PROJECT 7: EARTHQUAKE PREDICTION MODEL USING PYTHON

Project Title: Earthquake Prediction

Problem Definition:

The problem is to develop an earthquake prediction model using a Kaggle Dataset. The objective is to explore and understand the key features of earthquake data, visualize the data on a world map for a global overview, split the data for training and testing, and build a neural network model to predict earthquake magnitudes based on the given features.

Objective:

Earthquakes are well-studied events, with plenty of academic studies coverage, so only the basic concepts will be described here.

The majority of seismic activity happens between the movement of lithospheric plates (a.k.a. *tectonic* plates). This movement accumulates energy in the form of rock stress, and then it is suddenly released.

After the quake happens, it can be determined the location (longitude, latitude, and depth), time, and magnitude. Magnitude is the physical size of the earthquake, and the energy released can also be roughly estimated by converting the moment magnitude .

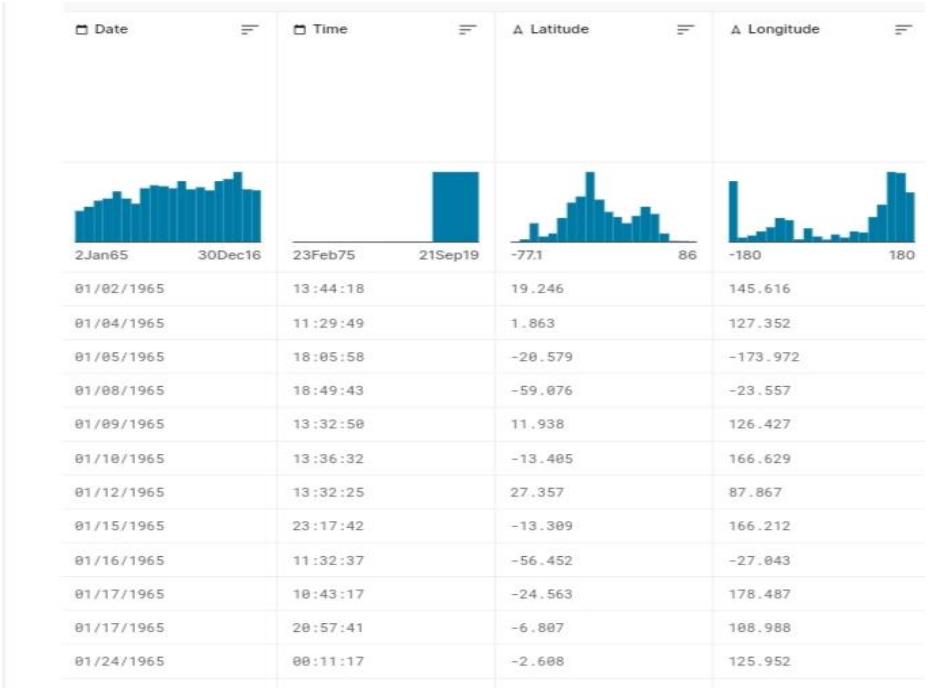
Earthquakes can cause destruction and loss of lives. Not only by the ground shaking event but also by secondary effects such as landslides, fissures, avalanches, fires and tsunamis.

Building a pre-emptive warning system can greatly increase risk management effectiveness. Being able to prepare for those rare events would help to minimise the harm caused, with actions such as local community alert and government provisioning.

Design Thinking:

1. DATA SOURCE:

Choose a suitable Kaggle dataset containing earthquake data with features like date, time, latitude, longitude, depth, and magnitude.



Date	Time	Latitude	Longitude
2Jan65 – 30Dec16	23Feb75 – 21Sep19	-77.1 – 86	-180 – 180
01/02/1965	13:44:18	19.246	145.616
01/04/1965	11:29:49	1.863	127.352
01/05/1965	18:05:58	-20.579	-173.972
01/08/1965	18:49:43	-59.076	-23.557
01/09/1965	13:32:50	11.938	126.427
01/10/1965	13:36:32	-13.405	166.629
01/12/1965	13:32:25	27.357	87.867
01/15/1965	23:17:42	-13.309	166.212
01/16/1965	11:32:37	-56.452	-27.043
01/17/1965	10:43:17	-24.563	178.487
01/17/1965	20:57:41	-6.807	108.988
01/24/1965	08:11:17	-2.608	125.952

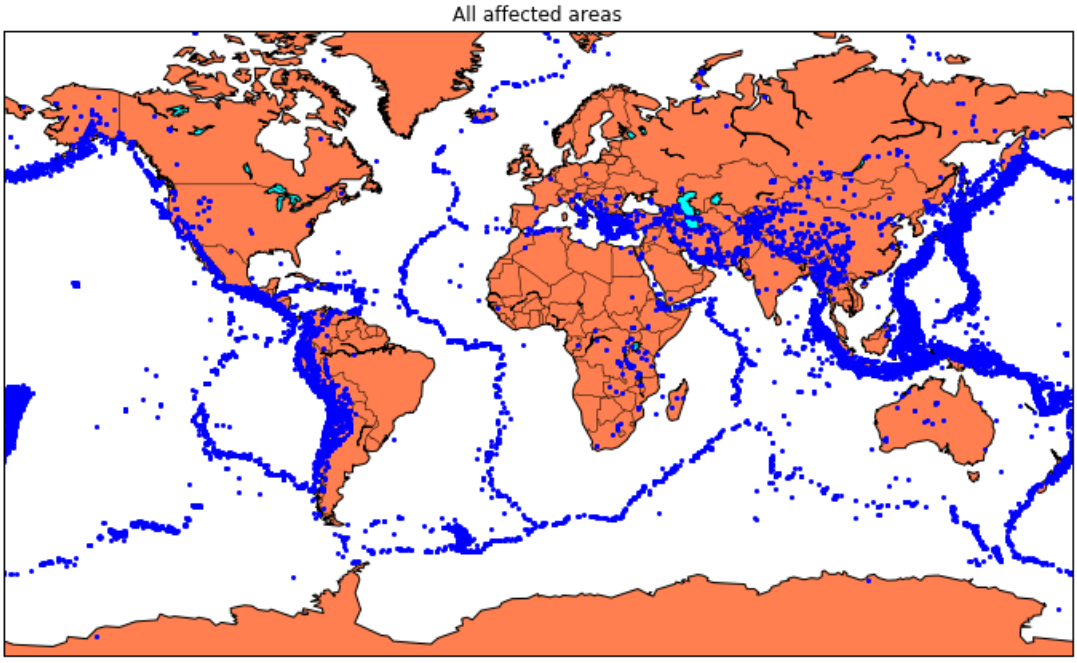
The Sample Dataset

LINK: https://www.kaggle.com/datasets/usgs/earthquake-database

This is the link to kaggle dataset which is used in this project

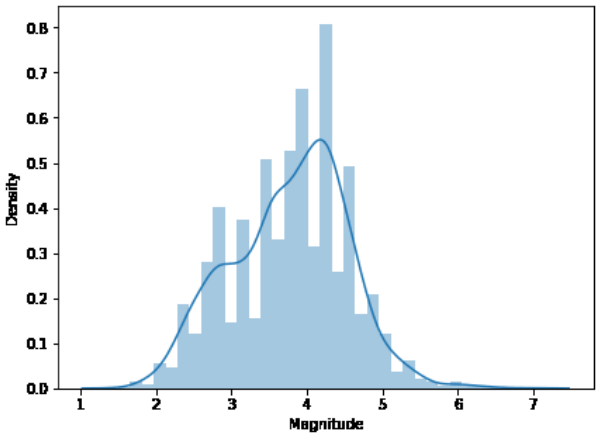
Feature Exploration:

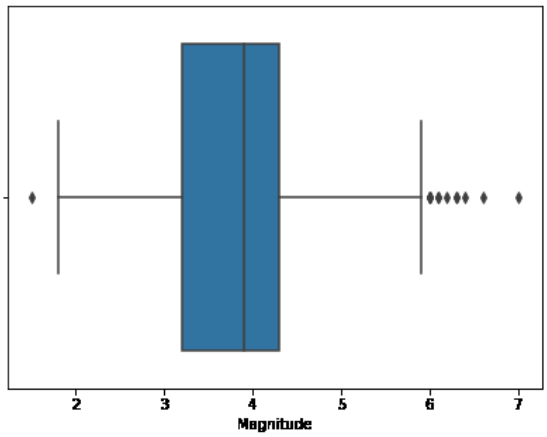
Analyze and understand the distribution, correlations, and characteristics of the key features.



Visualization:

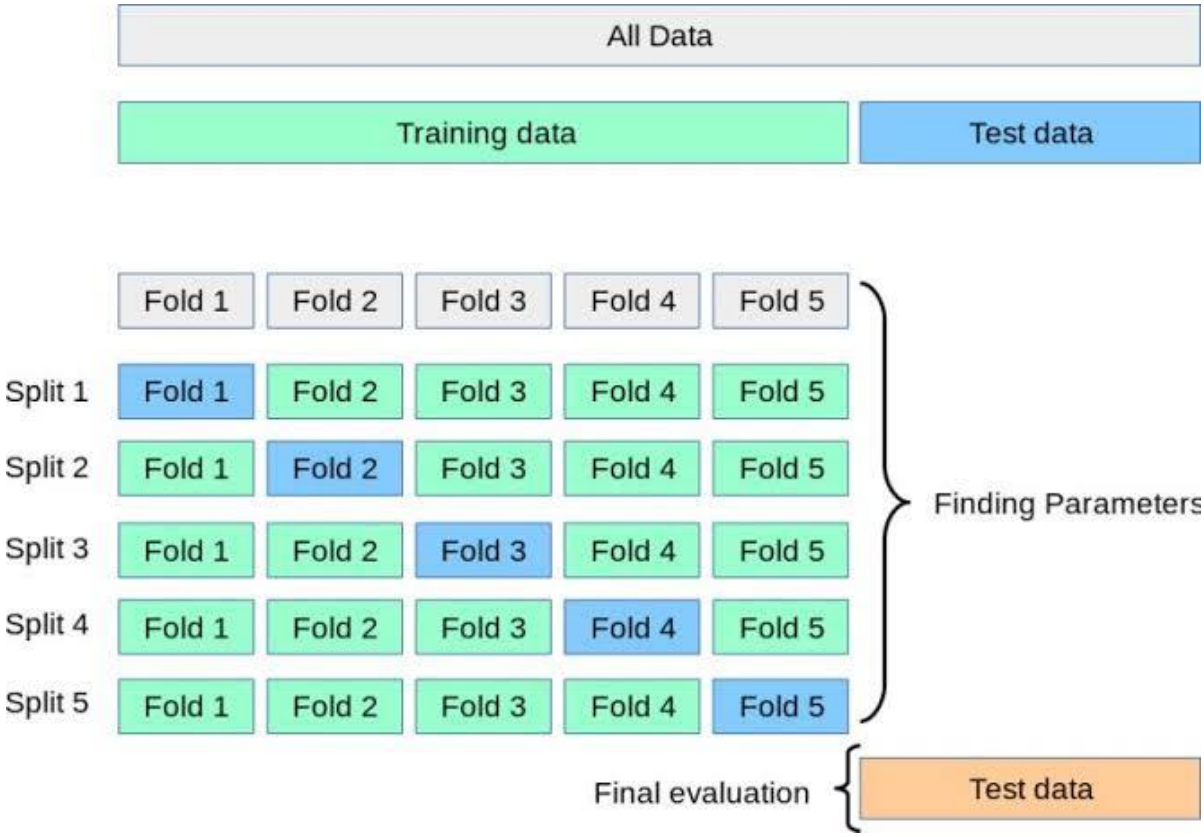
Create a world map visualization to display earthquake frequency distribution.





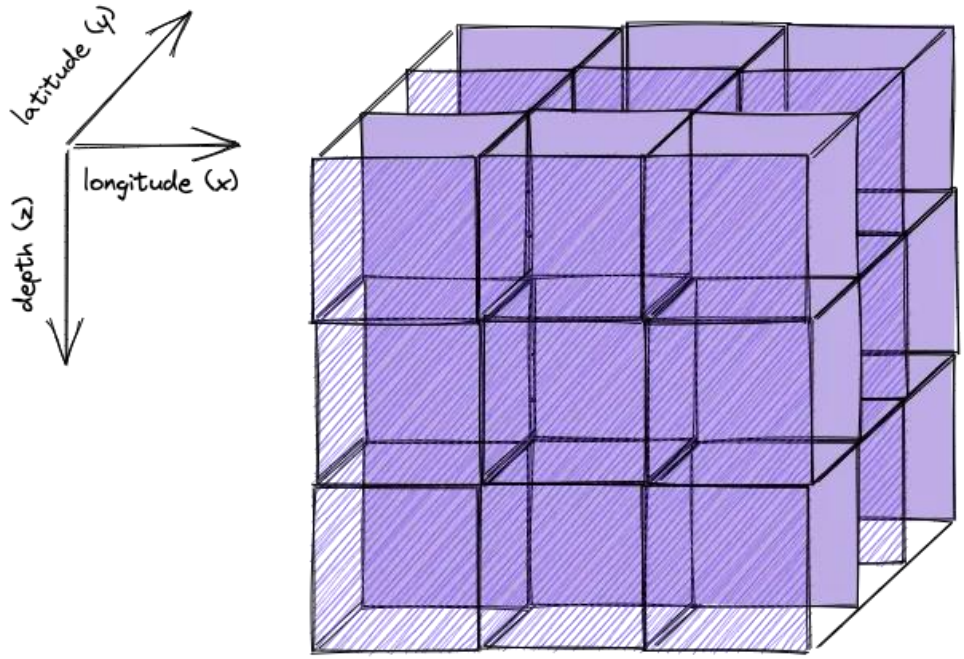
Data Spliting:

Split the dataset into a training set and a test set for model validation.



Model Development:

Build a neural network model for earthquake magnitude prediction.



Model Selection:

This explain the difference between those metrics. Precision is penalized from false alarms, and recall is penalized from missed events.

Weekly model

	Records	Balance	Events
Train	95,181 (90%)	6.95%	6,612
Test	11,084 (10%)	8.46%	938

Daily model

Records	Balance	Events
666,688 (90%)	1.72%	11,450
77,606 (10%)	2.16%	1,677

Training and Evaluation:

Train the model on the training set and evaluate its performance on the test set.

During the training phase, the model learns pattern and relationships within the training set, adjusting its parameters to optimize performance. Subsequently, he evaluation phase assesses the model's ability to generalize by testing it on the independent test set.