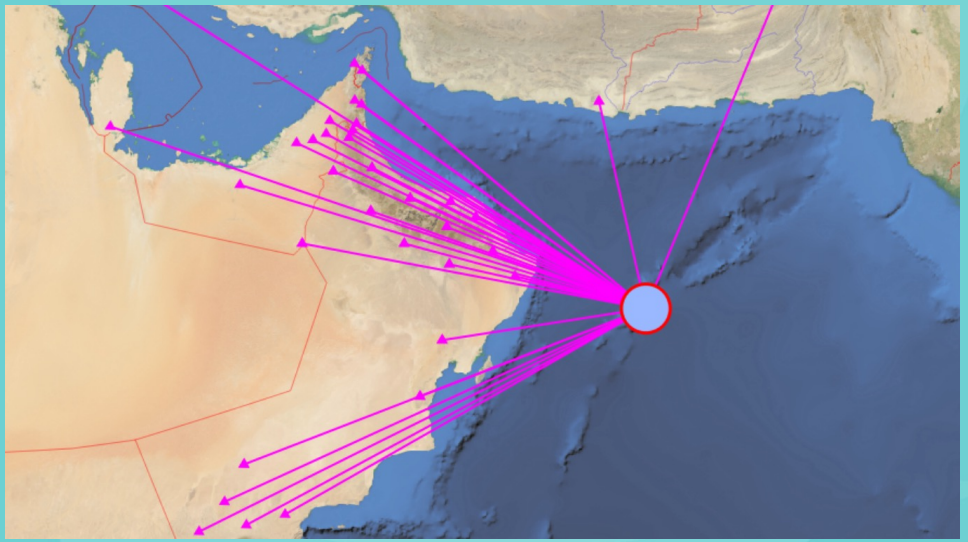
Predicting the Tsunami Behavior from Historical Data

Nouha osama Alsaleh

Table of content

Problem statement

OMAN 2011

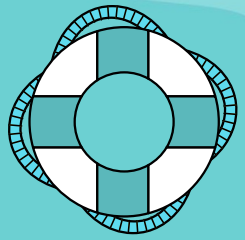


Data

- Tsunami historical events data from **(Homeland Infrastructure Foundation)**

- 2100 BC – present

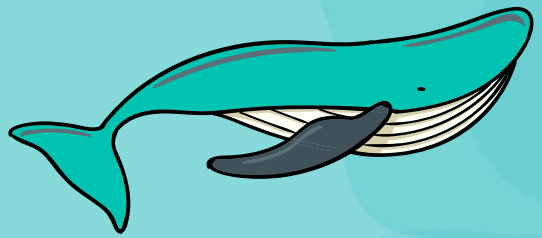
- 26,824 rows (entry) and 105 columns



Data cleaning
Replacing empty fields with NAAN

Handle missing data

Remove entries before AD

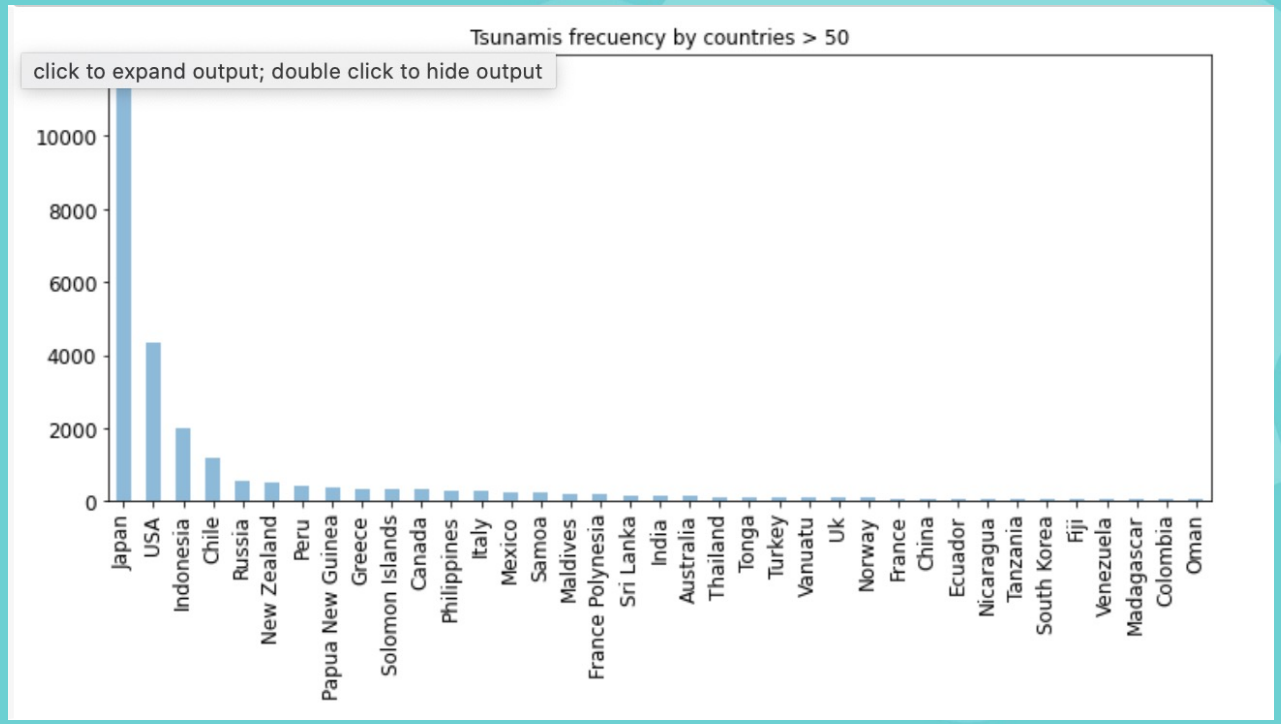


FEATURES

105 → 23

Findings

Higher tsunami frequincoes countries

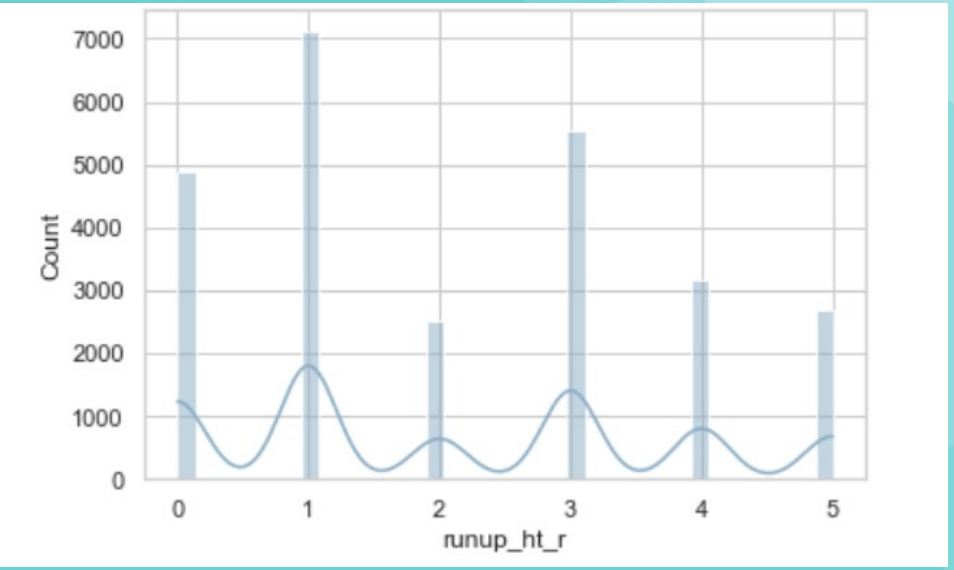


1.Japan

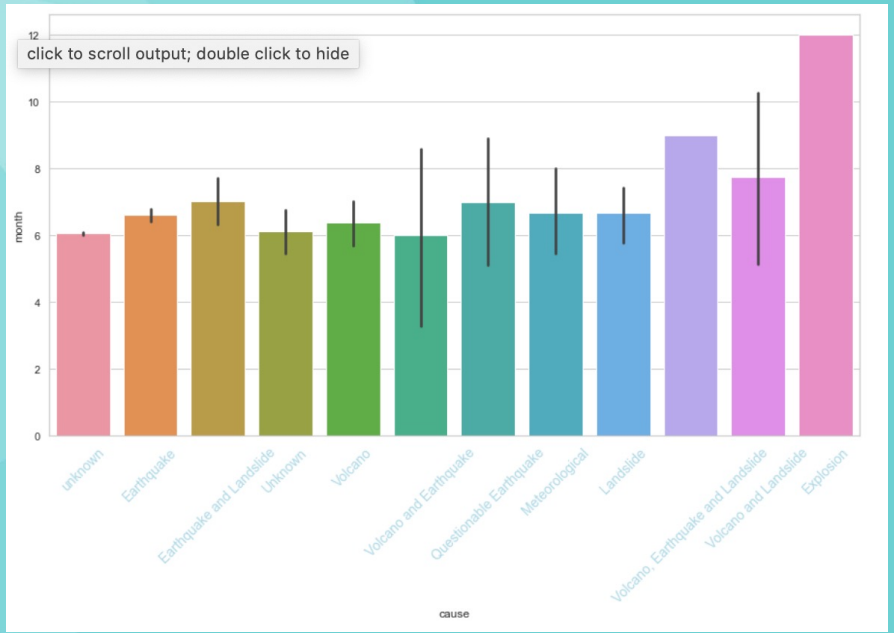
2. USA

3. Indonisia

Peaks of tsunami is mostly 1 or 3



Explosions and earthquakes are the most effective triggers for tsunami

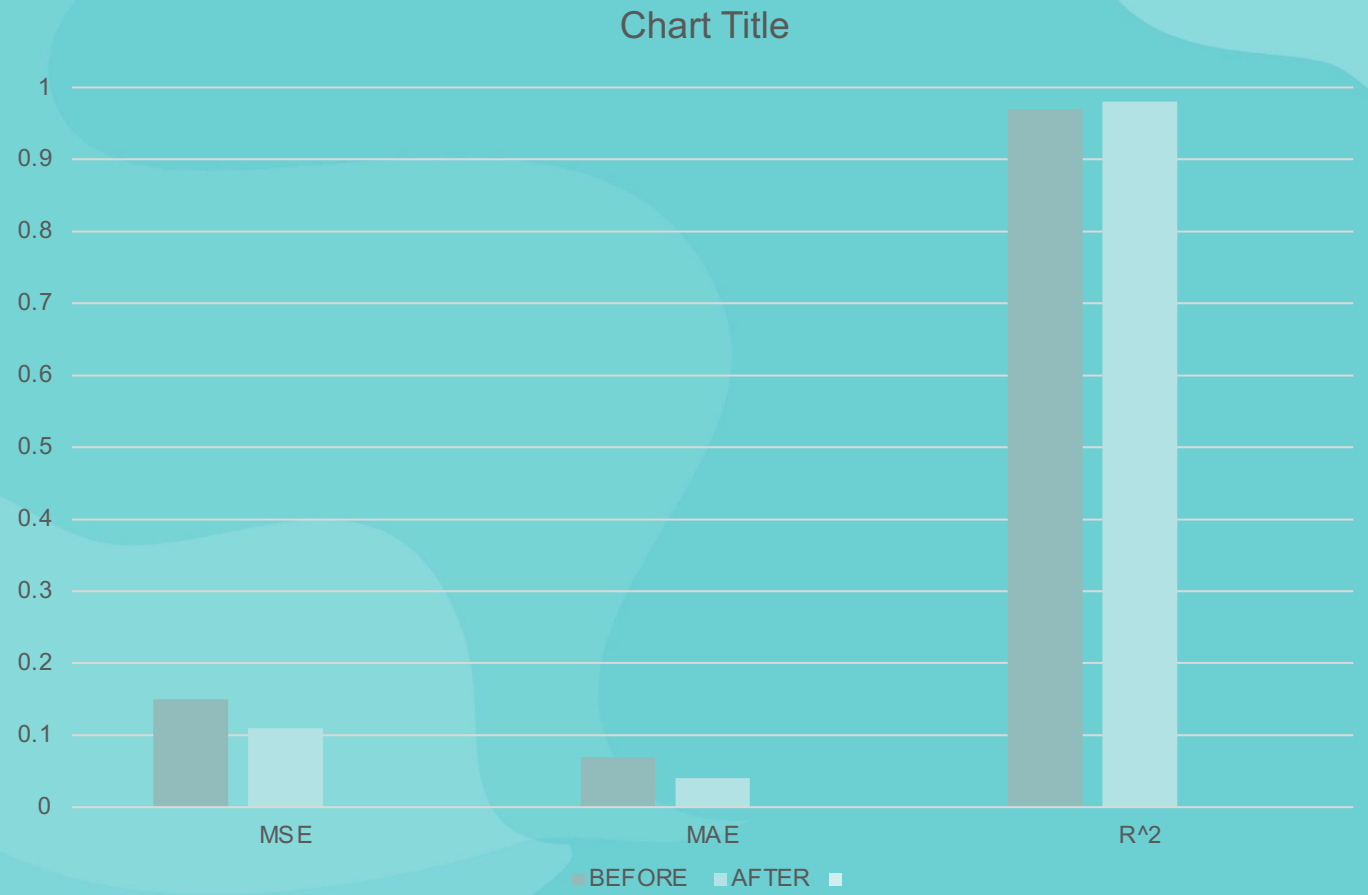


ALGORITHMS

1. linear regression

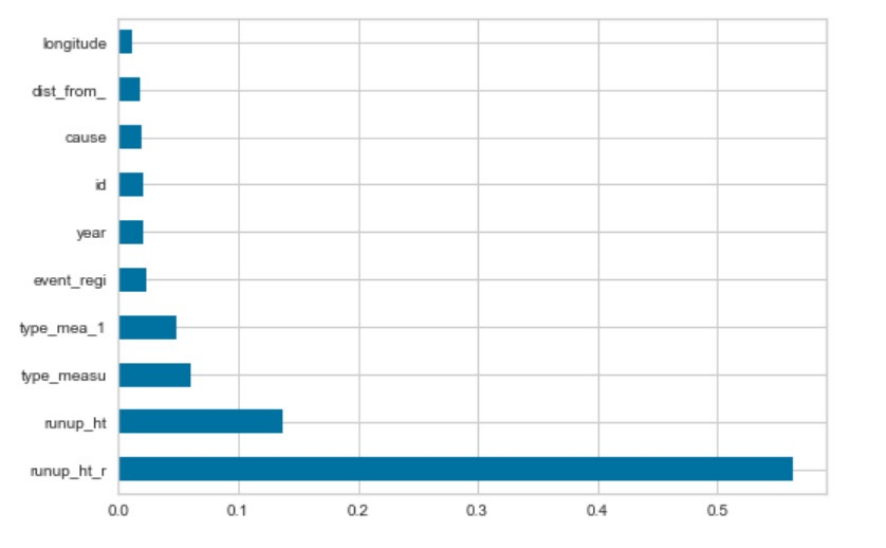
	MSE	MAE	R^2
Before feature engineering	7.88e-15	1.14e-28	1.0
After feature engineering	0.74	1.01	0.61

2. Sequential neural network

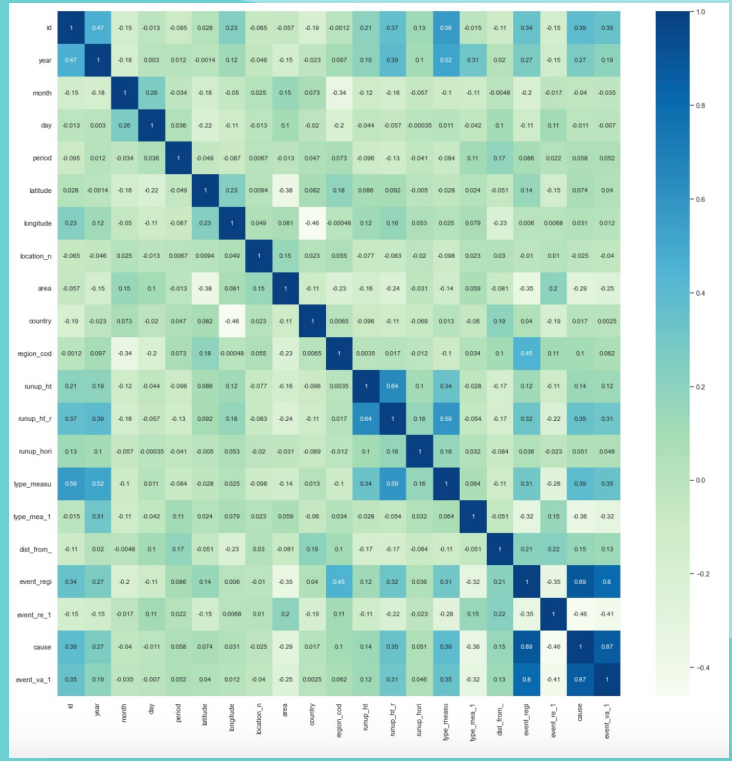


Feature engineering:

ExtraTreesClassifier



Top_corr_features



Tools

Data processing

Numpy

pandas

Visualization

Matplotlib

Seaborn

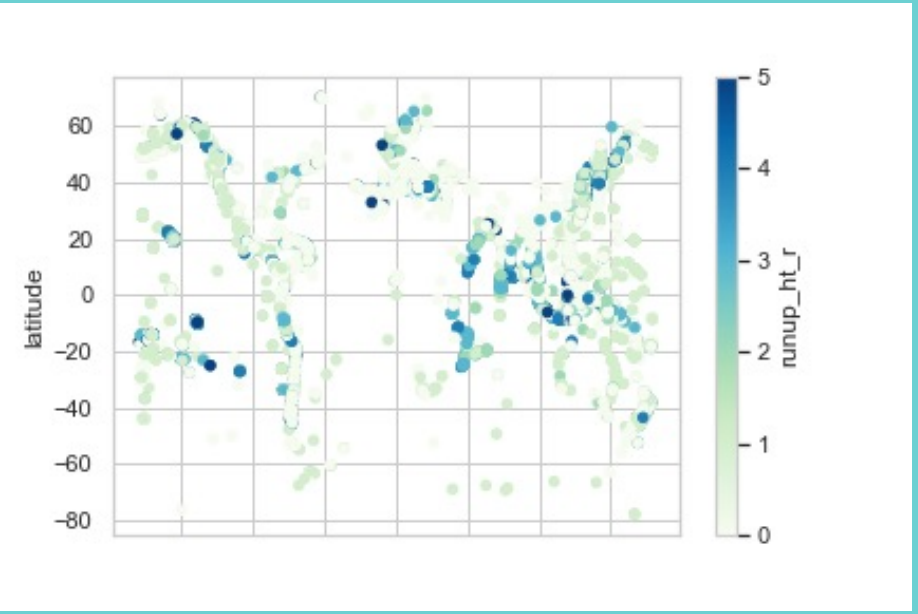
Modeling

Keras and tensorflow

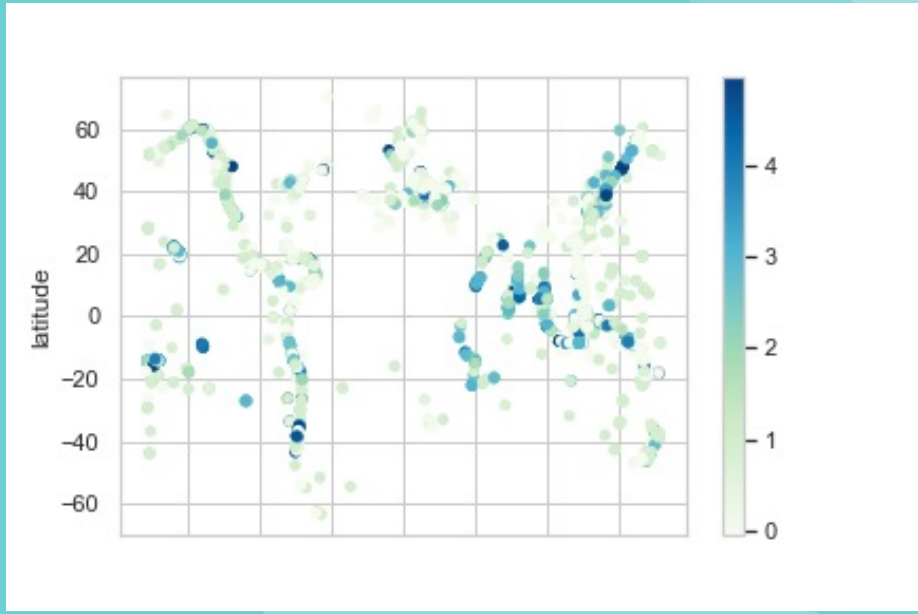
Scikit-learn

communication

original



predicted



recourses

- https://data.world/dhs/historical-tsunami-event

THANK YOU

ANY QUESTIONS?