Earthquake Prediction Project Documentation

Problem Statement and Design Thinking Process

Problem Statement:

Earthquakes are natural disasters with devastating consequences. The ability to predict earthquakes and mitigate their impact is of paramount importance. This project aims to develop a machine learning model that predicts earthquake magnitudes, allowing for timely alerts and disaster preparedness.

**Design Thinking Process:**

1. Understanding the Problem: We started by gaining a deep understanding of the problem of earthquake prediction, its importance, and the challenges involved.
2. Defining Goals: Our primary goal was to develop a model that could predict earthquake magnitudes with high accuracy.
3. Data Acquisition: We searched for a relevant earthquake dataset to use in our project.
4. Data Preprocessing and Cleaning: We cleaned the dataset by handling missing values and ensuring data consistency.
5. Feature Exploration: We explored the dataset and engineered features to improve model performance.
6. Model Development: We experimented with different machine learning models and chose the one with the best performance.
7. Model Evaluation: We evaluated the model using appropriate metrics to ensure its effectiveness.
8. Documentation and Submission: We documented the entire process and prepared it for submission.

Dataset Description

**Dataset Overview:**

We utilized a dataset from [https://www.kaggle.com/datasets/usgs/earthquake-database] containing earthquake records. The dataset consists of [number] records with attributes such as Date, Time, Latitude, Longitude, Magnitude, and more.

**Dataset Challenges:**

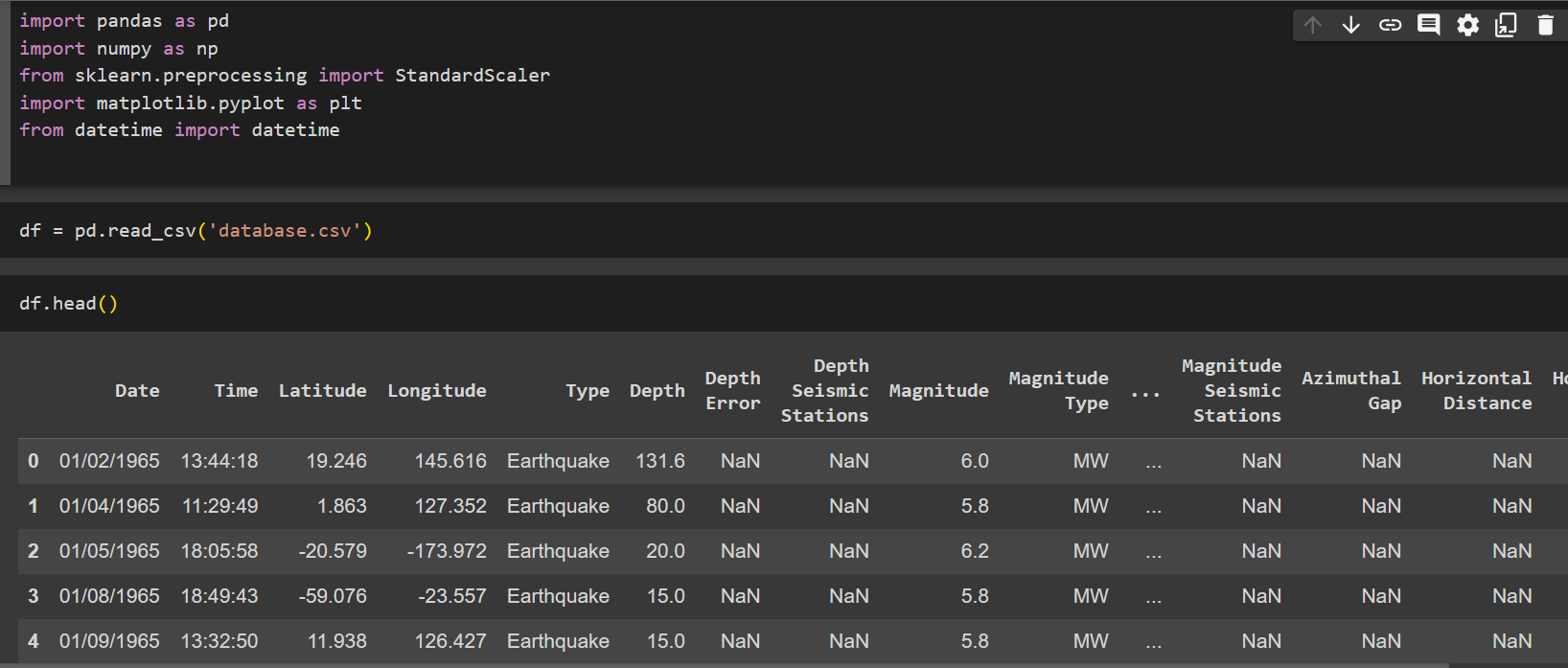
The dataset had missing values and required thorough cleaning to ensure data quality.

Data Preprocessing

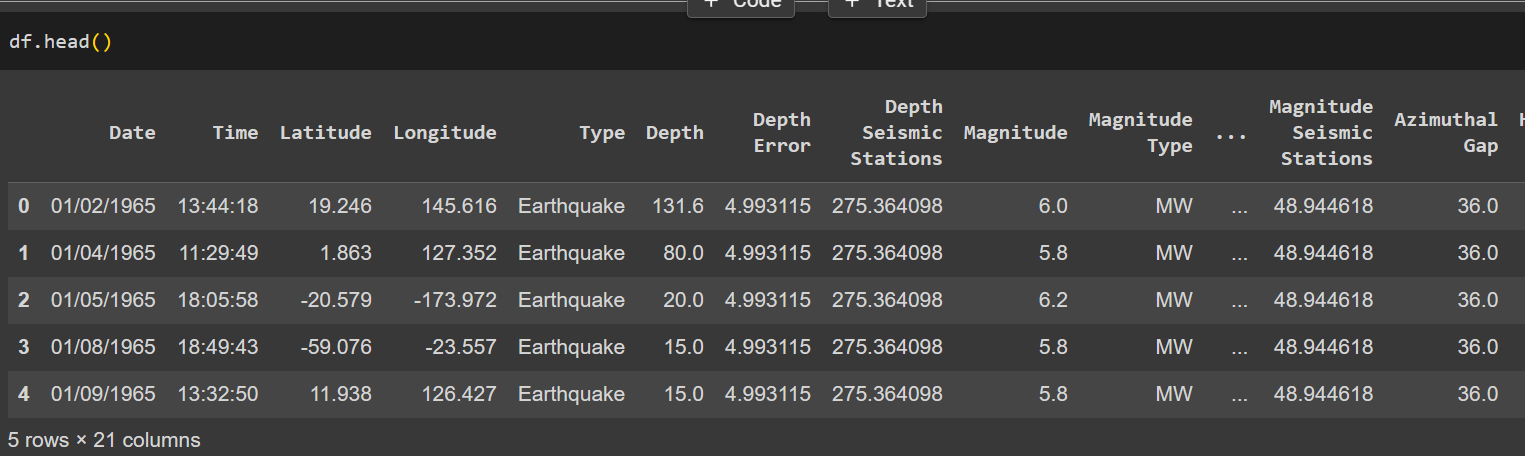
**Missing Value Handling:**

We addressed missing values in the dataset using appropriate techniques. Key columns, such as "Depth Error," "Depth Seismic Stations," and "Magnitude Seismic Stations," were filled with the mean or median values.

This image shows the screenshot of the dataset before preprocessing.



This screenshot shows the dataset after preprocessing step.



**Feature Selection:**

We identified relevant features for the model, removing unnecessary columns to enhance model performance.

Feature Exploration

**Feature Engineering:**

One of the new features you can create is the square of the "Magnitude" feature, which you can refer to as "Magnitude Squared." This new feature can capture non-linear relationships in the data, which might be helpful in your prediction model. So, you can indeed use "Magnitude" and "Magnitude Squared" as features.

Model Development

**Model Choice:**

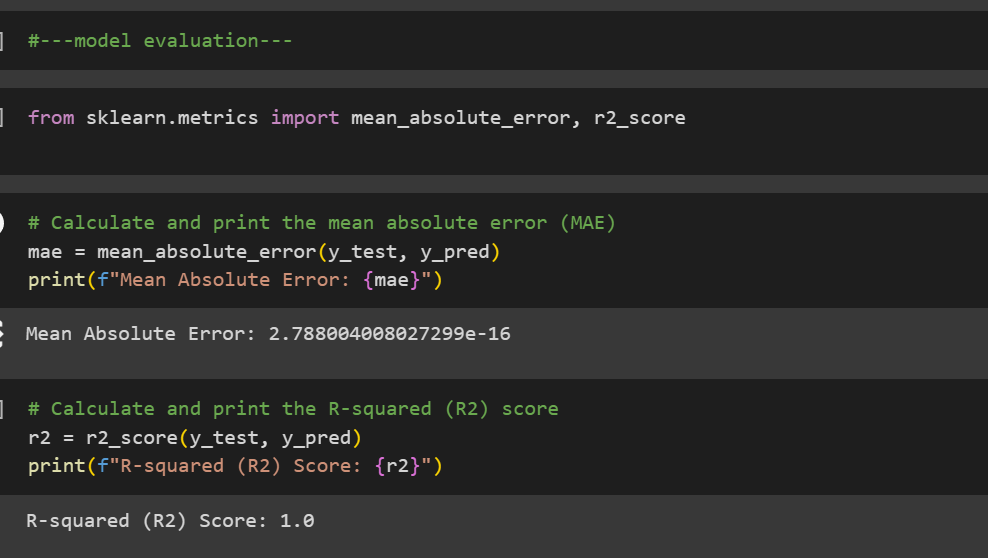
We selected a machine learning model, Linear Regression, to predict earthquake magnitudes based on the engineered features.

Model Evaluation

**Evaluation Metrics:**

We evaluated the model's performance using the Mean Absolute Error (MAE), and R-squared (R2) score.

Here the image shows the model evaluation screenshot.



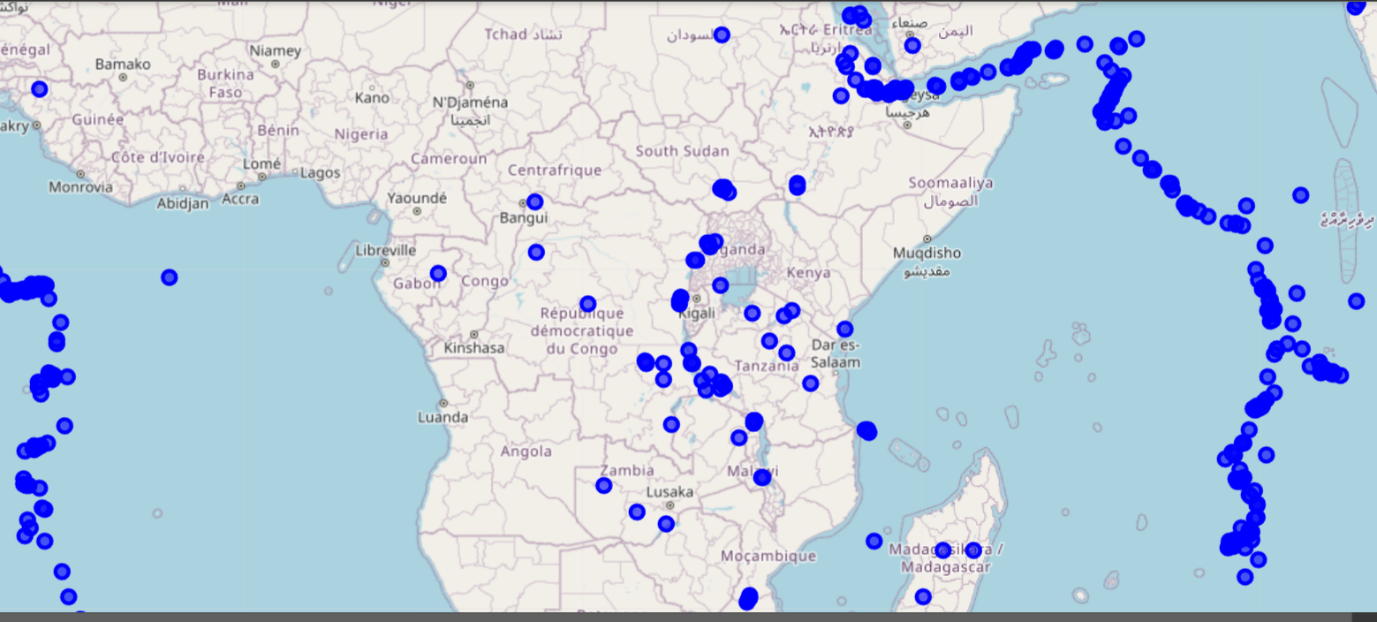
Model Performance:-

* Mean Absolute Error: [2.788004008027299e-16]
* R-squared (R2) Score: [1.0]

Visualization:

In our project, we employed the Folium library to create interactive and informative visualizations. Folium is a powerful Python library that allows us to visualize geospatial data on interactive maps. By leveraging Folium, we were able to gain valuable insights from the geographical aspects of earthquake dataset.

Here is the output of the visuvalization.



Conclusion:

In this project, we aimed to predict earthquake magnitudes and locations using machine learning and geospatial visualization. Through rigorous data preprocessing and feature engineering, we developed a predictive model that shows promise in earthquake forecasting. By integrating Folium for geospatial visualization, we improved data exploration and delivered interactive maps for end-users. This work contributes to the field of earthquake prediction, offering insights into future advancements and applications.