Phase 1: Problem Definition and Design Thinking

Problem Definition:

To Develop a machine learning model for earthquake magnitude using a dataset from kaggle.The goal is to build a neural network model that can accurately predict the magnitude of earthquakes .This predictive model will assist in assessing the potential impact of earthquakes in different regions.

Design Thinking:

1. **Data Source**: Choosing a suitable Kaggle dataset containing earthquake data with features like date, time, latitude, longitude, depth, and magnitude. The kaggle dataset **Significant Earthquakes, 1965-2016** provided by the National Earthquake Information Center (NEIC) has been chosen. (https://www.kaggle.com/datasets/usgs/earthquake-database)
2. **Feature Selection:** Identifying which features are most relevant for earthquake magnitude prediction. Features like latitude, longitude, depth, and historical earthquake data are crucial.
3. **Feature Transformation**: Performing transformations such as scaling, encoding categorical variables, and extracting additional features like seasonality or proximity to fault lines.
4. **Visualization**: Creating a world map visualization to display the distribution of earthquake frequencies. Colour coding or markers to represent the magnitude of earthquakes at different geographical locations. This visualization can help to easily identify earthquake-prone regions.
5. **Data Splitting**: Splitting the dataset into a training set and a test set. Typically, an 80-20 or 70-30 split is used, where the majority of data is allocated for training and a smaller portion for testing.
6. **Model Development:** Building a neural network model for earthquake magnitude prediction. Libraries like TensorFlow and PyTorch are used to create the model. Also the input features and output features(earthquake magnitude) are defined .
7. **Training and Evaluation**: Training the neural network model on the training set. Also the training process is monitored and adjustments are made as required .Then the model's performance is evaluated by using the test set For earthquake prediction, metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), or R-squared are used to assess how well the model predicts earthquake magnitudes.