University of Sulaymaniyah

Collage of science

Computer department

Data Science Project

About

Exploring Seismic Signals- Advanced Data Science Insights into Earthquake Dynamics

prepared by

Sonya mhamad

Shnyar Baxtyar

Astera Haji

Sarko Mhamad

Supervised by

Dr.Miran Taha

2023-2024

Table of Contents

Introduction 3

Problem statement 4

Solution Method 4

- **Data Gathering** 4
- **Data Cleaning** 5
- **Feature Selection** 5
- **Model Building** 5
- **Model Testing** 6
- **Model Improvement** 6

Implementation 7

..... 9

..... 9

Results 10

Project Conclusion 10

Introduction

Earthquakes, those powerful shakes in the ground, are natural disasters that can bring significant challenges to communities all around the world. Recognizing the seriousness of their impact, this data science project embarks on a mission to employ advanced tools and techniques to gain a deeper understanding of earthquakes and predict when they might occur. The goal is to contribute to the safety of communities by harnessing the power of data.

We understand that earthquakes can be incredibly dangerous, causing destruction and harm to people and places. Therefore, our primary objective is to delve into seismic data and historical records. We want to extract valuable insights that can significantly improve earthquake preparedness and mitigation strategies. Imagine it as trying to solve a puzzle where the pieces are made up of data and information.

Analyzing seismic data and historical records, our project seeks patterns and correlations to enhance earthquake prediction. Using smart tools, we piece together insights for a comprehensive understanding of earthquake dynamics. This scientific investigation aims to improve prediction accuracy, providing valuable information for minimizing the impact on communities.

Problem statement

Our project aims to solve the challenge of predicting earthquakes accurately. Earthquakes are natural disasters that can cause a lot of damage and harm to people and places. Even though scientists have learned a lot about earthquakes, it's still hard to predict when they will happen and how strong they will be.

This problem is important because earthquakes can be very dangerous, and being able to predict them better can help keep people safe. If we can predict earthquakes more accurately, we can prepare for them better and reduce the damage they cause to buildings and communities. Our goal is to use simple language and data science to improve how we understand and predict earthquakes, ultimately making communities safer.

Solution Method

To address the challenge of predicting earthquakes, we've devised a systematic approach that involves the following steps:

- **Data Gathering**

We initiate our process by acquiring data from diverse sources such as seismic databases, geological records, and governmental repositories.

This data encompasses a wide range of information about past earthquake events, including their locations, magnitudes, depths, and temporal occurrences.

- **Data Cleaning**

After obtaining the data, we meticulously clean and refine it to ensure its accuracy and reliability.

This involves identifying and rectifying any errors, inconsistencies, or missing values present in the dataset.

By ensuring the cleanliness of the data, we lay a robust foundation for subsequent analysis and modeling.

- **Feature Selection**

With the cleaned data in hand, we proceed to select the most pertinent features for our predictive models.

These features are carefully chosen based on their relevance to earthquake prediction and their potential to contribute to model accuracy.

Factors such as earthquake location, historical seismic activity, geological characteristics, and temporal trends are considered during this stage.

- **Model Building**

Leveraging state-of-the-art machine learning algorithms, we embark on the task of building predictive models.

These models are designed to utilize the selected features to learn patterns and relationships inherent in the data.

- **Model Testing**

Once the models are constructed, we subject them to rigorous testing to evaluate their predictive performance.

This involves comparing the model predictions with actual earthquake occurrences to assess their accuracy and reliability.

Through meticulous testing, we gain insights into the strengths and limitations of each model, enabling us to make informed decisions for further refinement.

- **Model Improvement**

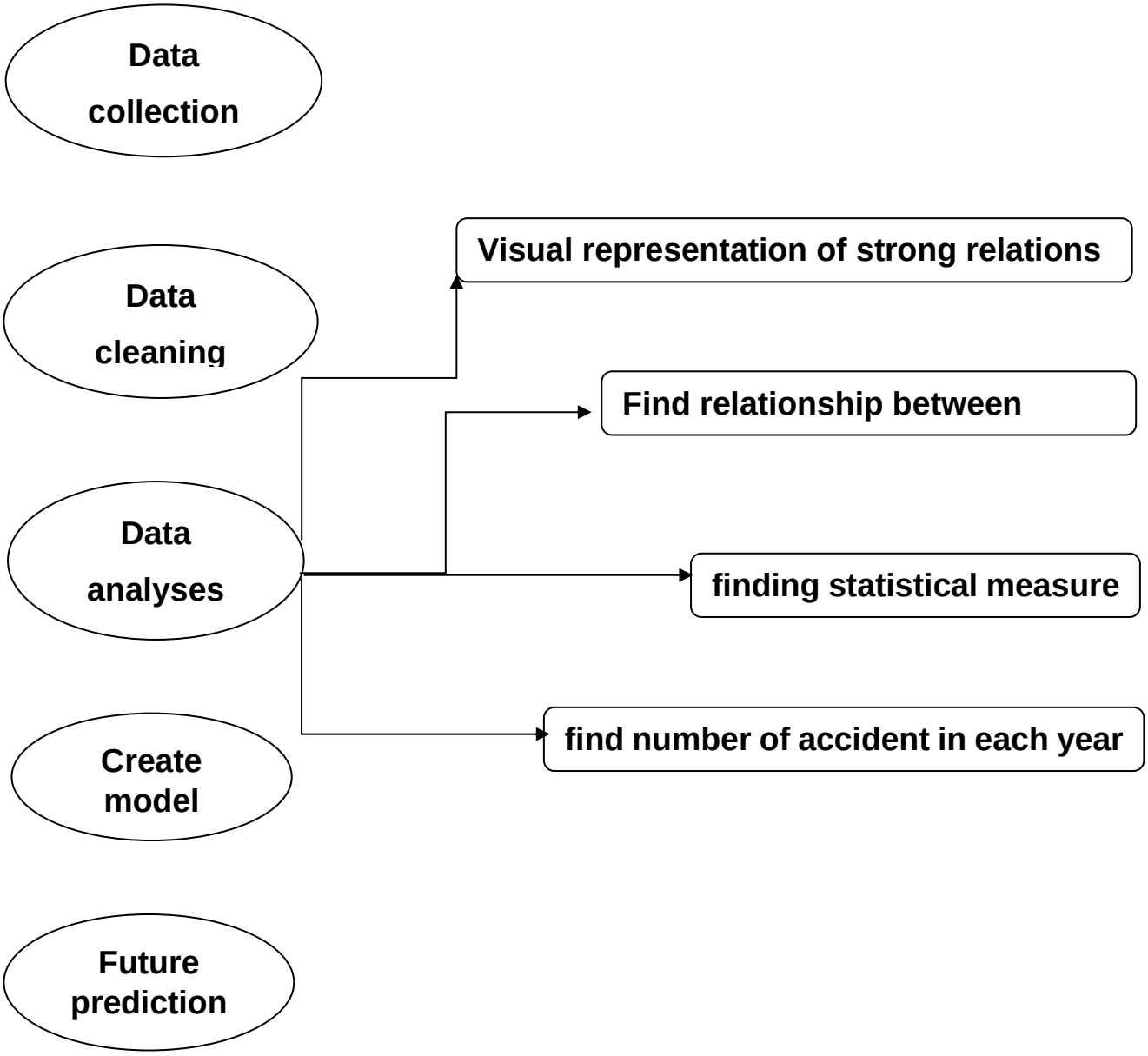
In cases where the initial models exhibit suboptimal performance, we embark on a process of continuous improvement.

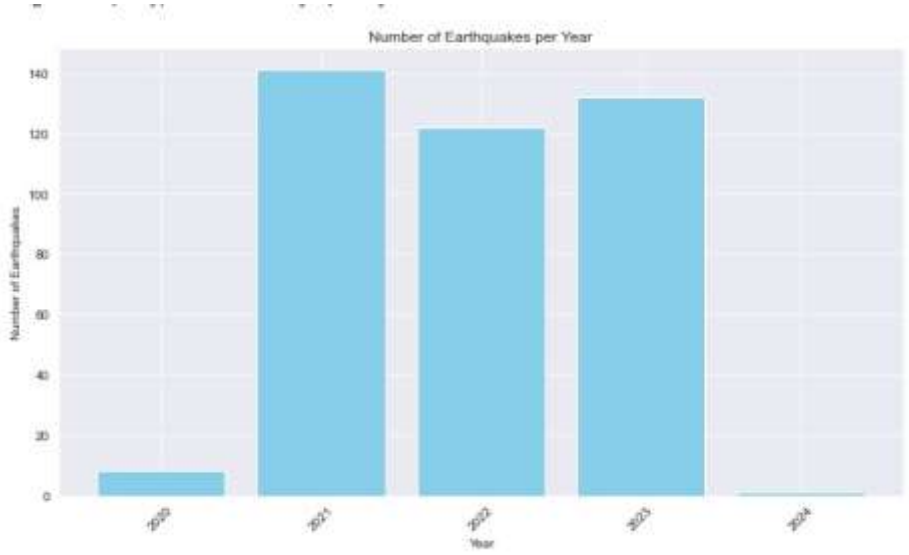
This may involve fine-tuning model parameters, incorporating additional features, or exploring alternative algorithms.

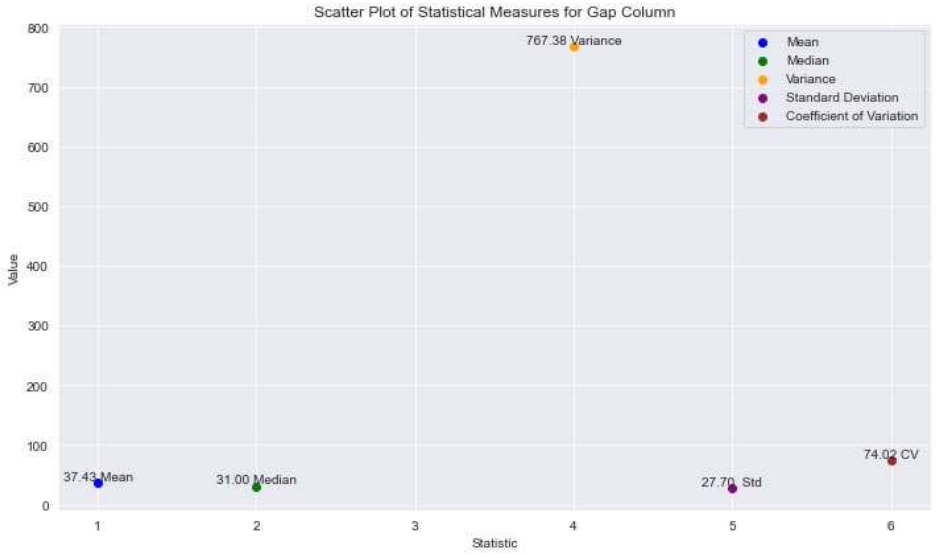
By iteratively refining the models based on testing results and domain expertise, we strive to enhance their predictive capabilities.

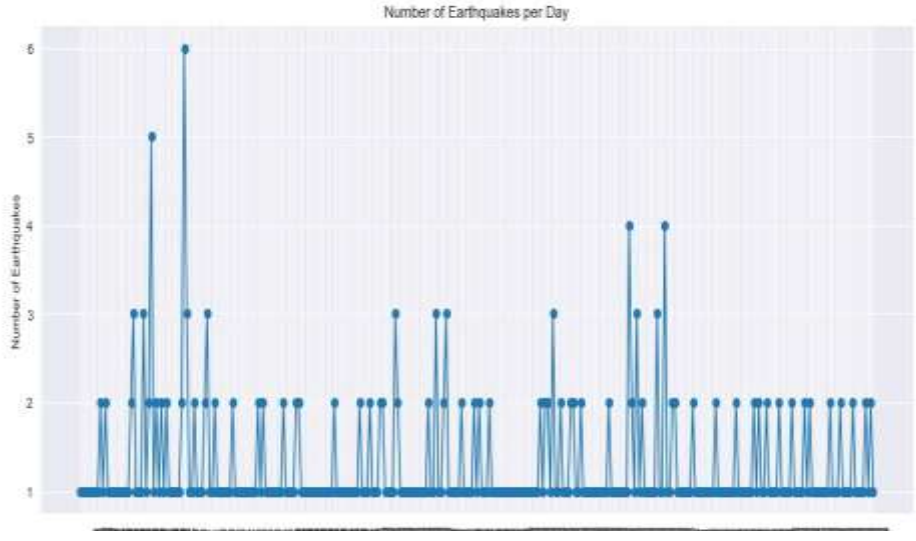
Through this methodical approach, we aim to develop robust predictive models that can provide valuable insights into earthquake occurrence, ultimately contributing to enhanced disaster preparedness and community safety.

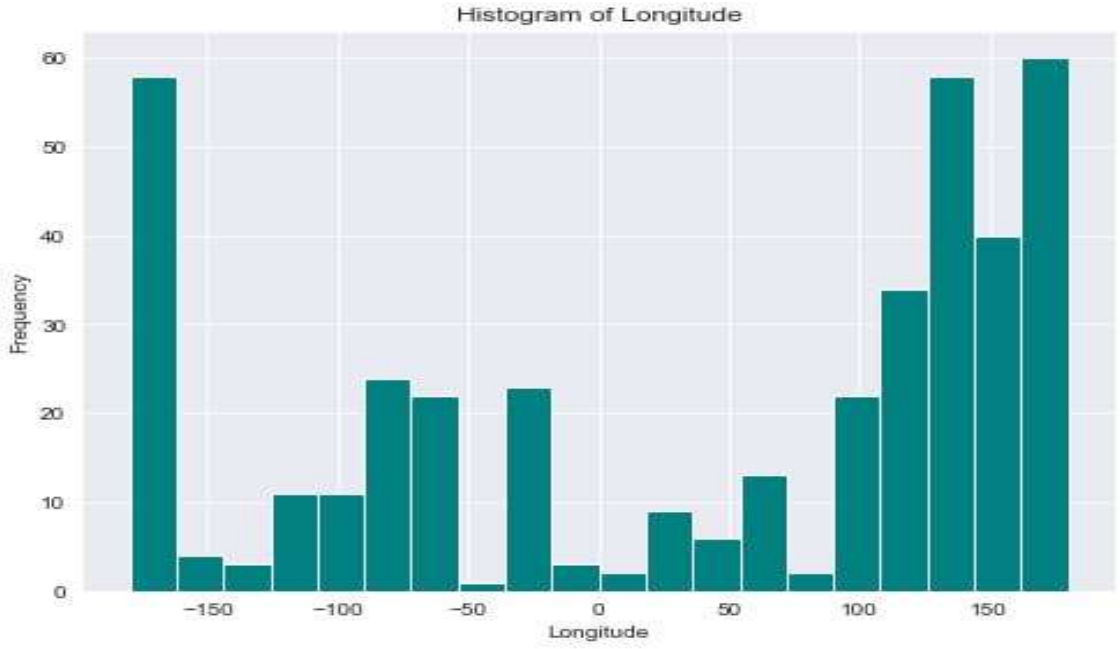
Implementation




Number of Earthquakes per Year
Number of Earthquakes
Year
2020
2021
2022
2023
2024


Scatter Plot of Statistical Measures for Gap Column
767.38 Variance
Mean
Median
Variance
Standard Deviation
Coefficient of Variation
37.43 Mean
31.00 Median
27.70 Std
74.02 CV
Value
Statistic


Number of Earthquakes per Day
Number of Earthquakes


Histogram of Longitude
Frequency
Longitude
0
10
20
30
40
50
60
−150
−100
−50
0
50
100
150


Correlation Heatmap
latitude
longitude
depth
nst
gap
dmin
rms
horizontalError
depthError
magError
magNst
1.00 0.16 -0.13 0.19 -0.01 -0.43 -0.09 -0.44 -0.04 -0.29 0.16
0.16 1.00 -0.00 0.08 -0.29 -0.02 -0.00 0.04 0.10 0.09 0.03
-0.13 -0.00 1.00 0.12 -0.21 -0.01 0.01 0.26 0.23 -0.01 0.06
0.19 0.08 0.12 1.00 -0.16 -0.01 -0.20 0.01 0.09 -0.08 0.12
-0.01 -0.29 -0.21 -0.16 1.00 -0.07 -0.17 -0.12 -0.09 0.03 -0.05
-0.43 -0.02 -0.01 -0.01 -0.07 1.00 -0.01 0.42 -0.02 0.20 -0.07
-0.09 -0.00 0.01 -0.20 -0.17 -0.01 1.00 0.15 0.07 0.16 0.02
-0.44 0.04 0.26 0.01 -0.12 0.42 0.15 1.00 0.25 0.26 -0.04
-0.04 0.10 0.23 0.09 -0.09 -0.02 0.07 0.25 1.00 0.15 0.14
-0.29 0.09 -0.01 -0.08 0.03 0.20 0.16 0.26 0.15 1.00 -0.38
0.16 0.03 0.06 0.12 -0.05 -0.07 0.02 -0.04 0.14 -0.38 1.00
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

Results

The culmination of our efforts is reflected in the results section, where the main findings are succinctly summarized. Emphasis is placed on significant correlations, accurate predictions, and any unexpected discoveries gleaned from the data. These results form the basis for the subsequent discussion and interpretation.

Project Conclusion

In the Project Conclusion , we provide a concise overview of the key aspects. We begin by summarizing the achievements of our project, emphasizing the milestones and outcomes reached during its course. Following that, we outline our plans for future development, addressing potential enhancements and improvements to our earthquake prediction approach. This section also incorporates reflections on the project's impact, highlighting its significance within the broader context of ongoing progress in the field of earthquake prediction. The aim is to provide a clear understanding of the project's contributions, future directions, and its overall importance in advancing knowledge and capabilities related to earthquake prediction.