Phase 1: Problem Definition and Design Thinking.

Earthquake Prediction Model using Python

Explore the key features of earthquake data and design an object for those features, such as date, time, latitude, longitude, depth, and magnitude. Before developing the prediction model, visualize the data on a world map to display a complete overview of where the earthquake frequency will be higher. Split the data into a training set and a test set for validation. Lastly, build a neural network to fit the data from the training set

Abstract

The goal of this project is to predict the date, time, latitude, longitude, depth, and magnitude of an earthquake for a region given by the user with the help of historical data. As we know the destruction caused by earthquakes is massive and causes loss of lives every year.

Earthquake's association with structural damage and loss of life is one that keeps on enduring and thus is the focal point of consideration for many fields, say, seismological research and environmental engineering yet not limited to these[1]. It's significance is stretched out to human life too, to sustain and to survive. A prediction that can be accurate and relied on is a requisite for all the areas prone to disasters and as well as for locations that have less to no chances. It will get us ready for all the worst possible scenarios and for necessary measures as well that can be taken beforehand to solve the upcoming crisis. As technology is evolving and helping humans for a better and a convenient lifestyle, the possibility of saving lives is taken up with the help of efficient ML algorithms and Data Science to give accurate forecasts.

Machine Learning is a subset of Artificial Intelligence. It permits the system to adapt to a behaviour of a particular kind based on its own learning and possesses the ability to improve itself naturally solely from experience without any explicit programming, human mediation or help[8]. Initialization of a machine learning process starts with feeding an honest quality data set to the algorithm(s), so as to build a ML prediction model. Algorithms perform knowledge discovery and statistical evaluation, determining patterns and trends in data. Selection of algorithms relies on data and on the task that requires automation.

Our target is foreseeing catastrophic events and improving the manner in which we react to them. Great forecasts and admonitions spare lives. A notice of an approaching calamity can be issued well ahead of time as it will help in reducing both death occurrence and structural loss. ML algorithms construct two types of predictive models, Regression and Classification models[6]. Each of them approaches data in a different way. Concerned system makes use of regression model whose core idea is forecasting a numerical value.

SYSTEM DEVELOPMENT

Data sampling

The data set is very huge so as we get the location of the area where we want to predict the earthquake. The following figure shows the flow of how data is collected. Figure[1]

PROPOSED WORK

Developing predictive modelling involves gradual procedure. Tools which are conventionally used for developing model are Python, Hadoop and R. Various steps involved are:

A. DATA ACQUISITION

Data acquisition is the process for bringing data for production use either from source outside the system and into the system, or from data produced by the system. This is the underlying advance to start and alludes to gathering required information. We obtain required data sets from government provided website such as

– • USGS.gov (United States Geological Survey)- Scientific agency of the United States government.

• IMD.gov (India Meteorological Department)- Agency of the Ministry of Earth Sciences of the Government of India.

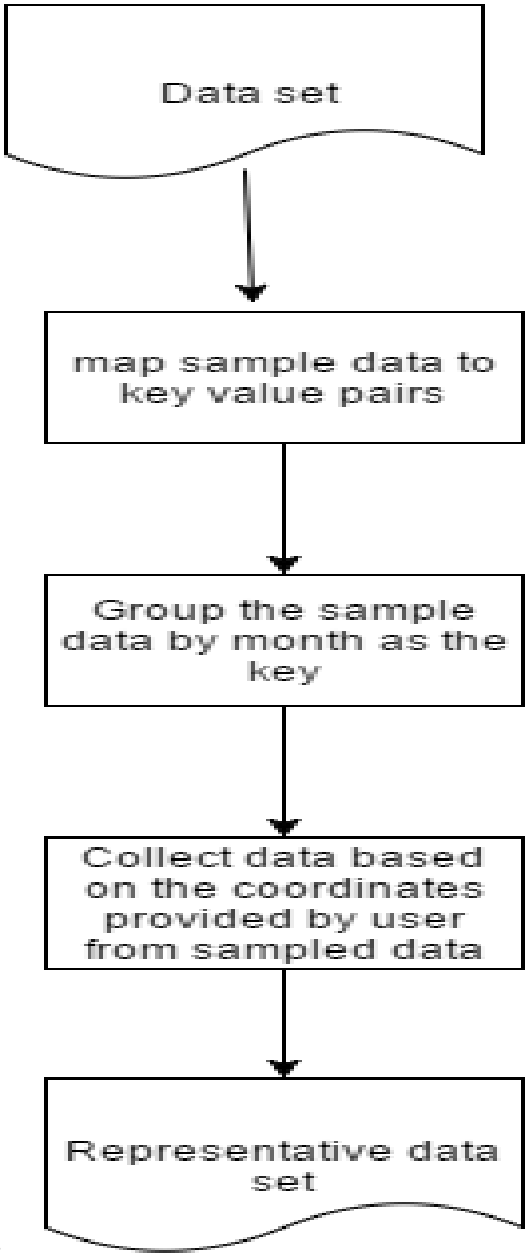
Google Acquired Kaggle contains data-set collected from different agencies of different governments.

The columns in the data-set are –

- Date
- Time
- Latitude
- Longitude
- Magnitude
- Depth

Algorithm:

1. Input data set and load libraries.
2. Data Pre-processing.
3. Model Building.
4. Making Predictions


Data set
map sample data to key value pairs
Group the sample data by month as the key
Collect data based on the coordinates provided by user from sampled data
Representative data set

Figure[1}