Significant Earthquakes over the years

Analyzing Patterns, Trends, and Predictive Modeling

Simone Dag
EC Utbildning

Thesis
202405

Abstract

This project aims to analyze significant earthquakes from magnitude 5 and above, spanning from 1900 to the present, to identify patterns, trends, and potential predictors of earthquake occurrences globally. The study will explore various characteristics of earthquakes, including magnitude, location, depth, and frequency, to gain insights into seismic activity over time.

Table of Contents

1 Introduction

Understanding Earthquakes: Measurements, Causes, and Human Impact

Earthquakes, the powerful natural phenomena that have occurred on our planet since its inception, are measured and understood through methods like the Richter Scale and seismograph.

The Richter Scale is a numerical scale used to measure the magnitude of earthquakes. Developed by Charles F. Richter in 1935, it quantifies the amplitude of seismic waves generated by earthquakes. It's important to note that the Richter Scale measures the energy released by an earthquake, not the damage it causes. This means that two earthquakes registering a magnitude of 6 on the Richter Scale can have vastly different impacts depending on factors such as depth, distance from populated areas, and local soil conditions. (Stefan Johansson., 2023)

Richter scale spans:

1. Less than 2.0 (Micro): Micro earthquakes are very subtle for people to feel but detected by instruments. They occur frequently, with an estimated 1.4 million happening annually worldwide.
2. 2.0 – 2.9 (Minor): Minor earthquakes occur often but rarely result in damage. Approximately 1.3 million of these events happen each year.
3. 3.0 – 3.9 (Light): Light earthquakes are frequently felt but seldom cause significant damage. Around 130,000 of these earthquakes happen annually.
4. 4.0 – 4.9 (Moderate): These earthquakes significantly shake indoor items and are followed by rattling noises. Remarkable damage is unlikely to happen and occur around 13,000 globally each year.
5. 5.0 – 5.9 (Strong): Strong earthquakes have the potential to cause major damage to buildings and other structures and happen around 1,300 annually.
6. 6.0 – 6.9 (Major): Major earthquakes cause substantial damage in populated areas. About 100 such events occur each year.
7. 7.0 and higher (Great): These earthquakes cause severe damage. They occur around 10-20 times per year globally. Typically, there is only one earthquake per year with a magnitude between 8 and 10. No earthquake with a magnitude of 10 or higher has ever been recorded. (Anne Helmenstine ,. 2023)

Tsunami in Thailand 2004 and the earthquake in Japan 2011 were both around magnitude 9.

Seismographs detect and record even the smallest vibrations in the ground, providing invaluable data for monitoring earthquakes. These instruments help determine the earthquake's epicenter and depth, crucial for understanding its impact. (Stefan Johansson., 2023)

Earthquakes arise from the movement of tectonic plates, volcanic eruptions, or human activities like mining. Understanding these causes is vital for mitigating their devastating effects. While humans can't directly cause earthquakes, activities like fracking or dam construction may indirectly contribute to seismic activity. However, natural processes primarily drive earthquakes, with human settlements exacerbating their impact. By comprehensively measuring, understanding, and preparing for earthquakes, we can better protect communities and minimize the risks of future disasters. (Herman Larsson., 2023).

1.1 Incorporating Machine Learning

Traditional methods for analyzing and predicting earthquakes often fall short in capturing complex, non-linear relationships in seismic data. This study employs machine learning techniques to enhance predictive accuracy and provide deeper insights into earthquake behaviors. Four models—Linear Regression, Random Forest Regression, Gradient Boosting Regression, and XGBoost Regression—are used to analyze and predict earthquake magnitudes and depths, offering both interpretable results and high predictive accuracy.

1.2 Research Questions

The purpose of this study is to explore and understand the patterns and trends of significant earthquakes worldwide and to develop predictive models for future earthquake occurrences.

I will explore that by answering the following questions:

1. What are the temporal and spatial patterns of significant earthquakes over the past century?
2. Can predictive models be developed to forecast future earthquake events with reasonable accuracy?

2 Theoretical Framework

In this project, I utilized various machine learning models to predict earthquake occurrences, specifically employing random forest regression, gradient boosting regression, XGBoost regression, and linear regression described below. These models were chosen for their effectiveness in handling diverse data types and their robustness in capturing patterns within the data. While linear regression provides a baseline with its simplicity, the ensemble methods (random forest, gradient boosting, and XGBoost) are particularly adept at capturing complex relationships. To evaluate the performance of these models, I used the Root Mean Square Error (RMSE) metric, which measures the average deviation of predicted earthquake magnitudes from the actual values. Additionally, I implemented cross-validation with a 5-fold split to ensure the reliability and robustness of the model evaluations.

2.1 Machine Learning

Let's begin by understanding what a machine learning (ML) model or algorithm is and how it operates. Although you might not realize it, we encounter them frequently in our daily lives, embedded in various applications. For instance, the recommendation systems on platforms like Netflix or Spotify are driven by such models.

Machine learning is a subset of artificial intelligence (AI), consisting of sophisticated algorithms designed to recognize patterns and make predictions, continuously improving as they process new data. They adapt and improve over time as they encounter new data.

ML works by inputting data to your written algorithm, in this case recorded earthquakes over time with their location, magnitude and depth. The model is first trained on a subset of this data, allowing the ML algorithm to learn patterns from it. Next, the model is tested on a separate dataset that was not used during training, to evaluate its performance. After training and testing, we can assess the model's performance using various metrics and evaluation methods, such as cross-validation and RMSE.

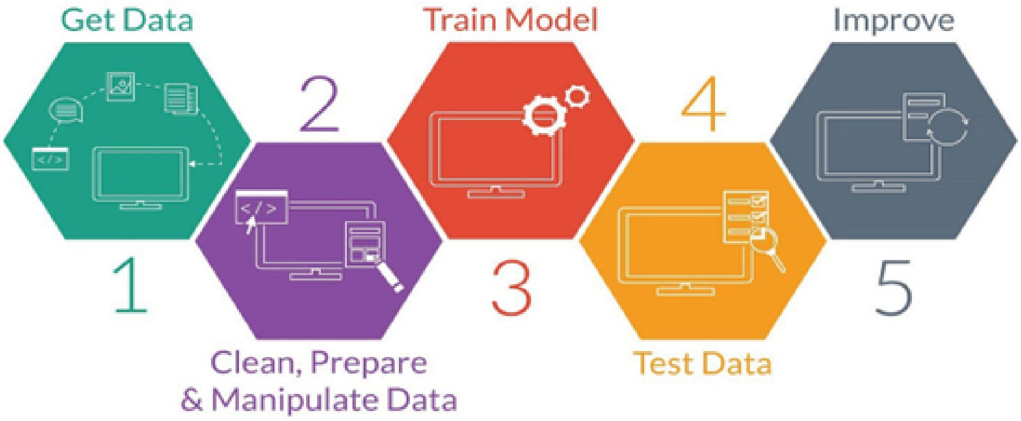

Figure 1: Machine learning process flow.

2.2 Regression models

Regression is a method used to determine a relationship between an outcome and a set of independent variables (i.e. inputs). A variable represents a piece of data; in this case, one variable is the earthquake magnitude. Regression is used in both machine learning and statistical modeling to predict future outcomes based on these variables. (Sagar Shukla ,. 2024)

2.2.1 Linear Regression

Linear regression is a type of supervised machine learning algorithm which learns from labeled datasets to map data points to optimized linear functions. Supervised machine learning involves algorithms learning from labeled data, where the target values are already known. It includes:

- Classification: Predicts categorical outcomes (e.g., determining if an animal in an image is a cat or dog).
- Regression: Predicts continuous outcomes (e.g., forecasting house prices based on parameters like age, location, and area).

Linear regression falls under regression, predicting continuous output variables from independent input variables meaning it finds the best straight line to predict outcomes based on data with continuous variables. (Mohit Gu,. 2024)

2.2.2 RandomForestRegression

RandomForestRegressor is an algorithm that combines predictions from many decision trees. Decision trees are like flowcharts used to make decisions. They follow a series of questions to reach conclusions. RandomForestRegressor creates lots of these decision trees during training and then combines their predictions. (sklearn.ensemble.RandomForestRegressor,. 2007-2024)

2.2.3 Gradient Boosting

Gradient boosting is a powerful technique in machine learning where multiple weak learners are combined to create a single, highly accurate strong learner. Weak learners are models that are slightly better than random guessing, like a decision tree that performs just above 50-60% accuracy. By combining many of these weak learners, gradient boosting can achieve over 95% accuracy on various tasks. Decision trees are often chosen as weak learners due to their versatility with different datasets. (Bex Tuychiev,. 2023)

2.2.4 XGBoostRegression (Extreme Gradient Boosting)

XGBoost is a powerful algorithm for regression and classification tasks. It's highly accurate due to its ability to capture complex relationships in data. It's efficient and scalable, making it ideal for large datasets. It works by combining multiple decision trees where each tree corrects errors of the previous ones. (Alexis Cook)

2.2.4.1 Hyperparameter optimization on XBG

For the XGB model I used a library called Optuna which is a tool that automatically tunes different hyperparameters like n_estimatos, max depth. N_estimators are the number of trees in the forest, meaning how many individual opinions do we gather before deciding. This method can help to improve the overall decision making but it can also take time to process. Max depth refers to the maximum depth allowed for each decision tree in the ensemble. The depth of a decision tree indicates how many splits it can make before reaching a leaf node, where a prediction is made.

Optuna runs through multiple trials within the intervals given and adjusts hyperparameters like n_estimators and max_depth to find the best combination that could help to minimize the RMSE. (Mario Filho,. 2023)

In this trial run, Optuna found that n_estimators should be set to 246 with a max depth of 10.

2.3 Comparison of Gradient Boosting and Extreme Gradient Boosting

2.3.1 Optimization

Gradient Boosting and XGBoost share similar methodologies but differ significantly in their optimization techniques. Gradient Boosting relies on the first derivative, which indicates the rate and direction of change in a function. In contrast, XGBoost utilizes the second derivative, providing insights into how the rate of change itself is changing. This additional layer of information allows XGBoost to optimize more effectively.

2.3.2 Regularization

XGBoost enhances Gradient Boosting by incorporating advanced regularization methods, specifically L1 and L2 regularization. This feature helps improve the model's generalization capabilities, reducing the risk of overfitting. (Gajendra., 2022)

L1 regularization, also known as Lasso, adds the absolute values of the coefficients as a penalty term to the loss function, promoting sparsity by shrinking some coefficients to zero. L2 regularization, or Ridge, adds the squared values of the coefficients as a penalty term, discouraging large coefficients and thus reducing model complexity. (Anuja Nagpal., 2017)

2.3.3 Advantages

Both algorithms perform well with large datasets. They can handle nonlinear, non-monotonic data and segregated clusters without the need for normalized features. They typically achieve high accuracy.

2.3.4 Disadvantages

These models can overfit, particularly if the trees are excessively deep and the data is noisy. They can be computationally intensive, requiring a substantial number of trees, which can be both time and memory consuming. (Gajendra., 2022)

2.4 Model Evaluation

Understanding model performance is crucial, and there are various methods to evaluate it effectively. I used RMSE and cross-validation techniques to ensure a robust assessment of my models' performance.

2.4.1 Cross-validation Kfold=5

Let start with the cross-validation technique where I used 5 K-folds to assess the performance of the models. This technique helped in evaluating the model's performance more robustly and identifying any potential issues such as overfitting or underfitting.

Figure 2 illustrates Cross-Validation with 5 folds. In this process, the dataset is divided into 5 equal parts, or folds. The model is trained on 4 of these folds and validated (tested) on the remaining fold. This process is repeated 5 times, each time with a different fold being used as the validation set and the other 4 as the training set. This method provides a robust estimate of the model's performance on unseen data and helps prevent overfitting.

Overfitting occurs when a model learns the training data too well, capturing noise and details that do not generalize to new data, leading to poor performance on unseen data. Conversely, underfitting happens when the model is too simple to capture the underlying patterns in the data, resulting in poor performance on both the training and validation sets. Cross-validation helps to find a balance, improving the model's ability to generalize to new data. (sklearn, Cross-validation: evaluating estimator performance).

You can think about overfitting like memorizing a textbook instead of understanding, whereas underfitting is like barely studying and missing the key concepts.

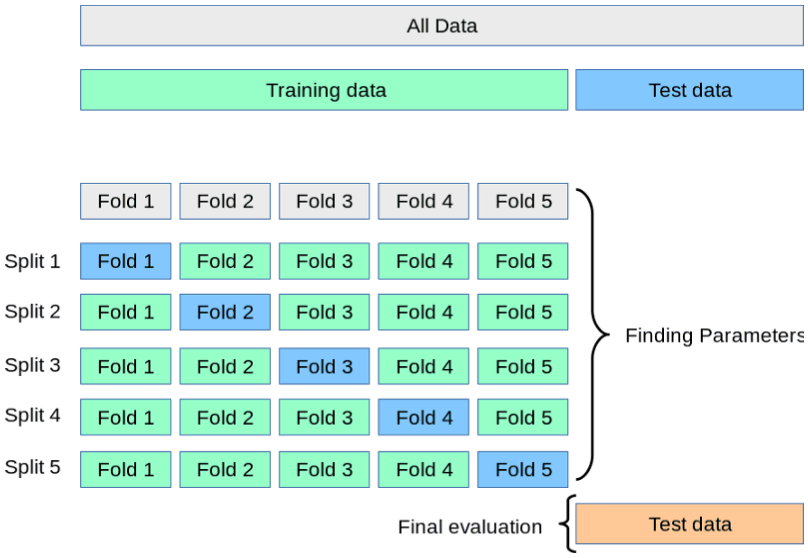

Figure 2: Cross-validation with 5 Kfolds

2.4.2 RMSE

Another important thing during modeling is to understand how well your model performs by calculating the accuracy. In this study, I have used the metric called Root Mean Square Error (RMSE) which calculates the differences between predicted and actual values, squares these differences, averages them, and then takes the square root of this average. The closer the RMSE value is to zero, the better the model's accuracy. (Jason Brownlee,. 2021)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(d_i - p_i\right)}^{2}$$

Figure 3: Mathematical equation of how RMSE is calculated.

2.5 Feature Importance Analysis

Understanding which features have the most significant impact on the model's prediction is essential for interpreting and improving model performance. I have used techniques like permutation importance and SHAP (SHapley Additive exPlanations) values to investigate the feature importance which will be explained in sections below with figures. These analyses help to identify which variables contribute most to the accuracy of the model's predictions.

2.5.1 Permutation importance

Permutation importance, see figure 4, shows us the impact of each feature on the model's performance by measuring how much the model's accuracy decreases when the feature's values are randomly shuffled. This is helping to identify the most influential features in the model. The higher the value, the more important the feature is for the model's predictions. These values indicate the relative importance of each feature in predicting the target variable. (sklearn, Permutation feature importance., 2007-2004

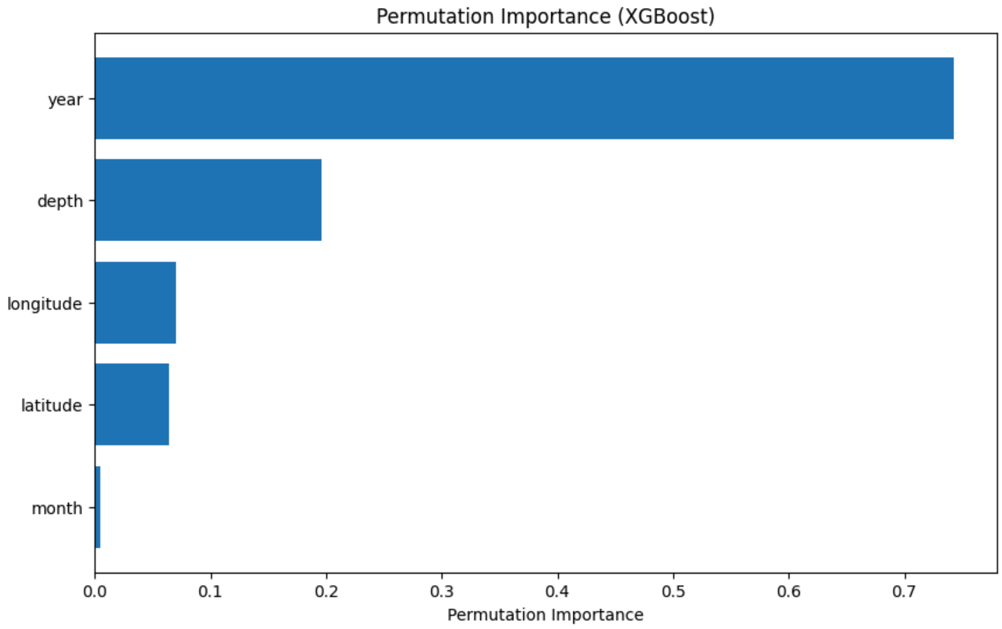

Figure 4: Permutation Importance for XGB.

2.5.2 SHAP

SHAP (SHapley Additive exPlanations) values, shown in Figure 5, offer an explanation of how each feature affects individual predictions. Each SHAP value represents the contribution of a feature to the difference between the actual prediction and the average prediction. Features with larger SHAP values have a more significant impact on the prediction, either positively or negatively. This visualization helps understand the model's decision-making process at a detailed level, highlighting how different values of a feature influence the predicted outcome.

By using SHAP values, we can make the model's predictions more transparent and understand the reasons behind specific predictions more clearly. This helps in better interpreting the machine learning model's behavior and decisions. (Abld All Awan,. 2023)

The features are ranked by their impact on the model's predictions, with the most influential feature at the top where each dot on the plot represents a data point from the dataset.

The x-axis represents the SHAP value. A negative SHAP value indicates that the model tends to underestimate the earthquake magnitude, while a positive SHAP value means the model tends to overestimate the magnitude. In the plot, red dots represent high feature values, and blue dots represent low feature values. The spread of data points shows their impact on the model's predictions: widely spread points indicate a significant impact, whereas points clustered around zero indicate a neutral impact.

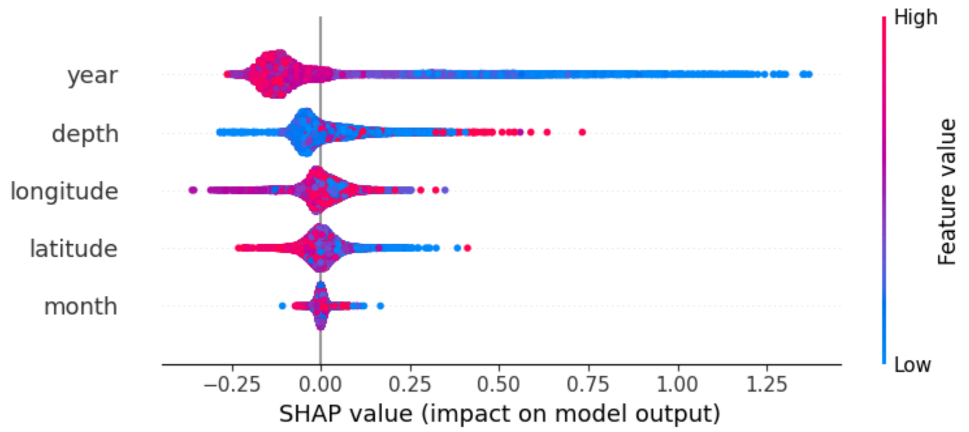

Figure 5: SHAP values for XGB as a beeswarm plot.

3 Methodology

3.1 Data mapping

To conduct my study, I obtained a comprehensive dataset from Kaggle, containing information on significant earthquakes worldwide from 1900 to present with a magnitude of 5 or above. This dataset, sourced from the United States Geological Survey (USGS), includes details such as location, date and time, magnitude, depth, and other relevant information about each earthquake. I preprocessed the data by retaining relevant columns, handling missing values, and extracting temporal features from the timestamp.

3.1.1 Data Preprocessing

Data preprocessing is crucial for ensuring the quality and suitability of the data for modeling. Using techniques like handling missing values, feature engineering (extracting temporal features from the timestamp), and scaling/normalization (using StandardScaler) are applied to prepare the data for modeling.

Normalization ensures that all features have a similar scale, preventing some features from dominating others. For instance, if we have features with different scales, some ranging from 10 to 1000, and others from 0.1 to 10, failing to normalize them could lead the model to focus on features with larger values. This imbalance could result in a poorly performing model that doesn't accurately represent the dataset. Normalization helps to level the playing field by giving equal importance to all features, thereby improving the model's performance and fairness in handling the data. (sklearn.preprocessing.StandardScaler., 2007-2004)

3.1.2 Data Exploration and Visualization

The initial phase involved exploring the dataset to better understand its structure, distribution, and relationships between variables. I have used different visualization techniques like histograms, bar plots, and line plots to analyze and understand the different aspects of the data, such as the distribution of earthquake magnitudes, occurrence over time, geographic distribution, and relationships between variables like magnitude, depth, latitude, and longitude.

3.1.2.1 Distribution Visualization

Figure 6 map presents a comprehensive overview of earthquakes registering a magnitude of 5 or higher across the globe since 1900. Each plotted point on the map signifies the precise epicenter coordinates of an earthquake. The map shows that earthquakes are not evenly distributed around the world but follow a specific pattern, with regions like the Pacific Ring of Fire and the Himalayas experiencing more frequent and severe earthquakes

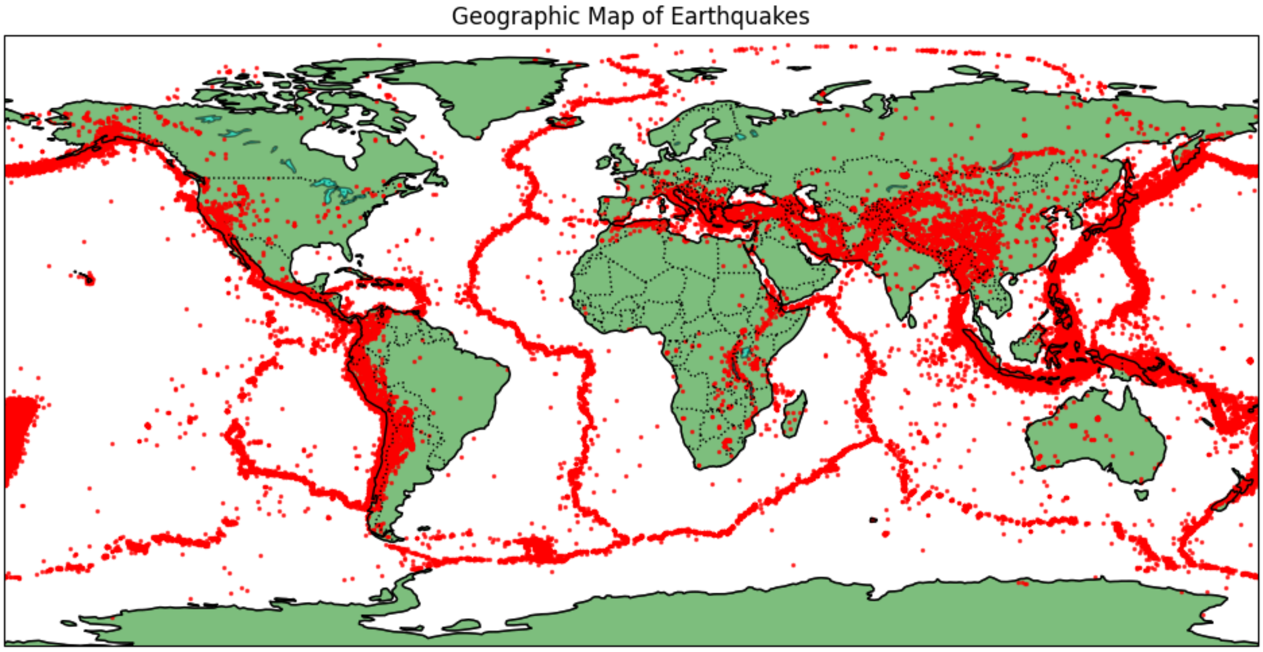

Figure 6: Geographic map of earthquakes occurring since 1900 to present.

3.1.2.2 Line plot

Figures 7 and 8 portray the trends in both earthquake magnitudes and depths over the years. The line plots showcase how these seismic attributes have evolved annually. The x-axis represents the years, while the y-axis displays the corresponding magnitudes or depths of earthquakes. Observing these trends helps in understanding the temporal distribution of seismic characteristics and identifying potential patterns or shifts in earthquake activity across different years.

The shadow shows how much the earthquake magnitudes and depths fluctuate from year to year. A wider shadow indicates higher variability, while a narrower shadow indicates more consistent values.

In Figure 7, the declining shadow and line indicate not only a reduction in average magnitudes but also a decrease in variability, suggesting more predictability in earthquake magnitudes over time. In Figure 8, the stable shadow width and line indicate consistent depths and variability over the years, with no significant trend in either direction.

The differences in the trends and shadow behaviors between Figures 7 and 8 highlight how earthquake magnitudes and depths exhibit different temporal patterns and variabilities. This information is critical for understanding the nature of seismic activities and improving earthquake prediction models.

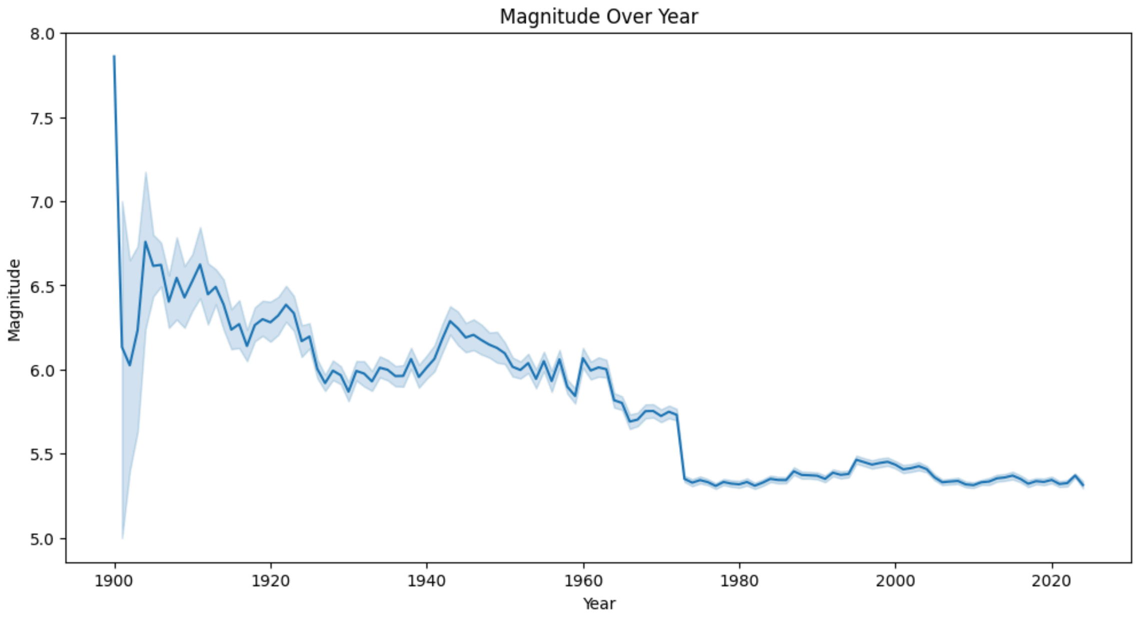

Figure 7: Magnitude over the years.

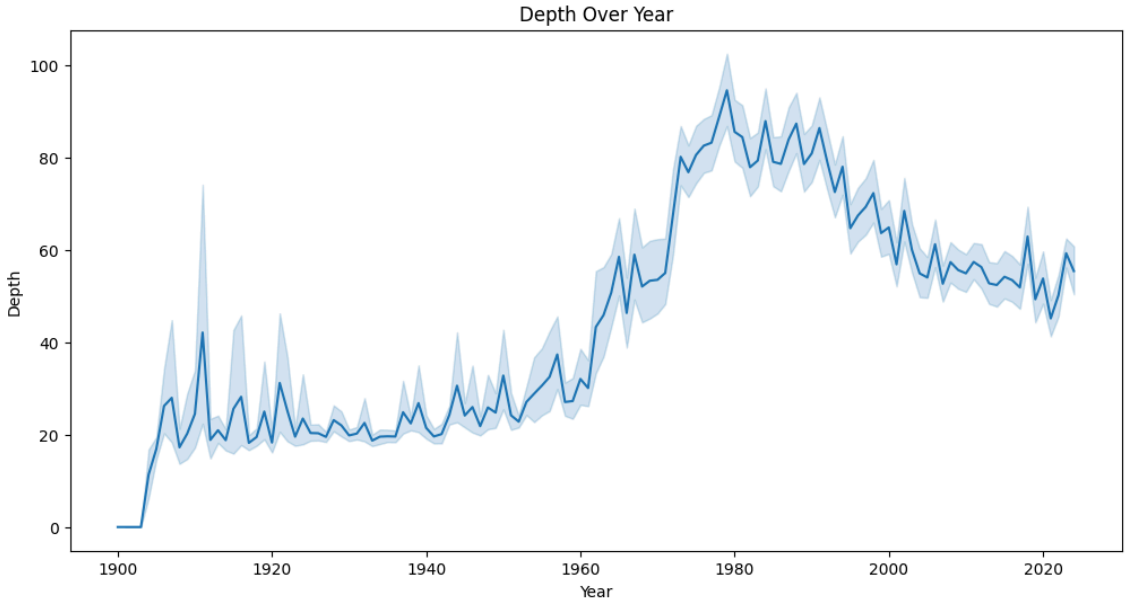

Figure 8: Depth of the earthquakes over the years in km.

3.1.2.3 Histograms

Looking at figure 9, we can see the distributions of magnitude, depth, latitude, and longitude of significant earthquakes. It shows the frequency of earthquakes occurring at different magnitudes which gives us insight into the characteristics and patterns of earthquake occurrences across magnitudes, depths, and geographical locations.

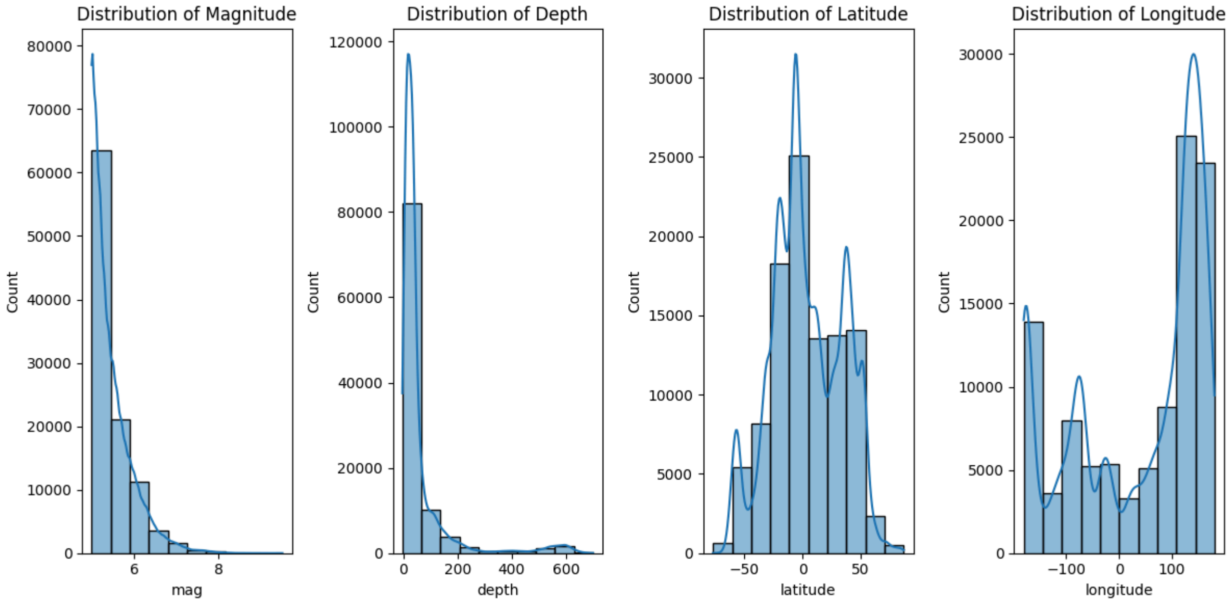

Figure 9: Distributions of magnitude, depth, latitude, and longitude of significant earthquakes.

3.1.2.4 Bar plot

Figure 10 displays the count of earthquakes for each year. Each bar represents a year, with the height indicating the corresponding number of earthquakes.

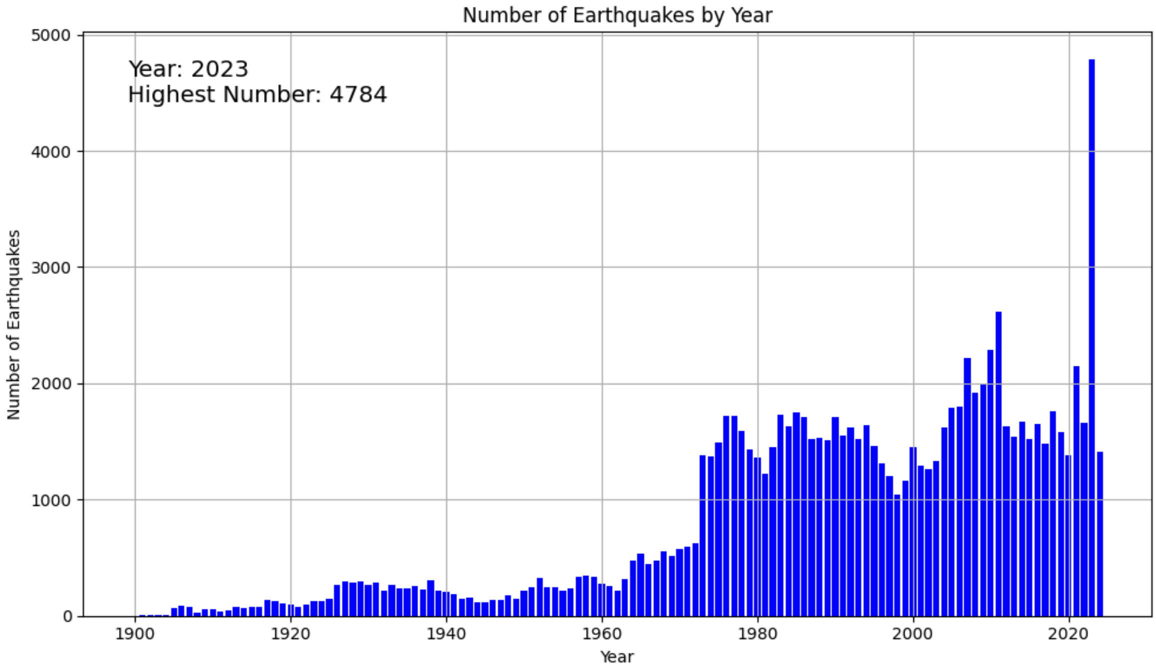

Figure 10: Number of earthquakes per year

4 Results and Discussion

Looking at the RMSE values for each model below, XGBoost outperformed the other models, achieving the lowest RMSE. After tuning the XGBoost model, I observed a slight further improvement in performance, with the RMSE decreasing to 0.389.

Linear regression was chosen as a baseline model due to its simplicity and interpretability. It provides a straightforward way to understand the relationship between the input features and the target variable. However, its performance is limited when dealing with complex, non-linear relationships.

Random forest regression was included because of its ability to handle high-dimensional data and its robustness against overfitting. By averaging the predictions of multiple decision trees, random forest reduces the variance of the model, leading to more stable and reliable predictions.

Gradient boosting regression was selected for its strength in improving predictive accuracy through sequential learning. Each tree in the sequence focuses on correcting the errors of the previous trees, which helps in capturing complex patterns in the data. This makes gradient boosting a powerful tool for achieving high accuracy.

XGBoost was chosen due to its superior performance and efficiency. It builds on the principles of gradient boosting but includes additional enhancements such as regularization, efficient handling of missing data, and parallel processing. These features make XGBoost particularly effective for large datasets and complex prediction tasks.

RMSE for the different models	
Linear Regression	0.444
RandomForest Regression	0.398
GradientBoosting Regression	0.401
XGBoost Regression before tuning hyper parameters	0.393
XGBoost Regression after tuning hyper parameters	0.389

Figure 11: RMSE values of the models.

Figure 12 visualizes the observed and predicted magnitudes of earthquakes over time. Each blue point represents the actual magnitude of earthquakes recorded in the test set, serving as the ground for evaluating the predictive performance of the model. The red dots represent the magnitudes predicted by the trained XGBoost model. To ensure clarity, the figure includes only 50 observations for a clear visual comparison. The plot highlights the model's ability to approximate earthquake magnitudes. When the red dots align closely with the blue points, it indicates that the model's predictions closely match the actual observations. Conversely, deviations between the red and blue points may suggest areas where the model performs less optimally, potentially due to less recorded data for higher magnitudes.

This comparison aids in assessing the model's efficacy in earthquake magnitude prediction, which is crucial for seismic risk assessment and disaster preparedness efforts.

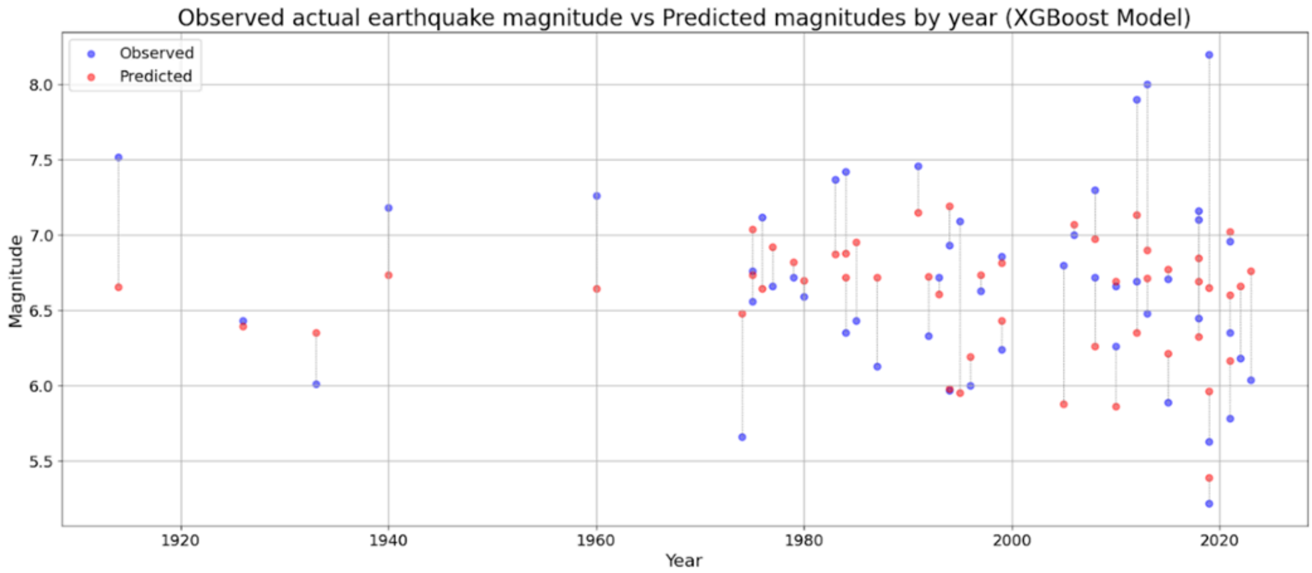

Figure 12: Comparison between observed and predicted earthquake magnitudes

Based on the theory and methods applied for this dataset, I have also conducted visualizations of predicted earthquake occurrences in 2024, as shown in figures 13, 14, and 15.

The combined visualization offers insights into the anticipated seismic activity for the year 2024. The map (figure 13) visualization represents the predicted earthquake occurrences across geographical locations. Each dot on the map symbolizes a forecasted earthquake event, with the color indicating the predicted magnitude.

The histogram (figure 14) displays the distribution of predicted earthquake magnitudes and depths density for the same time period. The histogram illustrates the frequency of forecasted magnitudes and depths, offering a summary of the range and distribution of predicted seismic activity.

Additionally, the timeline plot (figure 15) depicts the timeline of predicted earthquake occurrences. Here, each point on the plot corresponds to a forecasted earthquake event, with the x-axis representing time in days and the y-axis representing magnitude. This visualization allows for an understanding of when the earthquakes are predicted to occur throughout the year.

Together, these visualizations provide a comprehensive overview of the anticipated earthquake occurrences, magnitudes, and depths for the year 2024. They serve as valuable tools for earthquake preparedness and risk assessment efforts, aiding in understanding and mitigating potential seismic hazards.

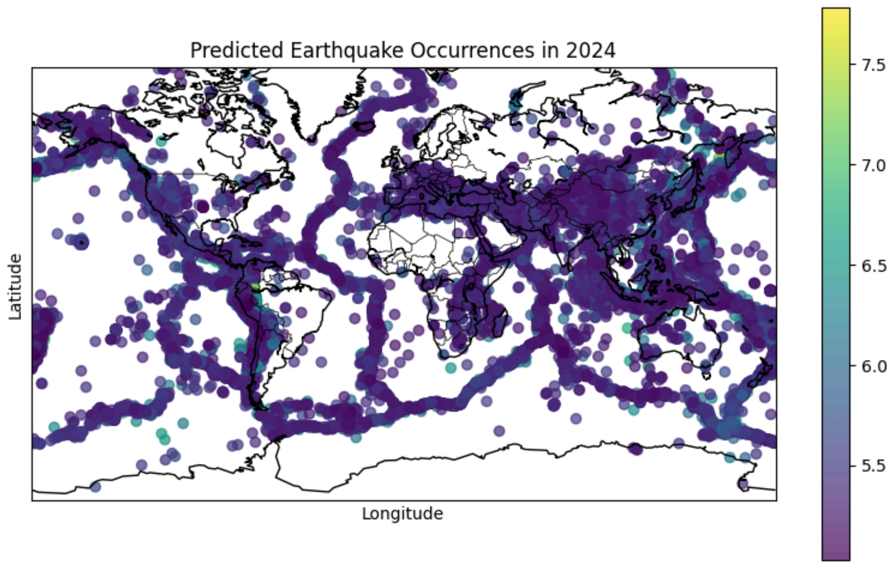

Figure 13: Predicted earthquake locations, latitude and longitude for 2024.

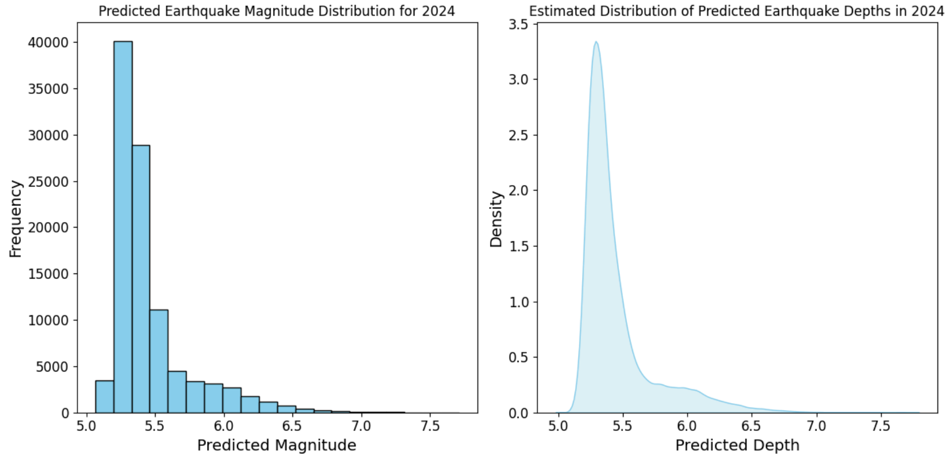

Figure 14: Predicted earthquake magnitudes and depth density for 2024

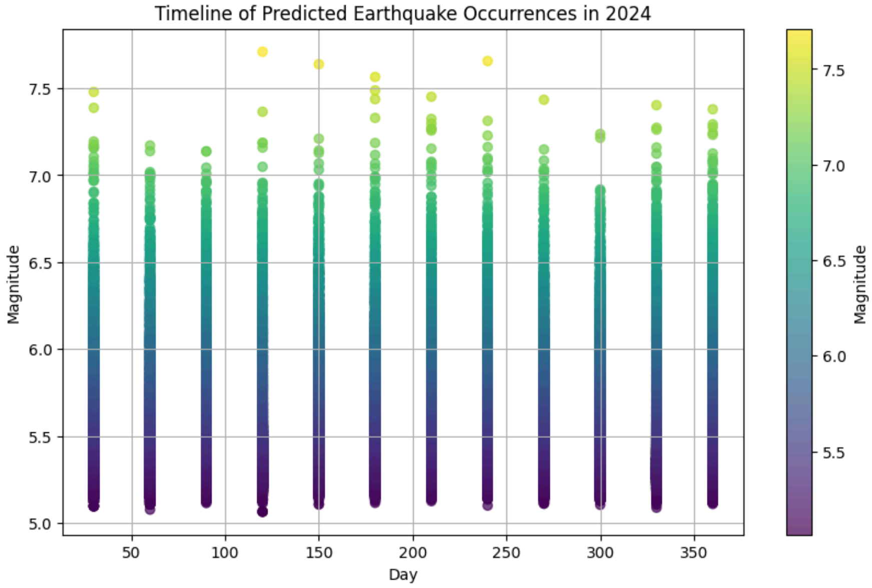

Figure 15: Predicted earthquake magnitudes and depth density for 2024

While the study provides valuable insights into earthquake prediction, it's important to acknowledge some limitations. The dataset, primarily sourced from the United States Geological Survey (USGS) via Kaggle, may not encompass all global seismic events. This limitation could introduce biases or gaps, potentially affecting the generalizability of the findings.

Even if the machine learning models, specifically XGBoost, shows good predictive performance, they might oversimplify the intricate dynamics of earthquakes. Exploring additional features or advanced modeling techniques could enhance prediction accuracy and robustness.

These limitations highlight opportunities for future research to improve the accuracy and applicability of earthquake prediction models.

5 Future work

Future improvements and areas for future work in earthquake prediction research could include the incorporation of additional data sources such as satellite imagery, geological surveys, and climate data to provide a more comprehensive understanding of seismic activity.

Exploring advanced machine learning techniques within deep learning, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), along with other machine learning models, could uncover more complex patterns in earthquake data and enhance predictive capabilities. This comprehensive approach allows for leveraging diverse algorithms to improve the accuracy and robustness of predictions.

Establishing a long-term monitoring and evaluation framework is crucial for continuously assessing the performance and effectiveness of earthquake prediction models over time, enhancing the reliability and usability of predictive tools in mitigating seismic risks.

By addressing these future improvements and considering ongoing research challenges, the field of earthquake prediction can advance towards more accurate, reliable, and actionable forecasting capabilities, ultimately contributing to enhanced disaster resilience and community safety.

6 Conclusion

This comprehensive study delves into the realm of earthquake prediction, seeking to uncover patterns, trends, and predictors of seismic events worldwide. By addressing the core questions outlined in the introduction, this research aims to deepen our understanding of earthquake dynamics and contribute to the development of predictive models for future seismic occurrences.

Through thorough data exploration and visualization techniques, this study reveals the temporal and spatial patterns of significant earthquakes over the past century. By examining attributes such as magnitude, depth, location, and frequency, it provides insights into the evolution of seismic activity, shedding light on the complex dynamics of earthquake occurrences.

Utilizing machine learning algorithms, including linear regression, random forest regression, gradient boosting regression, and XGBoost regression, the study develops predictive models to forecast earthquake events. Evaluation metrics such as root mean squared error (RMSE) offer quantitative assessments of model performance, with XGBoost emerging as the most effective model. Further refinement through hyperparameter tuning enhances predictive accuracy, showcasing the effectiveness of machine learning in earthquake prediction.

Despite the progress in earthquake prediction, significant challenges persist. Managing the vast volume of data and ensuring its quality and reliability pose considerable obstacles. Moreover, the intricate nature of seismic activity presents challenges in accurately modeling and predicting future earthquakes.

Looking ahead, continued research and refinement of predictive models are essential. Incorporating additional features and exploring advanced modeling techniques can bolster prediction accuracy and deepen our understanding of seismic events. Collaboration between researchers, data scientists, and seismologists will be pivotal in overcoming these challenges and advancing the field of earthquake prediction.

In conclusion, this study demonstrates the effectiveness of machine learning models in predicting earthquake occurrences. By leveraging these models and analyzing feature importance, we can enhance our understanding of earthquake patterns and trends, ultimately contributing to better disaster preparedness and mitigation strategies. Further research could focus on incorporating additional features or exploring advanced modeling techniques to improve prediction accuracy further.

The purpose of this study was to explore and understand the patterns and trends of significant earthquakes worldwide and to develop predictive models for future earthquake occurrences. Specifically, I sought to answer the following research questions:

1. What are the temporal and spatial patterns of significant earthquakes over the past century?
2. Can predictive models be developed to forecast future earthquake events with reasonable accuracy?

As the result and discussion is concluding, the answer to the research question is that significant earthquakes exhibit clear temporal and spatial patterns and can be systematically analyzed. Furthermore, the study demonstrated that predictive models, particularly those utilizing advanced machine learning techniques like XGBoost, can achieve reasonable accuracy in forecasting future earthquake events. This research thus provides a foundational step towards more accurate and reliable earthquake prediction, contributing to enhanced disaster preparedness and mitigation efforts.

7 Appendix

8 Bibliography

Abld All Awan (2023), An Introduction to SHAP Values and Machine Learning Interpretability
https://www.datacamp.com/tutorial/introduction-to-shap-values-machine-learning-interpretability
Accessed 2024-05-09

Alexis Cook, XGBoost
https://www.kaggle.com/code/alexisbcook/xgboost
Accessed 2024-05-07

Anuja Nagpal (2017), L1 and L2 Regularization Methods
https://towardsdatascience.com/l1-and-l2-regularization-methods-ce25e7fc831c
Accessed: 2024-05-21

Anne Helmenstine (2023), Richter Scale and Earthquake Magnitude
Richter Scale and Earthquake Magnitude (sciencenotes.org)
Accessed 2024-05-14

Bex Tuychiev (2023), A Guide to The Gradient Boosting Algorithm
https://www.datacamp.com/tutorial/guide-to-the-gradient-boosting-algorithm
Accessed 2024-05-07

Cross-validation: evaluating estimator performance
https://scikit-learn.org/stable/modules/cross_validation.html
Accessed 2024-05-09

Gajendra (2022), Gradient Boosting & Extreme Gradient Boosting (XGBoost)
https://medium.com/@gajendra.k.s/gradient-boosting-extreme-gradient-boosting-xgboost-de865b871203
Accessed 2024-05-21

Herman Larsson (2023), Varför uppkommer jordbävningar? | En djupdykning i geologiska fenomen (in swedish),
https://wvwv.se/varfor-uppkommer-jordbavningar-en-djupdykning-i-geologiska-fenomen/
Accessed 2024-05-14

Jason Brownlee (2021), Regression Metrics for Machine Learning
https://machinelearningmastery.com/regression-metrics-for-machine-learning/
Accessed 2024-05-07

Mario Filho (2023), XGBoost Hyperparameter Tuning With Optuna (Kaggle Grandmaster Guide)
https://forecastegy.com/posts/xgboost-hyperparameter-tuning-with-optuna/
Accessed 2024-05-07

Mohit Gu (2024), Linear Regression in Machine learning
https://www.geeksforgeeks.org/ml-linear-regression/?ref=next_article
Accessed 2024-05-07

Permutation feature importance
https://scikit-learn.org/stable/modules/permutation_importance.html
Accessed 2024-05-09

Sagar Shukla (2024), Regression in machine learning
https://www.geeksforgeeks.org/regression-in-machine-learning/
Accessed 2024-05-07

sklearn.ensemble.RandomForestRegresso, (2007-2004)
https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html
Accessed 2024-05-07

sklearn.preprocessing.StandardScaler, (2007-2004)
https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html
Accessed 2024-05-09

Stefan Johansson (2023), Seismografi och Richterskalan - Hur vi mäter jordbävningar (in swedish),
https://hurfungerar.se/vetenskap/geologi/seismografi-och-richterskalan
Accessed 2024-05-14