Date of publication 12.31. 2022

Earthquake Prediction Project by Using Linear Regression Model in Python Programming Language(December 2022)

Kübra Nur Tiryaki, Earthquake Prediction, Final Project, İstanbul, 2022
İstanbul Aydın University,, Engineering Faculty, Software Engineering, İstanbul/Türkiye

kubranurtiryakibusiness@gmail.com

Abstract

Scientists expect earthquakes in the near future. But since there are so many factors, it is impossible to forecast an earthquake, as everyone is aware. I developed a model utilizing historical earthquake data to forecast potential future earthquakes within a few parameters that are essential to our project. My dataset was discovered using earthquakes that took place between 1910 and 2017. Then, I only included earthquakes that occurred in Turkey by limiting the countries in my data set. After that, I changed the date and time information to match my model. I then eliminated information from my dataset that was not necessary for my model. I used the dataset to do out outlier detection. I then performed the required computations using the numbers in my data set. Then I divided the data into the eigenvalues, which represented the remaining data, and the Ritcher scale, which represented the value I wished to forecast. I then developed a linear regression model and used my eigenvalues to train the model. In the end, when I asked my model to predict results based on the [39.21,41.40,14.0,4.7,4.7] data, it did so with a very low margin of error.

Keywords *Earthquake, prediction, Turkey*

I. INTRODUCTION

Earthquake Prediction Project by Using Linear Regression Model in Python Programming Language is a study that establishing a model suitable for the data set we have and predicting results with high accuracy. The aim of the project is to create a Linear Regression Model using the data of past earthquakes. Of course, it is necessary to have a lot of variables to create an earthquake forecasting project. In addition, there are many factors that interfere with the calculations, except for the variables that we found and used. Predicting earthquakes and writing its project is a subject that may take many years of work, this version of my project is just the beginning and will continue to evolve.

II. MODEL, METHODOLOGY, and CODING

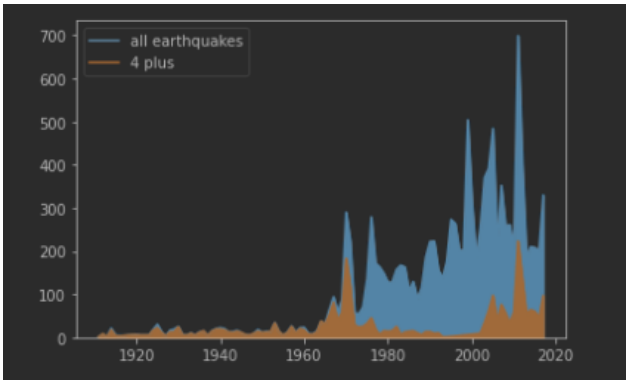
II.I Explanation of Data Set

lat = Latitude of the earthquake where it happened
long = Longitude of the earthquake
depth = The deepness of the earthquake where it happened
md = Time Dependent Magnitude -> How long the earthquake creates a vibration on the seismometer is measured and scaled with the distance from the earthquake center. This method is used for small (M<5.0) and near (Distance<300 km) earthquakes.
mb = Body wave magnitude. Body-waves consist of P-waves that are the first to arrive (see seismogram), or S-waves, or reflections of either. Body-waves travel through rock directly.
ritcher = a measure of the strength of earthquakes, developed by Charles Francis Richter.

II.II Programming Language and Libraries

In our project, we used the Linear Regression model in Python Programming language. We used many libraries during the coding of our project. These are classic libraries like NumPy, Pandas, scikit-klearn, matplotlib. We also used the seaborn library to generate some statistical data representations that you will see when you look inside the code. We used the datetime library to separate the date into day, month and year for data visualization. We used the NumPy library for mathematical tasks such as calculating errors. To import pandas library, comma separated values file into Data Frame, pull required values etc. we used. We used pandas library to import the Comma Separated Values (CSV) file to Dataframe, pull required values etc. We used the scikit-learn library to divide our dataset into two as training and testing, we also used the scikit-learn library to calculate the error and score, to scale the data we will use in training, and to set our model as Linear Regression. We also used the matplotlib library for data visualization

.



An example of data visualization with matplotlib library. Here, we looked at the number of earthquakes with a magnitude of 4 and above compared to all earthquakes in which years. The reason for choosing the value of 4 is that earthquakes reach a damaging level starting from 4 according to the Richer scale.

III. MATH

In this section, we will demonstrate the mathematical equations we use in our project.

III.I EQUATIONS

Mean Square Error

$$MSE = \frac{1}{n}\sum(y - \hat{y})^2$$

Mean Absolute Error

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}.$$

Root Mean Squared Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \|y(i) - \hat{y}(i)\|^2}{N}},$$

IV. DEVELOPMENTS

IV.I Future planned developments of the Project

•More graphics can be added.

•Apart from Linear Regression, new models can be created using Random Forest, Catboost Regressor, Support Vector Regressor, K Nearest Neighbor Classifier, K-Means Clustering etc.

•I'm thinking of transferring my project to Artificial Neural Networks and developing my project further.

V. CONCLUSION

This project was beyond my knowledge of understanding earthquake data and applying it in Python Programming Language. In the end, though, the importance of constructing a Linear Regression model, planning how it should be trained with earthquake data, fitting the data into the model, adjusting different parameters to increase accuracy score is obvious. I am truly grateful for being a part of this process and being able to contribute something to myself.

ACKNOWLEDGMENT

I really appreciate all contributions of Prof. Dr. Ali Okatan for his way of teaching. Also, I really thank Dr. Atınç Yılmaz for taking this course in a way that all the students can understand the math behind all these algorithms we use. And finally, I am grateful to myself, who strives to improve herself.

REFERENCES

[1] [Online] Available: https://tr.wikipedia.org/wiki/Richter_%C3%B6l%C3%A7e%C4%9Fi
[2] [Online] Available: https://www.kaggle.com/datasets/alpkoc/historical-earthquake-dataset-of-turkey
[3] [Online] Available: https://docs.google.com/document/d/1dotkQrbiH5sX2crGxXntHArMJuYD0mMe/edit
[4] [Online] Available: https://en.wikipedia.org/wiki/Body_wave_magnitude
[5] [Online] Available: http://ieeeshutpages.s3-website-us-west-2.amazonaws.com/

Kübra Nur Tiryaki was born in Istanbul, Turkey in 1999. After high school, she entered the top 28 thousand in the university placement exam and entered the Software Engineering Department at Istanbul Aydın University with a full scholarship, she is expected to finish school in 2023.

Although Ms. Tiryaki is interested in many fields due to her learning instinct, she wishes to advance in artificial intelligence.