Earthquake prediction model using python

ARTIFICIAL INTELLIGENCE Phase-4 [development-2]

Team member

210221104038 : Thirisha M

OBJECTIVE : To continue building the earthquake prediction model by:

i.   Visualizing the data on a world map.
ii.   Splitting it into training and testing sets

Data Source: Kaggle dataset containing earthquake data with features like date, time, latitude, longitude, depth, and Magnitude

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

Here we load our dataset using pandas's read_csv() method and then preprocess our data for further process and better performance

```
import pandas as pd

# Load the dataset from a CSV file

data  = pd.read_csv('earthquake_data.csv')

Earthquake_data = Earthquake_data.dropna()  # Remove rows with missing values
```

VISUALIZING THE DATASET ON WORLD MAP

Here, all the earthquakes from the database in visualized on to the world map which shows clear representation of the locations where frequency of the earthquake will be more.

```
from mpl_toolkits.basemap import Basemap


m = Basemap(projection='mill',llcrnrlat=-80,urcrnrlat=80, llcrnrlon=-180,urcrnrlon=180,lat_ts=20,resolution='c')


Longitudes = data["Longitude"].tolist()

Latitudes = data["Latitude"].tolist()

#m = Basemap(width=12000000,height=9000000,projection='lcc',

        #resolution=None,lat_1=80.,lat_2=55,lat_0=80,lon_0=-107.)

x,y = m(Longitudes,Latitudes)

fig = plt.figure(figsize=(12,10))

plt.title("All affected areas")
```

```
m.plot(x, y, "o", markersize = 2, color = 'blue')

m.drawcoastlines()

m.fillcontinents(color='coral',lake_color='aqua')

m.drawmapboundary()

m.drawcountries()

plt.show()
```

All affected areas
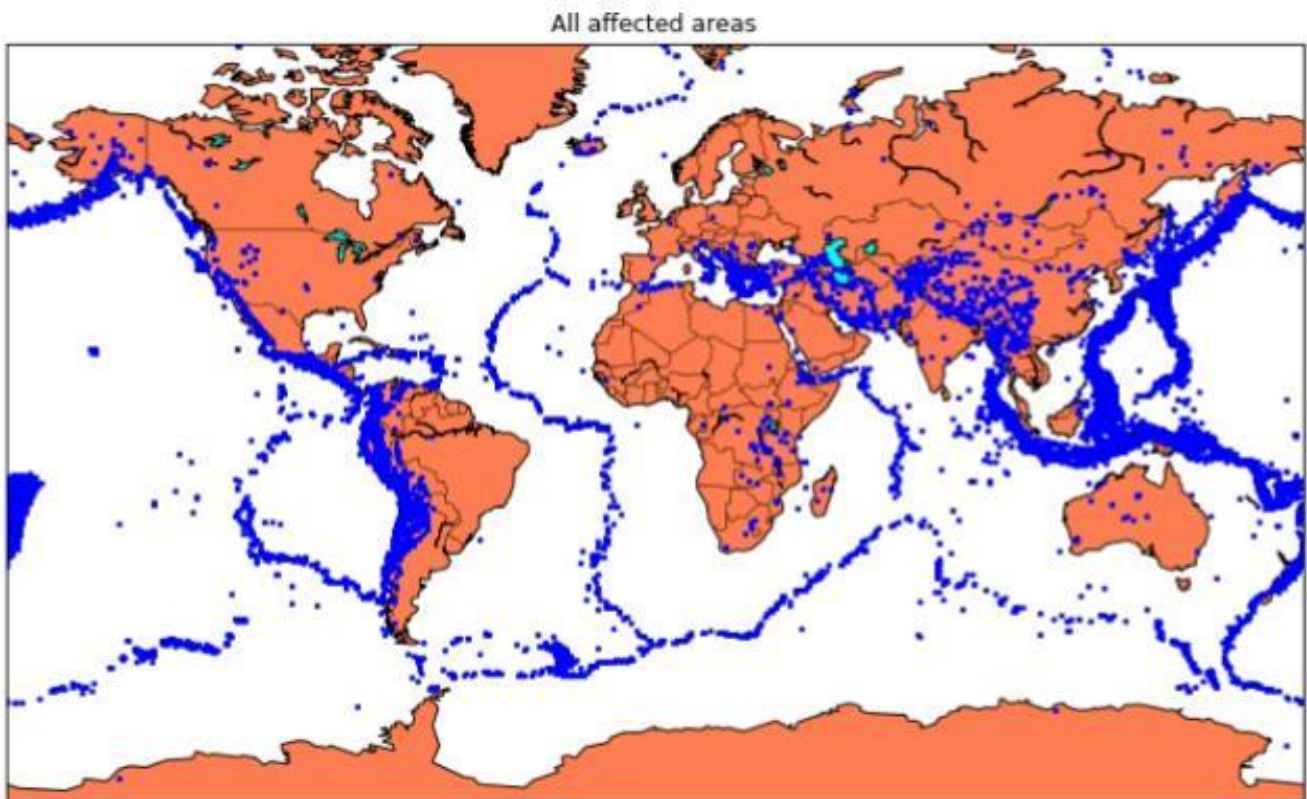
SPLITTING OF DATA

Firstly, split the data into Xs and ys which are input to the model and output of the model respectively. Here, inputs are Timestamp, Latitude and Longitude and outputs are Magnitude and Depth. Split the Xs and

ys into train and test with validation. Training dataset contains 80% and Test dataset contains 20%.

X = final_data[['Latitude', 'Longitude']]

Y = final_data[['Magnitude', 'Depth']]

from sklearn.cross_validation import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

print(X_train.shape, X_test.shape, y_train.shape, X_test.shape)

(18727, 3) (4682, 3) (18727, 2) (4682, 3)

/opt/conda/lib/python3.6/site-packages/sklearn/cross_validation.py:41: DeprecationWarning: This module was deprecated in version 0.18 in favor of the model_selection module into which all the refactored classes and functions are moved. Also note that the interface of the new CV iterators are different from that of this module. This module will be removed in 0.20.