# EARTHQUAKE PREDICTION MODEL USING PYTHON

## Phase 4 Submission Document:

**Project: Earthquake Prediction:**

## INTRODUCTION:

- Earthquakes, natural phenomena of profound consequence, continue to evoke awe and trepidation due to their unpredictable nature.

- Their ability to cause devastation underscores the urgency of understanding and predicting these geological events.

- This study embarks on a journey through data analysis, modeling, and visualization, striving to accomplish the following key objectives:

  - Data Exploration:
  - Feature Engineering:
  - Model Building:
  - Multiclass Classification:
  - Visualization:

## Data Exploration:

We initiate our exploration by delving into earthquake data sourced from reliable datasets. By meticulously examining this data, we gain insights into the temporal, spatial, and geological characteristics of earthquakes.

## Feature Engineering:

The data presents various attributes that could influence earthquake magnitudes, such as geographical coordinates (latitude and longitude) and depth. Through feature engineering, we extract the most relevant variables to feed into our predictive models.

## Model Building:

To predict earthquake magnitudes, we experiment with different machine learning models, beginning with a Linear Regression model to establish a baseline. We then advance to more sophisticated models like Support Vector Machines (SVM) to capture complex relationships within the data.

## Multiclass Classification:

Beyond magnitude prediction, we also explore multiclass classification to categorize earthquakes into different magnitude types (e.g., 'ML', 'MS', 'MB'). This not only enhances our predictive abilities but also provides insights into the diversity of seismic events.

## Visualization:

Throughout our analysis, we employ data visualization techniques to elucidate the geographical distribution of earthquakes and the performance of our models. Maps and scatter plots reveal patterns, correlations, and the efficacy of our predictions.

## Loading DataSet:

Data loading in earthquake prediction is a crucial step in the process. It involves gathering and organizing relevant information to train and test predictive models Downloading

## DataSet:

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

| date | depth | mag | place | latitude | longitude | depth_avg_22 | depth_avg_15 | depth_avg_7 | mag_avg_22 | mag_avg_15 | ma |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2020-07-14 | 6.70 | 1.58 | Oklahoma | 36.171483 | -97.718347 | 6.717727 | 6.560000 | 7.100000 | 1.352273 | 1.271333 | |
| 2020-07-14 | 7.55 | 2.07 | Oklahoma | 36.171483 | -97.718347 | 6.730000 | 6.682667 | 7.132857 | 1.372727 | 1.334667 | |
| 2020-07-14 | 7.39 | 1.89 | Oklahoma | 36.171483 | -97.718347 | 6.747727 | 6.708667 | 6.940000 | 1.396818 | 1.377333 | |
| 2020-07-15 | 7.75 | 1.48 | Oklahoma | 36.171483 | -97.718347 | 6.834545 | 6.764000 | 6.848571 | 1.383182 | 1.388667 | |
| 2020-07-15 | 7.81 | 1.50 | Oklahoma | 36.171483 | -97.718347 | 6.841364 | 6.854667 | 6.964286 | 1.404545 | 1.385333 | |

## Program :

```
import numpy as np
import pandas as pd
from sklearn import preprocessing;
from sklearn import model_selection;
from sklearn import linear_model;
import os
import datetime as dt
import matplotlib.pyplot as plt


df=pd.read_csv('database.csv')




new_column_names = ["Date(YYYY/MM/DD)",  "Time(UTC)", "Latitude(deg)",
                    "Longitude(deg)","Type","Depth(km)", "Depth Error(km)",
                    "Depth Seismic Stations(km)"," Magnitude(ergs)",
                     "Magnitude_type", "Magnitude Error",
                     "Magnitude Seismic Stations","Azimuthal Gap",
                     "Horizontal Distance", "Horizontal Error",
                     " RootMeanSquare ","ID ", "Source",
                     "Location Source", "Magnitude Source","Status"]




df.columns = new_column_names
ts = pd.to_datetime(df["Date(YYYY/MM/DD)"])
df = df.drop(["Date(YYYY/MM/DD)", "Time(UTC)"], axis=1)
df.index = ts
display(df)
```

```
                  Latitude(deg)  Longitude(deg)        Type  Depth(km)  \
Date(YYYY/MM/DD)
1965-01-02              19.2460        145.6160  Earthquake     131.60
1965-01-04               1.8630        127.3520  Earthquake      80.00
1965-01-05             -20.5790       -173.9720  Earthquake      20.00
1965-01-08             -59.0760        -23.5570  Earthquake      15.00
1965-01-09              11.9380        126.4270  Earthquake      15.00
...                         ...             ...         ...        ...
2016-12-28              38.3917       -118.8941  Earthquake      12.30
2016-12-28              38.3777       -118.8957  Earthquake       8.80
2016-12-28              36.9179        140.4262  Earthquake      10.00
2016-12-29              -9.0283        118.6639  Earthquake      79.00
2016-12-30              37.3973        141.4103  Earthquake      11.94

                  Depth Error(km)  Depth Seismic Stations(km)  \
Date(YYYY/MM/DD)
1965-01-02                    NaN                         NaN
1965-01-04                    NaN                         NaN
1965-01-05                    NaN                         NaN
1965-01-08                    NaN                         NaN
1965-01-09                    NaN                         NaN
...                           ...                         ...
2016-12-28                    1.2                        40.0
2016-12-28                    2.0                        33.0
2016-12-28                    1.8                         NaN
2016-12-29                    1.8                         NaN
2016-12-30                    2.2                         NaN

                   Magnitude(ergs) Magnitude_type  Magnitude Error  \
Date(YYYY/MM/DD)
1965-01-02                     6.0             MW              NaN
1965-01-04                     5.8             MW              NaN
1965-01-05                     6.2             MW              NaN
1965-01-08                     5.8             MW              NaN
1965-01-09                     5.8             MW              NaN
...                            ...            ...              ...
2016-12-28                     5.6             ML            0.320
2016-12-28                     5.5             ML            0.260
2016-12-28                     5.9            MWW              NaN
2016-12-29                     6.3            MWW              NaN
2016-12-30                     5.5             MB            0.029

                  Magnitude Seismic Stations  Azimuthal Gap  \
Date(YYYY/MM/DD)
```

```
1965-01-02                                   NaN           NaN
1965-01-04                                   NaN           NaN
1965-01-05                                   NaN           NaN
1965-01-08                                   NaN           NaN
1965-01-09                                   NaN           NaN
...                                          ...           ...
2016-12-28                                  18.0         42.47
2016-12-28                                  18.0         48.58
2016-12-28                                   NaN         91.00
2016-12-29                                   NaN         26.00
2016-12-30                                 428.0         97.00

                  Horizontal Distance  Horizontal Error   RootMeanSquare   \
Date(YYYY/MM/DD)
1965-01-02                        NaN               NaN              NaN
1965-01-04                        NaN               NaN              NaN
1965-01-05                        NaN               NaN              NaN
1965-01-08                        NaN               NaN              NaN
1965-01-09                        NaN               NaN              NaN
...                               ...               ...              ...
2016-12-28                      0.120               NaN           0.1898
2016-12-28                      0.129               NaN           0.2187
2016-12-28                      0.992               4.8           1.5200
2016-12-29                      3.553               6.0           1.4300
2016-12-30                      0.681               4.5           0.9100

                           ID   Source Location Source Magnitude Source  \
Date(YYYY/MM/DD)
1965-01-02        ISCGEM860706  ISCGEM          ISCGEM           ISCGEM
1965-01-04        ISCGEM860737  ISCGEM          ISCGEM           ISCGEM
1965-01-05        ISCGEM860762  ISCGEM          ISCGEM           ISCGEM
1965-01-08        ISCGEM860856  ISCGEM          ISCGEM           ISCGEM
1965-01-09        ISCGEM860890  ISCGEM          ISCGEM           ISCGEM
...                        ...     ...             ...              ...
2016-12-28          NN00570710      NN              NN               NN
2016-12-28          NN00570744      NN              NN               NN
2016-12-28          US10007NAF      US              US               US
2016-12-29          US10007NL0      US              US               US
2016-12-30          US10007NTD      US              US               US

                     Status
Date(YYYY/MM/DD)
1965-01-02        Automatic
1965-01-04        Automatic
1965-01-05        Automatic
1965-01-08        Automatic
1965-01-09        Automatic
...                     ...
2016-12-28         Reviewed
2016-12-28         Reviewed
```

```
2016-12-28         Reviewed
2016-12-29         Reviewed
2016-12-30         Reviewed

[23412 rows x 19 columns]



df.info()
df.head()



<class 'pandas.core.frame.DataFrame'>
Index: 23412 entries, 1965-01-02 00:00:00 to 2016-12-30 00:00:00
Data columns (total 19 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   Latitude(deg)                 23412 non-null  float64
 1   Longitude(deg)                23412 non-null  float64
 2   Type                          23412 non-null  object
 3   Depth(km)                     23412 non-null  float64
 4   Depth Error(km)               4461 non-null   float64
 5   Depth Seismic Stations(km)    7097 non-null   float64
 6    Magnitude(ergs)              23412 non-null  float64
 7   Magnitude_type                23409 non-null  object
 8   Magnitude Error               327 non-null    float64
 9   Magnitude Seismic Stations    2564 non-null   float64
 10  Azimuthal Gap                 7299 non-null   float64
 11  Horizontal Distance           1604 non-null   float64
 12  Horizontal Error              1156 non-null   float64
 13   RootMeanSquare               17352 non-null  float64
 14  ID                            23412 non-null  object
 15  Source                        23412 non-null  object
 16  Location Source               23412 non-null  object
 17  Magnitude Source              23412 non-null  object
 18  Status                        23412 non-null  object
dtypes: float64(12), object(7)
memory usage: 3.6+ MB

                     Latitude(deg)  Longitude(deg)        Type  Depth(km)  \
Date(YYYY/MM/DD)
1965-01-02 00:00:00         19.246         145.616  Earthquake      131.6
1965-01-04 00:00:00          1.863         127.352  Earthquake       80.0
1965-01-05 00:00:00        -20.579        -173.972  Earthquake       20.0
1965-01-08 00:00:00        -59.076         -23.557  Earthquake       15.0
1965-01-09 00:00:00         11.938         126.427  Earthquake       15.0

                     Depth Error(km)  Depth Seismic Stations(km)  \
Date(YYYY/MM/DD)
1965-01-02 00:00:00              NaN                         NaN
```

```
1965-01-04 00:00:00             NaN                          NaN
1965-01-05 00:00:00             NaN                          NaN
1965-01-08 00:00:00             NaN                          NaN
1965-01-09 00:00:00             NaN                          NaN

                      Magnitude(ergs) Magnitude_type  Magnitude Error  \
Date(YYYY/MM/DD)
1965-01-02 00:00:00              6.0             MW              NaN
1965-01-04 00:00:00              5.8             MW              NaN
1965-01-05 00:00:00              6.2             MW              NaN
1965-01-08 00:00:00              5.8             MW              NaN
1965-01-09 00:00:00              5.8             MW              NaN

                     Magnitude Seismic Stations  Azimuthal Gap  \
Date(YYYY/MM/DD)
1965-01-02 00:00:00                         NaN            NaN
1965-01-04 00:00:00                         NaN            NaN
1965-01-05 00:00:00                         NaN            NaN
1965-01-08 00:00:00                         NaN            NaN
1965-01-09 00:00:00                         NaN            NaN

                     Horizontal Distance  Horizontal Error   RootMeanSquare
\
Date(YYYY/MM/DD)
1965-01-02 00:00:00                  NaN               NaN               NaN
1965-01-04 00:00:00                  NaN               NaN               NaN
1965-01-05 00:00:00                  NaN               NaN               NaN
1965-01-08 00:00:00                  NaN               NaN               NaN
1965-01-09 00:00:00                  NaN               NaN               NaN

                              ID   Source Location Source Magnitude Source  \
Date(YYYY/MM/DD)
1965-01-02 00:00:00  ISCGEM860706  ISCGEM          ISCGEM           ISCGEM
1965-01-04 00:00:00  ISCGEM860737  ISCGEM          ISCGEM           ISCGEM
1965-01-05 00:00:00  ISCGEM860762  ISCGEM          ISCGEM           ISCGEM
1965-01-08 00:00:00  ISCGEM860856  ISCGEM          ISCGEM           ISCGEM
1965-01-09 00:00:00  ISCGEM860890  ISCGEM          ISCGEM           ISCGEM

                        Status
Date(YYYY/MM/DD)
1965-01-02 00:00:00  Automatic
1965-01-04 00:00:00  Automatic
1965-01-05 00:00:00  Automatic
1965-01-08 00:00:00  Automatic
1965-01-09 00:00:00  Automatic
```

```
df.shape
```

```
(23412, 19)
```

```
df.describe()

       Latitude(deg)  Longitude(deg)     Depth(km)  Depth Error(km)  \
count   23412.000000    23412.000000  23412.000000      4461.000000   
mean        1.679033       39.639961     70.767911         4.993115   
std        30.113183      125.511959    122.651898         4.875184   
min       -77.080000     -179.997000     -1.100000         0.000000   
25%       -18.653000      -76.349750     14.522500         1.800000   
50%        -3.568500      103.982000     33.000000         3.500000   
75%        26.190750      145.026250     54.000000         6.300000   
max        86.005000      179.998000    700.000000        91.295000   

       Depth Seismic Stations(km)  Magnitude(ergs)  Magnitude Error  \
count                 7097.000000     23412.000000       327.000000   
mean                   275.364098         5.882531         0.071820   
std                    162.141631         0.423066         0.051466   
min                      0.000000         5.500000         0.000000   
25%                    146.000000         5.600000         0.046000   
50%                    255.000000         5.700000         0.059000   
75%                    384.000000         6.000000         0.075500   
max                    934.000000         9.100000         0.410000   

       Magnitude Seismic Stations  Azimuthal Gap  Horizontal Distance  \
count                 2564.000000    7299.000000          1604.000000   
mean                    48.944618      44.163532             3.992660   
std                     62.943106      32.141486             5.377262   
min                      0.000000       0.000000             0.004505   
25%                     10.000000      24.100000             0.968750   
50%                     28.000000      36.000000             2.319500   
75%                     66.000000      54.000000             4.724500   
max                    821.000000     360.000000            37.874000   

       Horizontal Error  RootMeanSquare  
count       1156.000000    17352.000000  
mean           7.662759        1.022784  
std           10.430396        0.188545  
min            0.085000        0.000000  
25%            5.300000        0.900000  
50%            6.700000        1.000000  
75%            8.100000        1.130000  
max           99.000000        3.440000  

df.info()

<class 'pandas.core.frame.DataFrame'>
Index: 23412 entries, 1965-01-02 00:00:00 to 2016-12-30 00:00:00
Data columns (total 19 columns):
 #   Column                      Non-Null Count  Dtype  
---  ------                      --------------  -----  
 0   Latitude(deg)               23412 non-null  float64
```

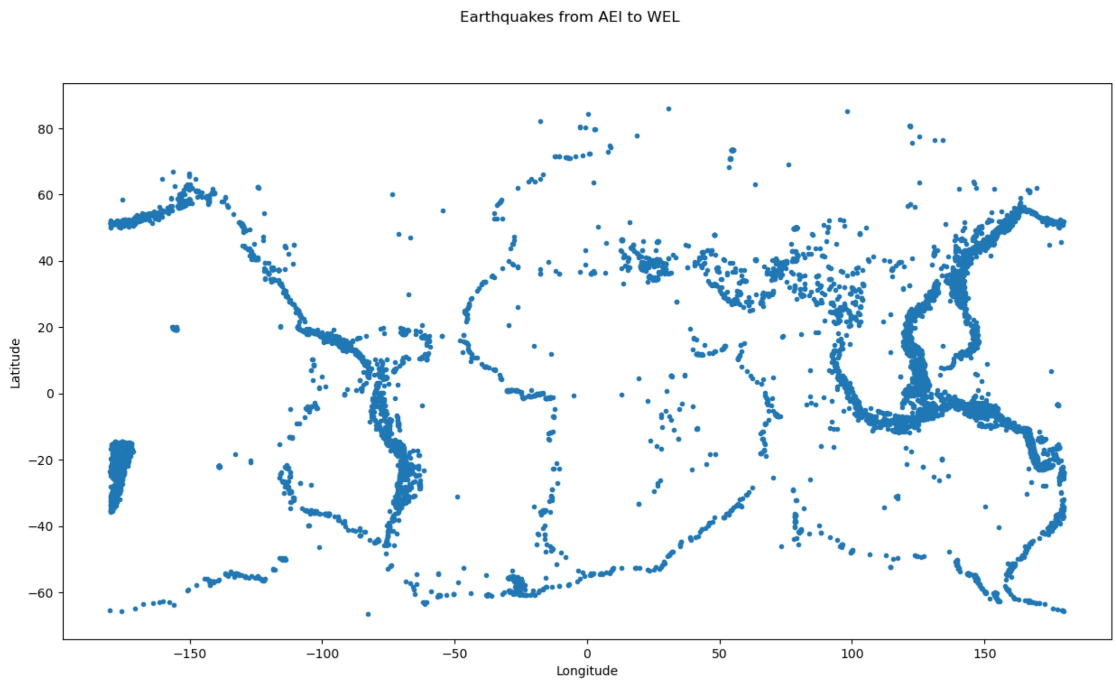
```
 1   Longitude(deg)              23412 non-null  float64
 2   Type                        23412 non-null  object
 3   Depth(km)                   23412 non-null  float64
 4   Depth Error(km)             4461 non-null   float64
 5   Depth Seismic Stations(km)  7097 non-null   float64
 6    Magnitude(ergs)            23412 non-null  float64
 7   Magnitude_type              23409 non-null  object
 8   Magnitude Error             327 non-null    float64
 9   Magnitude Seismic Stations  2564 non-null   float64
 10  Azimuthal Gap               7299 non-null   float64
 11  Horizontal Distance         1604 non-null   float64
 12  Horizontal Error            1156 non-null   float64
 13   RootMeanSquare             17352 non-null  float64
 14  ID                          23412 non-null  object
 15  Source                      23412 non-null  object
 16  Location Source             23412 non-null  object
 17  Magnitude Source            23412 non-null  object
 18  Status                      23412 non-null  object
dtypes: float64(12), object(7)
memory usage: 3.6+ MB
```

```
df.isnull().sum()
```

```
Latitude(deg)                   0
Longitude(deg)                  0
Type                            0
Depth(km)                       0
Depth Error(km)             18951
Depth Seismic Stations(km)  16315
 Magnitude(ergs)                0
Magnitude_type                  3
Magnitude Error             23085
Magnitude Seismic Stations  20848
Azimuthal Gap               16113
Horizontal Distance         21808
Horizontal Error            22256
 RootMeanSquare              6060
ID                              0
Source                          0
Location Source                 0
Magnitude Source                0
Status                          0
dtype: int64
```

```
rounding_factor = 10
fig, ax = plt.subplots(figsize=(15,8))
# latitude and longitude of earthquake site of top 10500 samples.
plt.plot(np.round(df['Longitude(deg)'].head(10500),rounding_factor),
         np.round(df['Latitude(deg)'].head(10500),rounding_factor),
         linestyle='none', marker='.')
```

```
plt.suptitle('Earthquakes from ' + str(np.min(df['Location Source']))[:20] +
' to ' + str(np.max(df['Location Source']))[:20])
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.show()
```



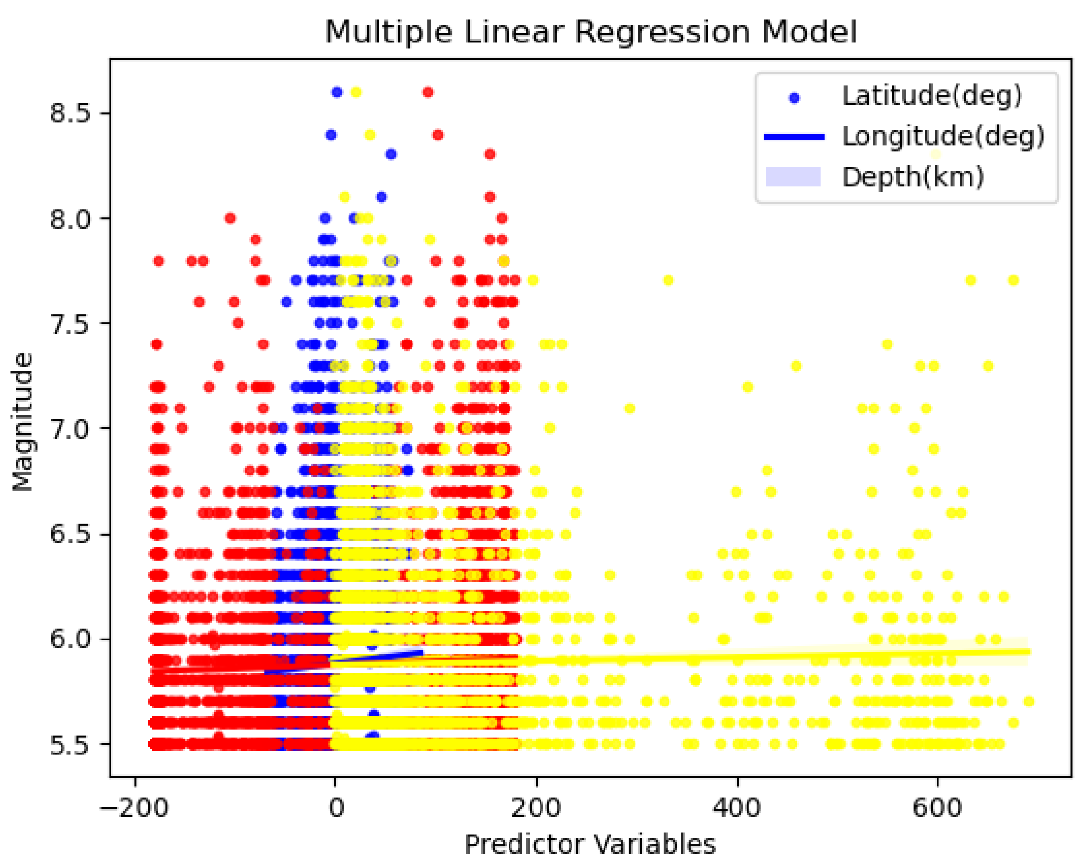

```
df = df.sort_values('Location Source', ascending=True)
#Date extraction
df['Date'] = df['Location Source'].str[0:10]
df.head()
```

```
                     Latitude(deg)  Longitude(deg)        Type  Depth(km)  \
Date(YYYY/MM/DD)
1992-05-12 00:00:00         59.691        -153.482  Earthquake      138.8
1991-05-01 00:00:00         62.476        -151.413  Earthquake      114.2
1995-12-30 00:00:00         63.212        -150.605  Earthquake      137.3
1995-10-06 00:00:00         65.170        -148.565  Earthquake        9.1
1993-11-20 00:00:00         60.025        -153.003  Earthquake      116.3

                     Depth Error(km)  Depth Seismic Stations(km)  \
Date(YYYY/MM/DD)
1992-05-12 00:00:00              NaN                          NaN
1991-05-01 00:00:00              NaN                          NaN
1995-12-30 00:00:00              NaN                          NaN
1995-10-06 00:00:00              NaN                          NaN
1993-11-20 00:00:00              NaN                          NaN

                      Magnitude(ergs) Magnitude_type  Magnitude Error  \
Date(YYYY/MM/DD)
```

```
1992-05-12 00:00:00                5.6             MW              NaN
1991-05-01 00:00:00                6.3             MW              NaN
1995-12-30 00:00:00                5.7             MW              NaN
1995-10-06 00:00:00                5.8             MS              NaN
1993-11-20 00:00:00                5.9             MW              NaN

                     Magnitude Seismic Stations  Azimuthal Gap  \
Date(YYYY/MM/DD)
1992-05-12 00:00:00                         NaN            NaN
1991-05-01 00:00:00                         NaN            NaN
1995-12-30 00:00:00                         NaN            NaN
1995-10-06 00:00:00                        64.0            NaN
1993-11-20 00:00:00                         NaN            NaN

                     Horizontal Distance  Horizontal Error   RootMeanSquare
\
Date(YYYY/MM/DD)
1992-05-12 00:00:00                  NaN                NaN                NaN
1991-05-01 00:00:00                  NaN                NaN                NaN
1995-12-30 00:00:00                  NaN                NaN                NaN
1995-10-06 00:00:00                  NaN                NaN                NaN
1993-11-20 00:00:00                  NaN                NaN                NaN

                            ID  Source Location Source Magnitude Source  \
Date(YYYY/MM/DD)
1992-05-12 00:00:00  USP000577C     US             AEI              HRV
1991-05-01 00:00:00  USP0004R34     US             AEI               NC
1995-12-30 00:00:00  USP00079V0     US             AEI              HRV
1995-10-06 00:00:00  USP00074KV     US             AEI               US
1993-11-20 00:00:00  USP00063Z5     US             AEI              HRV

                       Status Date
Date(YYYY/MM/DD)
1992-05-12 00:00:00  Reviewed  AEI
1991-05-01 00:00:00  Reviewed  AEI
1995-12-30 00:00:00  Reviewed  AEI
1995-10-06 00:00:00  Reviewed  AEI
1993-11-20 00:00:00  Reviewed  AEI

file_name = 'database.csv'
print('DataFrame is written to Excel File successfully.')
print(df.columns)

DataFrame is written to Excel File successfully.
Index(['Latitude(deg)', 'Longitude(deg)', 'Type', 'Depth(km)',
       'Depth Error(km)', 'Depth Seismic Stations(km)', ' Magnitude(ergs)',
       'Magnitude_type', 'Magnitude Error', 'Magnitude Seismic Stations',
       'Azimuthal Gap', 'Horizontal Distance', 'Horizontal Error',
       ' RootMeanSquare ', 'ID ', 'Source', 'Location Source',
```

```
       'Magnitude Source', 'Status', 'Date'],
      dtype='object')



import warnings
warnings.filterwarnings('ignore')



from sklearn.model_selection import train_test_split

# Select relevant columns
X = df[['Latitude(deg)', 'Longitude(deg)', 'Depth(km)']]
y = df[' Magnitude(ergs)']

# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=0)




from sklearn.linear_model import LinearRegression

# Train the linear regression model
regressor = LinearRegression()
regressor.fit(X_train, y_train)


LinearRegression()

from sklearn.metrics import r2_score, mean_squared_error

scores= {"Model name": ["Linear regression", "SVM", "Random Forest"], "mse":
[], "R^2": []}

# Predict on the testing set
y_pred = regressor.predict(X_test)

# Compute R^2 and MSE
r2 = r2_score(y_test, y_pred)
mse = mean_squared_error(y_test, y_pred)

scores['mse'].append(mse)
scores['R^2'].append(r2)

print("R^2: {:.2f}, MSE: {:.2f}".format(r2, mse))
```

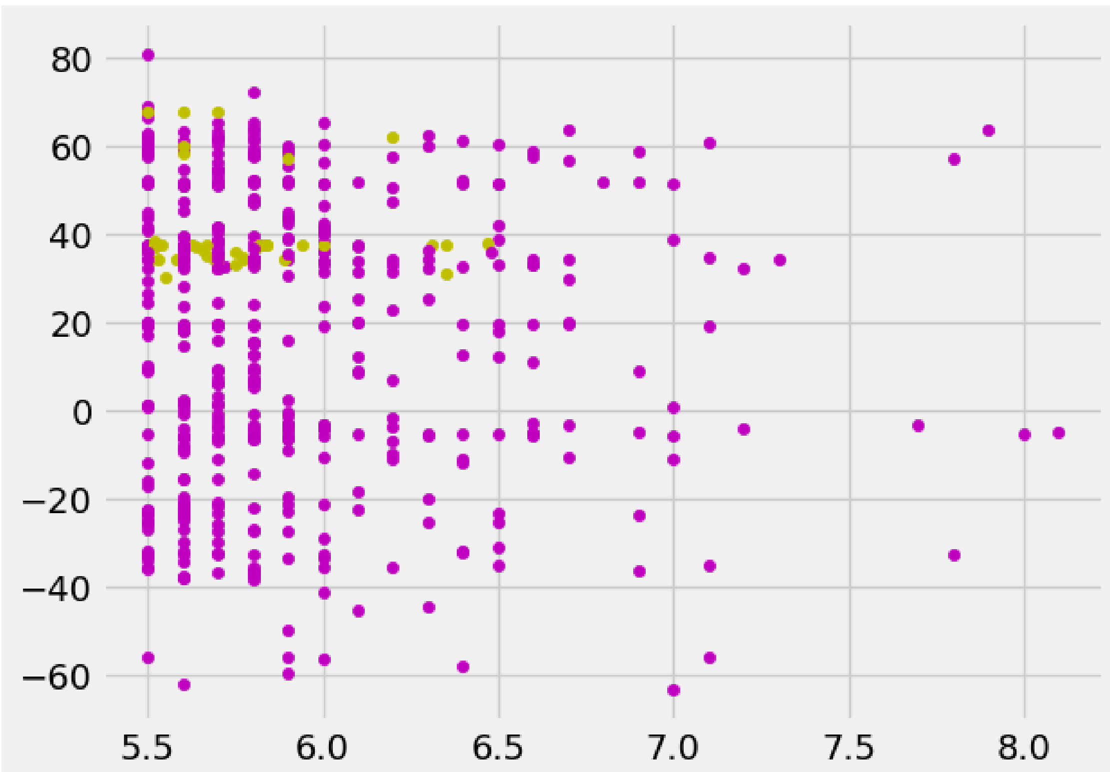
```
R^2: 0.00, MSE: 0.18

new_data = [[33.89, -118.40, 16.17], [37.77, -122.42, 8.05]]
new_pred = regressor.predict(new_data)
print("New predictions:", new_pred)

New predictions: [5.87417003 5.87453261]


import seaborn as sns
import matplotlib.pyplot as plt

# Plot the regression line
sns.regplot(x=X_test['Latitude(deg)'], y=y_test, color='blue',
scatter_kws={'s': 10})
sns.regplot(x=X_test['Longitude(deg)'], y=y_test, color='red',
scatter_kws={'s': 10})
sns.regplot(x=X_test['Depth(km)'], y=y_test, color='yellow',
scatter_kws={'s': 10})
plt.legend(labels=['Latitude(deg)', 'Longitude(deg)', 'Depth(km)'])
plt.xlabel('Predictor Variables')
plt.ylabel('Magnitude')
plt.title('Multiple Linear Regression Model')
plt.show()
```



```
from sklearn.svm import SVR

# Select a subset of the training data
subset_size = 500
X_train_subset = X_train[:subset_size]
y_train_subset = y_train[:subset_size]

# Create an SVM model
svm = SVR(kernel='rbf', C=1e3, gamma=0.1)

# Train the SVM model on the subset of data
svm.fit(X_train_subset, y_train_subset)

# Evaluate the model on the test set
score = svm.score(X_test, y_test)
print("Test score:", score)


Test score: -0.1688371824477879

# Predict on the testing set
y_pred_svm = svm.predict(X_test)

# Compute R^2 and MSE
r2_svm = r2_score(y_test, y_pred_svm)
mse_svm = mean_squared_error(y_test, y_pred_svm)

scores['mse'].append(mse_svm)
scores['R^2'].append(r2_svm)

print("SVM R^2: {:.2f}, MSE: {:.2f}".format(r2_svm, mse_svm))

SVM R^2: -0.17, MSE: 0.21



# Predict on new data
new_pred_svm = svm.predict(new_data)
print("New SVM predictions:", new_pred_svm)


New SVM predictions: [5.88892086 5.88874738]

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib import style
from sklearn.svm import SVC

style.use('fivethirtyeight')
```

```
# create mesh grids
def make_meshgrid(x, y, h =.02):
    x_min, x_max = x.min() - 1, x.max() + 1
    y_min, y_max = y.min() - 1, y.max() + 1
    xx, yy = np.meshgrid(np.arange(x_min, x_max, h), np.arange(y_min, y_max,
h))
    return xx, yy

# plot the contours
def plot_contours(ax, clf, xx, yy, **params):
    Z = clf.predict(np.c_[xx.ravel(), yy.ravel()])
    Z = Z.reshape(xx.shape)
    out = ax.contourf(xx, yy, Z, **params)
    return out

# color = ['y', 'b', 'g', 'k']

subset_size = 500

# modify the column names based on the dataset
features = df[[' Magnitude(ergs)','Latitude(deg)']][:subset_size].values
classes = df['Magnitude_type'][:subset_size].values

# create 3 svm with rbf kernels
svm1 = SVC(kernel ='rbf')
svm2 = SVC(kernel ='rbf')
svm3 = SVC(kernel ='rbf')
svm4 = SVC(kernel ='rbf')

# fit each svm's
svm1.fit(features, (classes=='ML').astype(int))
svm2.fit(features, (classes=='MS').astype(int))
svm3.fit(features, (classes=='MB').astype(int))

fig, ax = plt.subplots()
X0, X1 = features[:, 0], features[:, 1]
xx, yy = make_meshgrid(X0, X1)

# plot the contours
'''
plot_contours(ax, svm1, xx, yy, cmap = plt.get_cmap('hot'), alpha = 0.8)
plot_contours(ax, svm2, xx, yy, cmap = plt.get_cmap('hot'), alpha = 0.3)
plot_contours(ax, svm3, xx, yy, cmap = plt.get_cmap('hot'), alpha = 0.5)
'''
color = ['y', 'b', 'g', 'k', 'm']
```

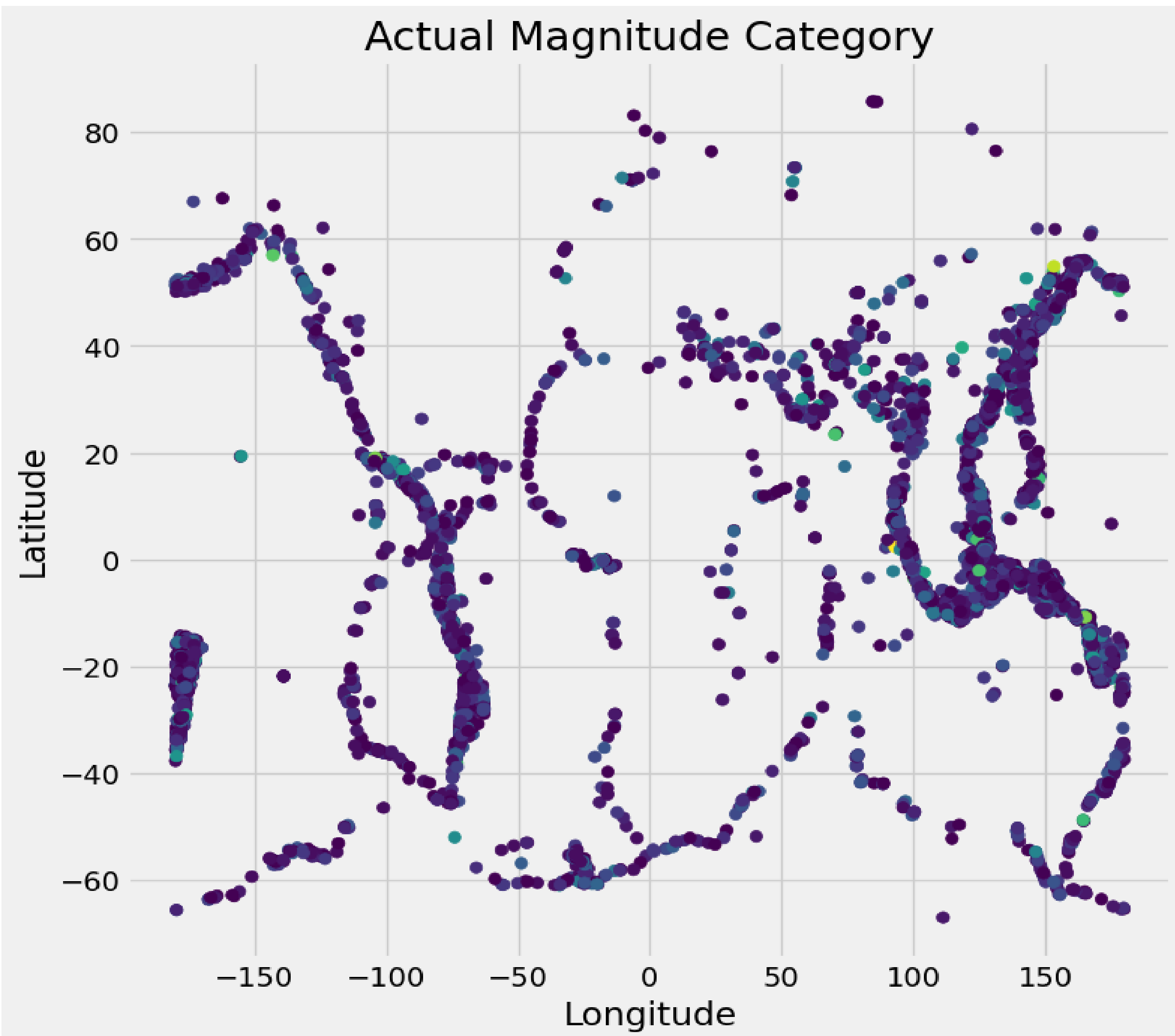
```
for i in range(subset_size):
    if classes[i] == 'ML':
        plt.scatter(features[i][0], features[i][1], s = 20, c = color[0])
    elif classes[i] == 'Mx':
        plt.scatter(features[i][0], features[i][1], s = 20, c = color[1])
    elif classes[i] == 'Md':
        plt.scatter(features[i][0], features[i][1], s = 20, c = color[2])
    else:
        plt.scatter(features[i][0], features[i][1], s = 20, c = color[4])
plt.show()
```

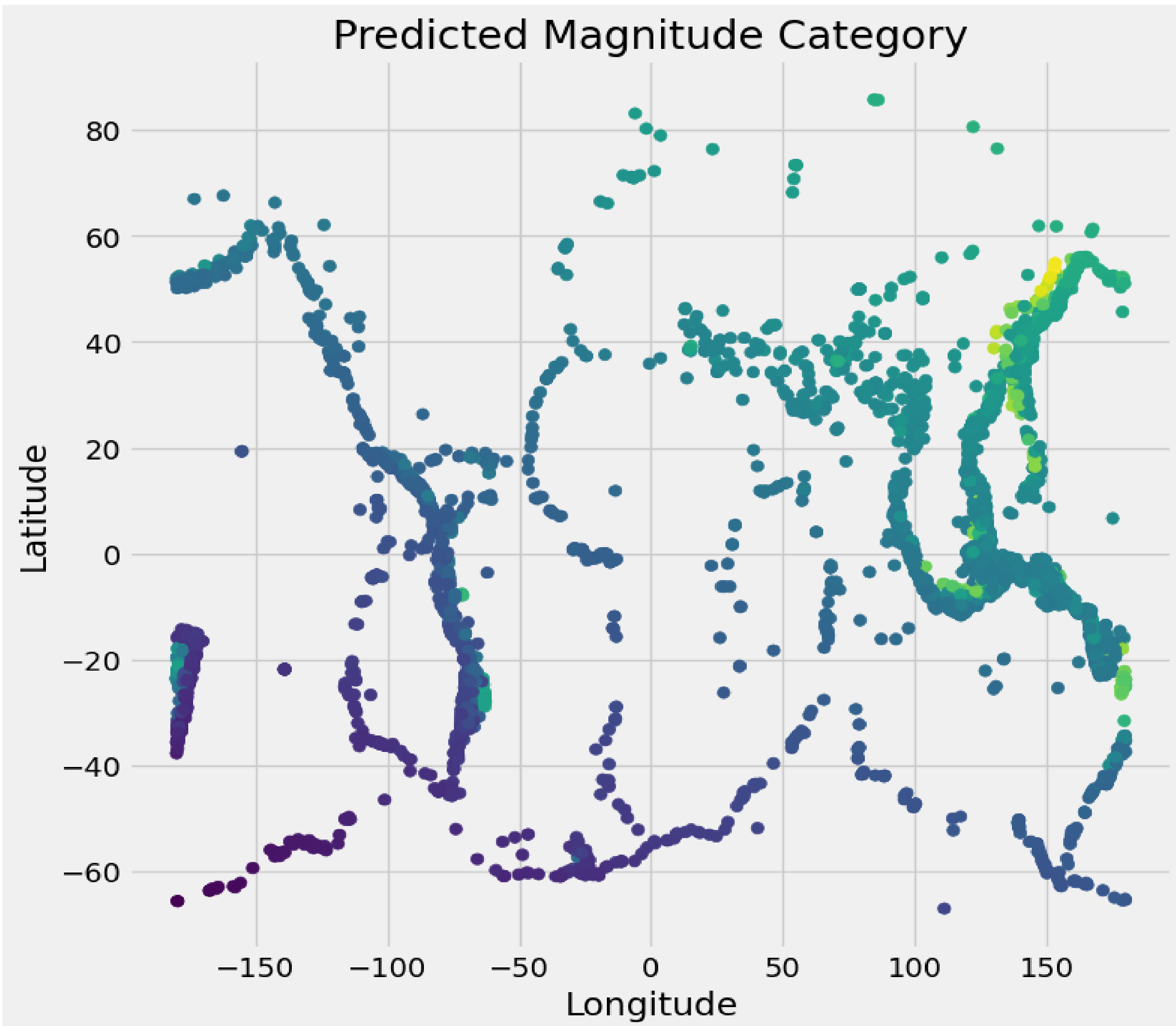



```
print(df.columns)
df['Magnitude_type'].unique()
```

```
Index(['Latitude(deg)', 'Longitude(deg)', 'Type', 'Depth(km)',
       'Depth Error(km)', 'Depth Seismic Stations(km)', ' Magnitude(ergs)',
       'Magnitude_type', 'Magnitude Error', 'Magnitude Seismic Stations',
       'Azimuthal Gap', 'Horizontal Distance', 'Horizontal Error',
       ' RootMeanSquare ', 'ID ', 'Source', 'Location Source',
       'Magnitude Source', 'Status', 'Date'],
      dtype='object')

array(['MW', 'MS', 'MWC', 'MWW', 'MWB', 'MB', 'ML', nan, 'MH', 'MWR',
       'MD'], dtype=object)
```

```
plt.figure(figsize=(8, 8))
plt.scatter(X_test['Longitude(deg)'], X_test['Latitude(deg)'], c=y_test,
cmap='viridis')
plt.title('Actual Magnitude Category')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.show()
print(" ")
plt.figure(figsize=(8, 8))
plt.scatter(X_test['Longitude(deg)'], X_test['Latitude(deg)'], c=y_pred,
cmap='viridis')
plt.title('Predicted Magnitude Category')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.show()
print(" ")
```




Predicted Magnitude Category
80
60
40
20
0
−20
−40
−60
Latitude
−150
−100
−50
0
50
100
150
Longitude

## Conclusion:

- In conclusion, our analysis shed light onthe potential of machine learning in earthquake prediction. While the models offered valuable insights and predictability, they represent just the beginning of a much broader field of research.
- Earthquake prediction remains a challenging domain, and our study serves as a foundational effort to enhance the accuracy and reliability of such predictions.