<u>Phase 4 project : Significant Earthquakes</u>

Introduction

An earthquake occurs when two blocks of the Earth's crust suddenly slip past each other along a fault or fault plane. The point where the earthquake originates below the surface is called the hypocenter, while the corresponding point directly above it on the Earth's surface is the epicenter. To measure earthquakes, scientists use scales such as the Richter scale, which quantifies magnitude based on the amplitude of seismic waves, and the Modified Mercalli Intensity scale, which assesses the effects and damage at specific locations. Understanding these measurements helps in assessing earthquake impacts and planning effective response strategies.

On March 27, 1964, at 5:36 pm an earthquake of magnitude 9.2 occurred in Alaska. The depth of Alaska's (64) earthquake was 25 km beneath the surface. The earthquake lasted approximately 4.5 minutes. It was followed by multiple Tsunamis recorded as high as 67 meters. It still holds the title of the most powerful recorded earthquake in the history of the United States. It is also the second-largest earthquake ever recorded, next to the M9.5 earthquake in Chile in 1960. This catastrophic event was a great leap forward to the modern age of earthquake science. Most of what we know about earthquakes can be traced back to the geological research done after the great Alaskan earthquake.

Challenges

When earthquakes occur in unprepared areas, several challenges arise. These include structural damage to buildings and infrastructure, leading to potential collapse and hindering rescue efforts. Loss of life and injuries can occur due to the sudden and violent shaking, particularly in densely populated areas without adequate emergency response systems. Disruption of critical services such as electricity, water supply, and communication networks further complicates rescue and relief operations. Secondary hazards like landslides, tsunamis, and aftershocks can compound the damage and pose additional risks. Limited resources and capacity in affected areas hamper effective response efforts, while psychological and socioeconomic impacts can be long-lasting for survivors.

Business Understanding

In February 6th, 2023, Turkey was struck by a devastating earthquake that was strongly felt in the southern and central parts of Turkey, the northern and western parts of Syria with mild yet deadly shaking in Cyprus, Israel, and Lebanon. In Turkey alone, there was widespread damage of property and loss of life with a death toll of more than 55,000 people. Further, more than 130,000 people were left injured and millions were displaced from their homes. With a magnitude of 7.8, the earthquake's impact was greatly felt on infrastructure with destruction amounting to billions of dollars.

An earthquake of such magnitude is a reminder of the need to be prepared in terms of forecasting earthquakes and the magnitude to be expected. Therefore, our project aims to enhance earthquake prediction and analysis to mitigate the impact of these natural disasters. By improving the accuracy of the earthquake prediction model through advanced time series forecasting techniques and optimizing model parameters, the project seeks to provide reliable warnings and timely information. Additionally, comprehensive analysis of earthquake patterns, including factors such as magnitude, depth, and time of occurrence, will enable the identification of significant trends and correlations. The project aims to provide actionable insights for improving earthquake preparedness and risk assessment, ultimately saving lives and minimizing damage.

Problem statement

Earthquakes pose a significant threat to human lives, infrastructure, and the environment. Timely and accurate prediction of earthquake occurrences can greatly assist in mitigating the impact of these natural disasters. However, predicting earthquakes with high precision remains a challenging task due to the complex nature of seismic events and the limited understanding of the underlying processes.

Data understanding

For this project, we analyzed the Significant earthquakes 1965-2016 dataset. The dataset consists of 23412 observations and 21 columns. With our target being the Magnitude column. We noticed a large amount of missing data in the columns:Depth Error,Depth Seismic stations,Magnitude error,Magnitude seismic stations,Azimuthal gap,horizontal distance,Horizontal error,Root mean square. With magnitude error having 98% of missing data. Our date and time were also recorded as object data-types and three rows had wrongly formatted date and time.

Data cleaning

In this section we involved fixing the wrongly formatted dates and combining the date and time columns . We then set the DateTime as our index. We also dropped the columns that had numerous missing data.

Data analysis

We explored the following areas:

1)Magnitude type frequency and MW(moment magnitude scale) was the most frequent.

2)Type. Earthquake had the most value counts of about 90% so we dropped the other values that weren't earthquakes, such as, Rock bursts and explosions.

3)Status. More than 70% if the earthquakes were reviewed , and a small percentage Automatic.

4)we plotted maps to show the location of the earthquakes around the world

5)We plotted a trend analysis of magnitude for decades and also created separate trends for every decade. It showed changes in magnitudes in every decade

6)We plotted trend analysis of depth for every decade and concluded that the depth at which the earthquake occurred bellow the earth surface is at its highest from 1980s to 1990s ,from then, there's a gradual drop towards 2010s

Data prepossessing

We checked for stationarity for the Magnitude and Depth column and both were stationary. We then performed a time-series split on our data. It is common practice not to split the data in time series but we did it so that we can later test for performance.

Feature Engineering

This sections entailed the following processes

- We first categorized our data into seasons: Winter, spring,summer and fall.

- We then generate binary features for the type column. This shows the types as earthquake,Nuclear Explosion, Explosion or rock Burst
- We classified the depth at which the earthquakes occurred below the earth surface as : Shallow, intermediate,deep or very deep
- We created a column that labels magnitudes as : Minor, light ,moderate or strong

Modeling

We created Facebook Prophet Model and an ARIMA model.
Our Facebook Model had an RMSE value Of 0.45 which indicated good performance of the model,Even though the data was big and we had a challenge having very clear plots on the same.

Our ARIMA Model had an RMSE 0.4456 . This indicated a good model performance too.

Our plots have not performed to expectations but with more research and additional tuning , we hope to build better models with better plots.

Recommendations

1. *Improve data collection and monitoring systems:* Enhance data collection and monitoring systems by deploying additional sensors, seismographs, and other monitoring equipment. This will improve the quality and quantity of earthquake data, enabling more accurate analysis and modeling.

2. *Invest in advanced prediction models:* Explore and invest in advanced prediction models to enhance the accuracy and reliability of earthquake forecasting.

3. *Enhance collaboration and information sharing:* Foster collaboration and information sharing among researchers, government agencies, and international organizations involved in earthquake research and preparedness. Establish platforms and networks for sharing data, insights, and best

practices, facilitating collaborative research efforts, and enabling a more comprehensive understanding of earthquake behavior and impacts.

4. *Strengthen infrastructure resilience:* Focus on enhancing the resilience of infrastructure in earthquake-prone areas. The potential damage and loss of life caused by earthquakes can be significantly reduced.

5. *Increase public awareness and education:* Promote public awareness and education initiatives to increase understanding of earthquake risks and preparedness. Develop educational programs, public campaigns, and community drills that provide guidance on earthquake response, evacuation plans, and emergency preparedness. Empowering individuals and communities with knowledge about earthquakes can save lives and mitigate the impact of future events.

Conclusions

1. *Earthquake data analysis provides valuable insights for risk assessment:* The analysis of earthquake data offers valuable insights into the patterns, trends, and characteristics of seismic activity. This information is crucial for assessing and understanding earthquake risks, informing mitigation strategies, and developing effective emergency response plans.

2. *Quality and availability of earthquake data are crucial for research:* The quality, quantity, and accessibility of earthquake data significantly impact the accuracy and reliability of research and modeling efforts. Therefore, continuous efforts should be made to improve data collection, standardization, and sharing mechanisms to support comprehensive and robust earthquake studies.

3. *Collaboration is essential for advancing earthquake research and preparedness:* Collaboration among researchers, government agencies, and international organizations is vital for advancing earthquake research, improving prediction models, and enhancing preparedness measures. Sharing knowledge, data, and expertise can lead to more effective earthquake risk reduction strategies and better-informed decision-making.

Next Steps

1. *Refine and optimize prediction models for better accuracy:* Further refine and optimize prediction models, considering factors such as additional variables, feature engineering, and ensemble techniques. Continuously evaluate and update models with new data to improve accuracy and robustness.

2. *Conduct in-depth statistical analysis to identify trends and factors:* Perform in-depth statistical analysis on earthquake data to identify long-term trends, recurring patterns, and factors influencing seismic activity. Investigate the relationship between seismic events and various geophysical, geological, and environmental variables to gain deeper insights into earthquake behavior.

3. *Develop an integrated early warning system using real-time data:* Build an integrated early warning system that leverages real-time data from seismic sensors and other monitoring devices. Implement algorithms and decision support systems that can quickly and accurately assess earthquake risks and issue timely warnings to potentially affected regions.

4. *Educate and engage the public in earthquake awareness and preparedness:* Launch comprehensive public awareness campaigns to educate communities about earthquake risks, safety measures, and preparedness strategies. Conduct drills, workshops, and training sessions to empower individuals and communities with the knowledge and skills necessary to respond effectively during earthquakes and minimize potential impacts.