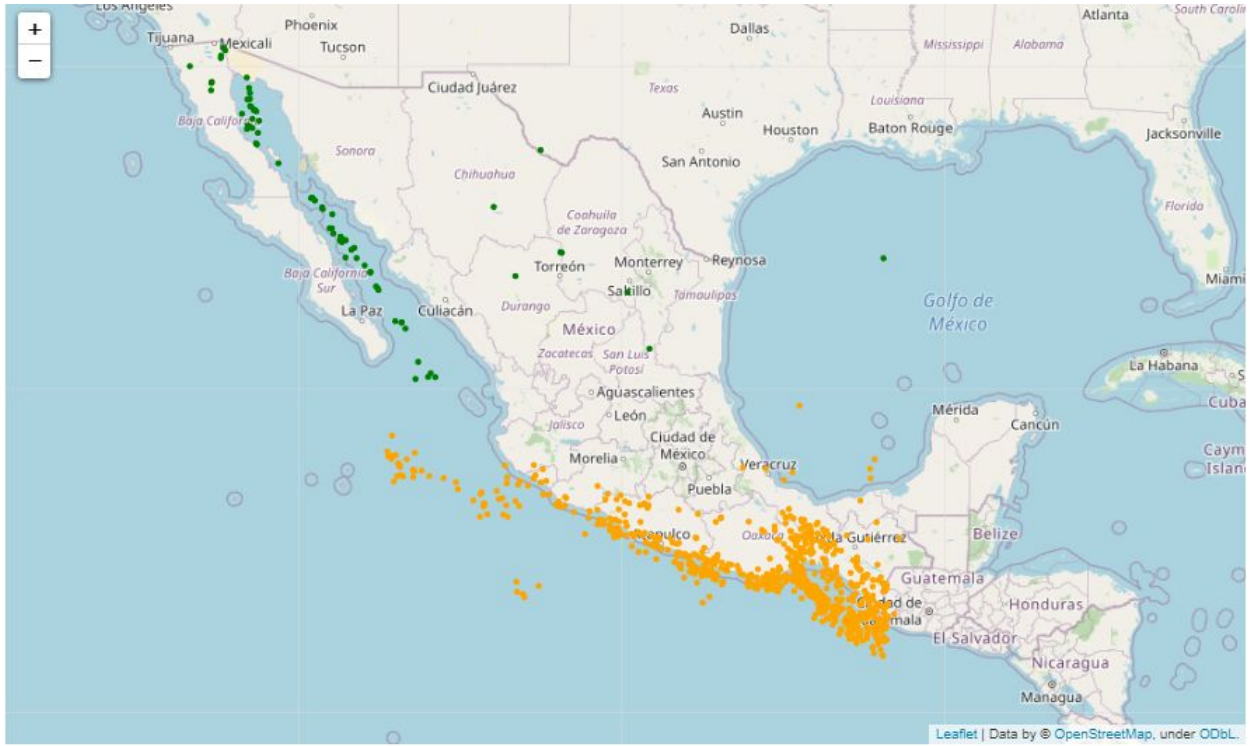
Disclaimer

I am not a geologist, a physicist, or a geophysicist.

Data -- Earthquakes

1000 most significant seismic events from 2018-2020.

Magnitudes ranging from 4.3-7.1

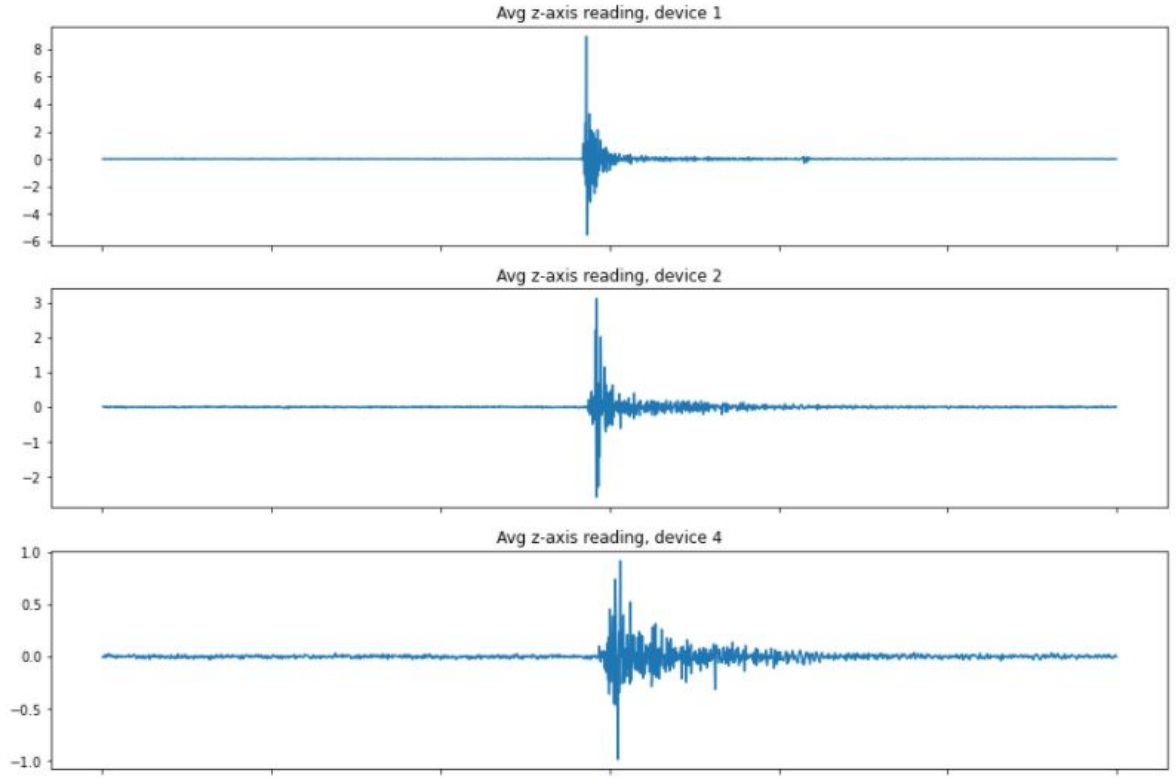


Data -- Accelerometers

28 active devices.

Record 32 measures per second on x,y,z-axis.

For each seismic event, readings were processed into 3 x 300 **waveform.**



Modeling

Waveforms were used to train a convolutional neural network in Keras to regress on latitude and and longitude of epicentre.

Kernels shaped to identify most discernible features across x-y-z axis, maximized, then convoluted again across time dimension.

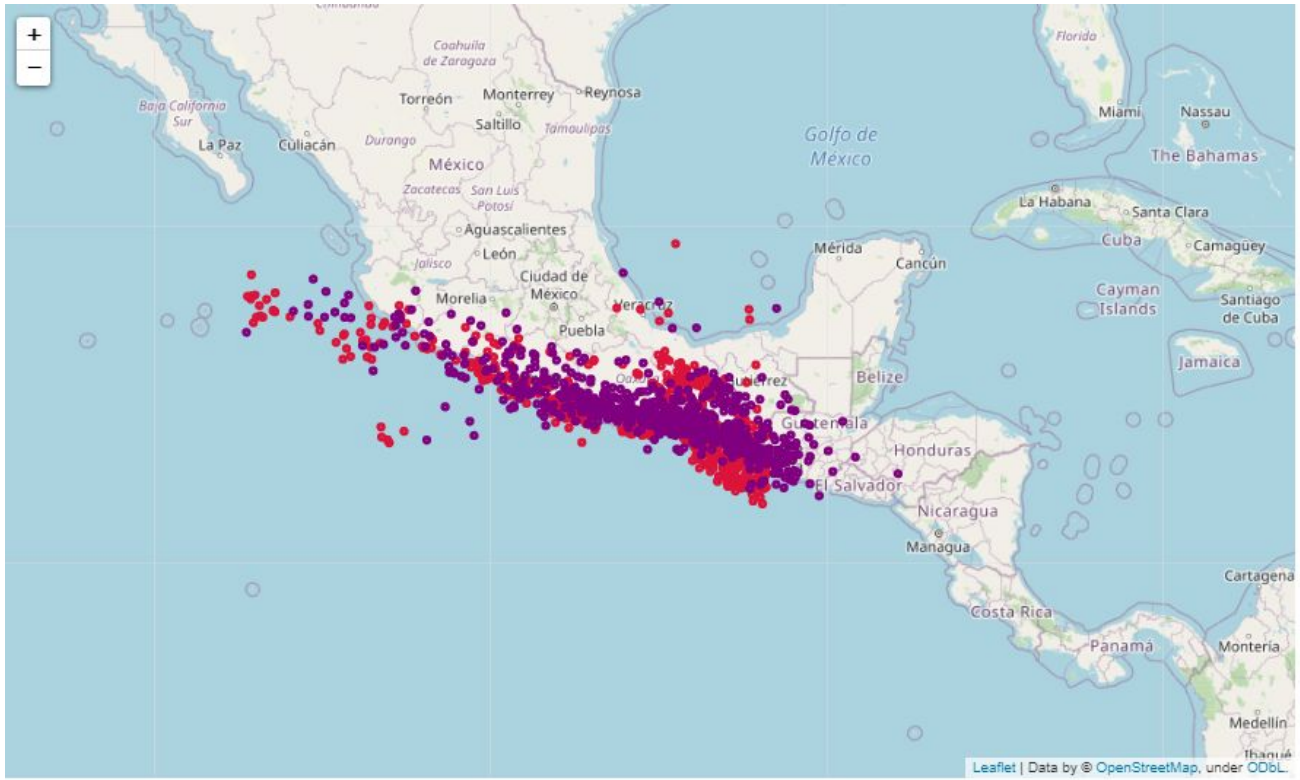
This data was then passed through a dense feed forward neural network.

```
Model: "sequential_4"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_8 (Conv2D)            (None, 3, 300, 32)        5408
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 1, 300, 32)        0
_________________________________________________________________
conv2d_9 (Conv2D)            (None, 1, 300, 64)        4160
_________________________________________________________________
flatten_4 (Flatten)          (None, 19200)             0
_________________________________________________________________
dense_8 (Dense)              (None, 2048)              39323648
_________________________________________________________________
dropout_4 (Dropout)          (None, 2048)              0
_________________________________________________________________
dense_9 (Dense)              (None, 2)                 4098
=================================================================
Total params: 39,337,314
Trainable params: 39,337,314
Non-trainable params: 0
_________________________________________________________________
```

Predictions (Training Data)

Earthquake locations in red.

Predicted locations in purple.



Predictions (Test Data)

Earthquake locations in red.

Predicted locations in purple.

Predictions (App)

Model predictions for the test set were deployed to a Heroku app.

https://ga-openeew.herokuapp.com/

Does This Model Make Good Predictions?

No!

Superficially the predictions seem reasonable.

However, numerical metrics indicate negligible predictive ability.

Many networks and modeling formulations were tried, but none produced conclusive results.

What Went Wrong?

Earthquake epicentres are conventionally located by triangulating data from seismographs.

Research literature indicates that accelerometers are very poor for measuring distance.

- Accelerometers measure….acceleration (not offset). This doesn't convert to a distance measurement readily, or allow direct construction of seismic wave.
- Strongly influenced by local geology in ways that don't generalize across a large area.
- Significantly less precise than lab grade seismographs.

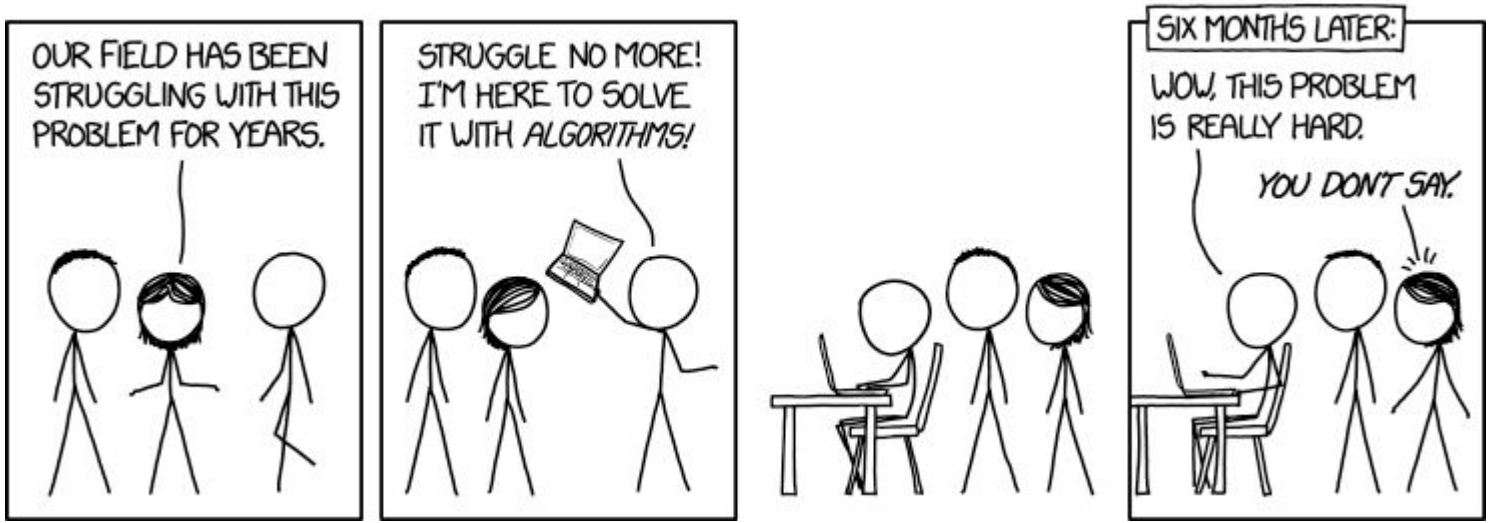
What Went Wrong?

The majority of seismic events analyzed were too small in magnitude to be effectively detected.

- Of the 928 earthquakes analyzed, only 27 classified as “strong” (magnitude >= 5.5) by the SSN.
- Most of the seismic events had magnitude near 4. These are barely perceptible to humans.
- Majority of earthquakes occured offshore.
- Many of these produced extremely low or negligible accelerometer readings.

I probably should have done more EDA.

What Went Wrong?



https://xkcd.com/1831/

Further Opportunities

Using convolutional neural networks to analyze waveform shape seems plausible.

I would like to try this methodology out on seismograph readings and see if it performs better.

OpenEEW data is time based data primarily used for warning systems, not for geolocation. Time series analysis to identify prediction threshold would be a promising project.

Questions?