

INGV – Volcanic Eruption Prediction

By Josh Webster



Problem: Volcanic Eruptions often come as a surprise to the scientific community and can cause excessive damage to communities near the eruption if not warned promptly

Data/Task: The goal of this project is to produce a model that can predict the eruption time of a volcano using these readings. If successful, scientists would be able to warn communities faster and reduce casualties from volcanic eruptions. The data comes from 10 seismic sensors placed around the surface of a volcano taking readings for 10 minutes.

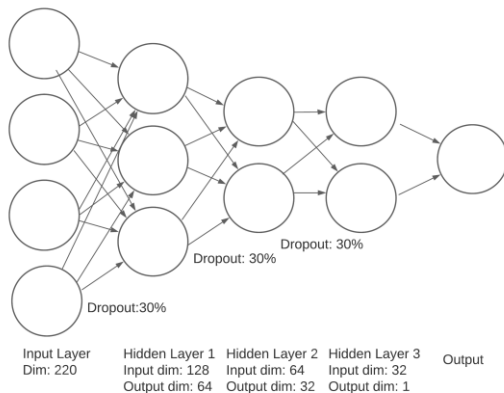
Approach:

Data Preprocessing:

- Dimensions per label is [4431, 10, 60001]
- Reduce dimension to [4421, 220]
 - Take different statistics of the sensor data
 - Sum, mean, std, quartiles,....
- Regularize the new data
- 80/20 Train/Test split

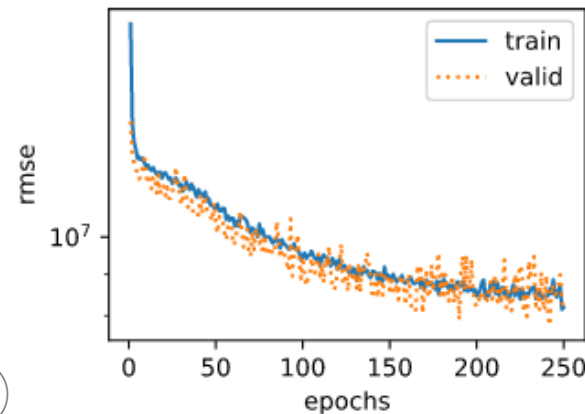
Deep Learning Regression Model:

- Multilayer Perceptron
- Epochs 250
- Learning Rate 0.05
- Adam optimizer
- 3 hidden layers
- Batch size 32
- RMSE Loss



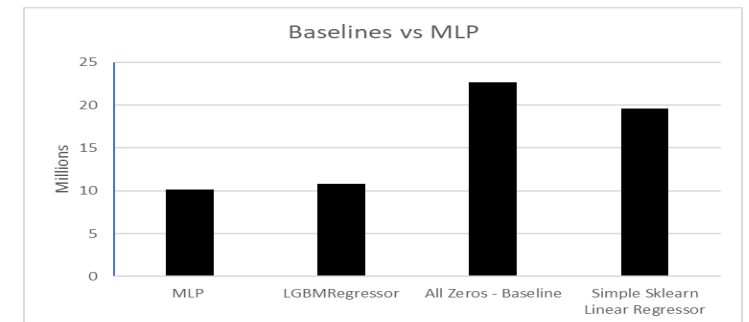
Results:

- Train loss reduced by 49.8%
- Val loss reduced by 42.6%



Analysis:

- Compared the competition score of MLP to several Baselines
- MLP – 10151414
- LGBM Regressor – 10832907
- Sklearn Linear Regressor – 19608030
- All zero submission - 22658131



Conclusions:

- MLP does not improve the baselines enough to be the best model for this problem
- Ranked 286 out of 356