École Polytechnique Fédérale de Lausanne

Earthquake Prediction using Deep Learning Algorithms

by Fatih Mutlu

Master Semester Project Report

Approved by the Examining Committee:

Dr. Gaëtan Frusque
Project Supervisor

Prof. Dr. Olga Fink
Head of Laboratory

EPFL ENAC IIC
IMOS
GC A2 402 (Bâtiment GC)
Station 18
CH-1015 Lausanne

June 9, 2023

Acknowledgments

First of all, I would like to express my deepest gratitude to my supervisor, Dr. Gaëtan Frusque. His guidance was indispensable and crucial throughout the project. His listening, knowledge, intuition, insights and directions provided me with the ideal conditions to stay on track and move forward during this academic journey. The experience that I have gained under his supervision will be of great value for my future projects.

I would also like to express my profound thanks to Prof. Olga Fink, the Head of Laboratory. Her faith in the significance and potential of this project was essential for its realization.

Furthermore, I am extremely recognizing for the source of knowledge made accessible by the current and past researchers. Their work and publications gave me the possibility to structure and conduct my work in the best possible way.

On a more personal note, I want to acknowledge the support of my close relatives, friends and family. Their encouragement was of great importance in times of challenge and difficulty.

Lausanne, June 9, 2023 Fatih Mutlu

Abstract

This project explores the application of deep learning models for earthquake prediction. USGS's Earthquake Catalog is used to collect raw earthquake data from all regions of the world that have magnitude $M \geq 2.5$. After a thorough review of previous research in this field, the study attempts to replicate and improve some of the results using both classification and regression types of predictions and recent deep-learning architectures such as LSTM. While it is to some extent successful in this task, it also reveals certain limitations showing the complexity inherent to seismic activities.

Contents

Chapter 1

Introduction

Eartquakes are a natural geological events that are the result of the release of seismic energy due to the pressure accumulating in the friction of tectonic plates. While earthquakes do occur on a daily basis, the majority of them are of small magnitude, typically measuring less than 3 on the Richter scale. These smaller earthquakes often do not cause significant damage. However, larger earthquakes can be extremely destructive and have sometimes had catastrophic societal and economical consequences. It is therefore crucial to have reliable methods to be able to predict these.

Traditional methods rely on the observation of accessible natural phenomena such as sea level changes and animal behaviour. Modern methods use available technical data. Key techniques include:

- Observation of seismicity patterns

  Statistical methods can be employed to analyze seismicity patterns provided by the seismic waves and identify notable changes on them a few days before an earthquake (Vallianatos et al., 2021). However, these patterns can vary from region to region and from event to event (H. Kanamori, 1981), limiting their practical use in earthquake prediction.

- Satellite observation

  Satellite data can be helpful: we can monitor strain accumulation with high accuracy through geodetic observations (Ekbal Hussain, 2018) and changes in ground temperature through the measurement of infrared radiation to monitor the movement of plates and make predictions based on this.

- Machine Learning and Artifical Intelligence

  Modern state-of-the-art approaches rely on the processing of large data do detect long-term patterns that humans may not be able to identify alone. It is now possible to provide the models with large datasets from various sources to enhance the performance, since they have the ability to learn complex patterns and relationships that may not be immediately apparent.

Chapter 2

Plan

In this project, we're going to use the 2013 paper "Neural networks to predict earthquakes in Chile" published in Applied Soft Computing by Reyes et al. as a starting point. The authors developed a multilayer perceptron neural network model to predict earthquakes by using a set of features based on seismicity laws and magnitude patterns. They are in the number of seven which is a relatively small number for such predictions. We'll investigate the integration of additional relevant features, the use of larger datasets or more recent models in order to improve the results.

Ultimately, we're aiming to replicate, evaluate and ideally improve the results of the 2023 paper "An attention-based LSTM network for large earthquake prediction" [1] by Berhich et al. who developed a prediction model based on state-of-the-art methods for timeseries prediction.

To make an earthquake prediction, we need three elements : the time, magnitude and location of its anticipated occurrence. On this study, we'll focus on predicting the time and magnitude.

Chapter 3

Related Work

In 2011, Moustra et al. [2] predicted using MLP and only magnitude time series data if magnitude M of the following day is ≥ 5.2.

Two years later, Martínez-Álvarez et al. [3] predicted the maximum magnitude of earthquake of an observed region in the next five days using 7 features derived from Gutenberg-Richer's law. with a decent accuracy but not enough high to avoid false alarms.

Asencio-Cortés et al., in 2017, derived more features from GR Law and fed an MLP model similar to Martinez-Alvarez et al's, slighlty improving their accuracy but evaluating on a different dataset. [4]

Again in 2017, Asim et al. [5] aimed to predict earthquakes of magnitude 5.5 and above in Hindukush on a monthly basis using machine learning approaches in combination with eight seismicity indicators. Four machine learning techniques were used, including pattern recognition neural network, recurrent neural network, random forest and linear programming boost ensemble classifier. The best results were achieved using the LPBoost ensemble classifier, which achieved the highest accuracy and sensitivity of 65% and 91%, respectively.

In a 2018 Nature article, DeVries et al. [6] used deep learning to predict aftershock locations. Training a neural network on over 131,000 earthquake pairs, they outperformed traditional methods in predicting over 30,000 aftershock locations in a separate test dataset. This system also provided insight into earthquake triggering mechanisms.

In 2019, González et al. [7] published a paper on earthquake magnitude prediction using LSTM recurrent neural networks. The study analyzed the Italian seismic catalog of earthquakes from 1995 to 2018 and found that their model could predict the maximum magnitude in the next hour with minimum error. However, predictions for longer time periods were less accurate.

A meta-analysis of ANN-based earthquake prediction literature from 1994-2019 was conducted by researchers from ETHZ in collaboration with the Swiss Seismological Service in Zurich, Mignan et al. in 2020 [8]. They found two trends: growing interest and increasingly complex models. Despite promising results, simpler models offered similar or better predictions. Given the structured data in earthquake catalogs and few considered features, simpler machine learning models are currently preferred. These models follow classical statistical seismology's empirical laws, which have limited large earthquake prediction abilities according to the authors.

In 2021, Banna et al. [9] published a paper on earthquake prediction using an Attention-Based LSTM model. They used earthquake data from 1950 to 2019 and computed 8 seismic indicators based on GR Law calculated on a monthly basis considering the previous 50 events. Their model, which predicted the occurrence of an earthquake of magnitude higher than 4.7 prior to each month achieved a 74.67% accuracy rate.

Berhich et al., in 2023, used an attention-based LSTM to obtain results for predicting magnitude, time and location of the next earthquake event $t+1$ using only the raw features, i.e. magnitude, time, longitude, latitutde and depth with good results. [1]

You can find a direct comparison on Table 3.1.

Table 3.1: Performance of Different Models in the literature for Predicting Earthquakes

Reference	**Model**	**Dataset Used & Features**	**Performance**	**Prediction Target**
Moustra et al. (2011) [2]	MLP	Time Series of Earthquake Magnitude Data of Greece	52.81% accuracy	If magnitude M of the following day is ≥ 5.2
Martínez-Álvarez et al. (2013) [3]	MLP	Seismic data from four seismic regions of Chile. 7 features derived from Gutenberg-Richer's law.	67.15% accuracy	If maximum magnitude M in the next five day is greater than the mean magnitude in the training data +0.6 times the standard deviation.
Asencio-Cortés et al. (2017) [4]	MLP	Time Series of Earthquake Events (time, magnitude, depth, coordinates) of Japan. Numerous features derived from GR Law	72% accuracy	Magnitude $M \geq 5.0$ in the observed region, in the next seven days.

Continued on next page

Table 3.1 Continued from previous page

Reference	Model	Dataset Used & Features	Performance	Prediction Target
Asim et al. (2017) [5]	LPBoost ensemble classifier	Hindukush region with 441 data vectors corresponding to each month from 1977 to 2013. 8 features from GR law and other seismicity parameters.	Accuracy of 65% and sensitivity of 91%	Earthquakes of magnitude $M \geq 5.5$ in Hindukush on a monthly basis
DeVries et al. (2018) [6]	Deep Neural Network with binary classification (sigmoid activation function)	More than 131,000 mainshock–aftershock pairs. Features: Magnitudes of the six independent components of the co-seismically generated static elastic stress-change tensor calculated at the centroid of a grid cell and their negative values	Area under curve of 0.849	Location of aftershocks following large earthquakes. Each 5 km × 5 km × 5 km grid cell in the volume around each mainshock is classified as either 'containing aftershocks' or 'not containing aftershocks'.
Gonzalez et al. (2019) [7]	LSTM	Italian seismic catalog of earthquakes with magnitude equal to or larger than 1.5 from 1995 to 2018. Features: latitude, longitude, and depth of the hypocenter, time of occurrence, and magnitude treated as a function that represents a set of samples over time	MSE of 0.003	Maximum earthquake magnitude within next hour
Banna et al. (2023) [9]	Attention-based Bi-LSTM	Earthquake catalog from Bangladesh meteorological department and USGS from 1950 to 2019 without foreschocks and afterschocks. 8 seismic indicators calculated on a monthly basis considering the previous 50 events in the calculation based on GR law	Accuracy of 74.67% for earthquake occurrence prediction	Earthquake occurrence of magnitude $M \geq 4.7$ and location prediction prior to each month

Continued on next page

Table 3.1 Continued from previous page

Reference	Model	Dataset Used & Features	Performance	Prediction Target
Berhich et al. (2023) [1]	Attention-based LSTM	Japan earthquake dataset of magnitude $M \geq 5.0$ from 1900 to October 2021. Features : earthquake magnitude, depth, location, and time	MSE of 0.0103	Time, location, magnitude of next earthquake

Chapter 4

Datasets

4.0.1 Data Collection

The data used in this project was collected from the United States Geological Survey (USGS), which is one of the world's most comprehensive and updated resources for earthquake data. The USGS maintains a catalog that documents all earthquake events of magnitude $M \geq 2.5$ globally, making it an invaluable asset for seismologists, geoscientists, emergency responders, and researchers interested in the study of seismic activities.

For this study, we focused on the data starting from 1920, as it was from this year that the earthquake data was collected consistently in the regions of interest.

Despite the absence of an official API provided by USGS, their portal allows users to query and download up to 20,000 earthquake events at once. To handle this, a Python script has been written to send requests to the USGS data source, retrieve the data, and append it into a single csv file. This process was carried out iteratively until all data for the specified time range was collected.

Using this method, we successfully collected and merged all earthquake events available on USGS database from around the world from 1920 till 2023, maintained in chronological order. The final dataset amounts to approximately 165 MB and is available on the following link :

https://media.githubusercontent.com/media/fmutlu/EarthquakePred/main/earthquake_data_from_1920.csv

Each row is composed of following key entries:

- `'latitude'`
- `'longitude'`

- `'depth'` : The depth at which the earthquake event occurred. This is crucial as it can be related to the severity of the event.
- `'mag'` : The magnitude of the earthquake event. Only earthquakes of magnitude $M \geq 2.5$ are reported.

Other columns of less importance are:

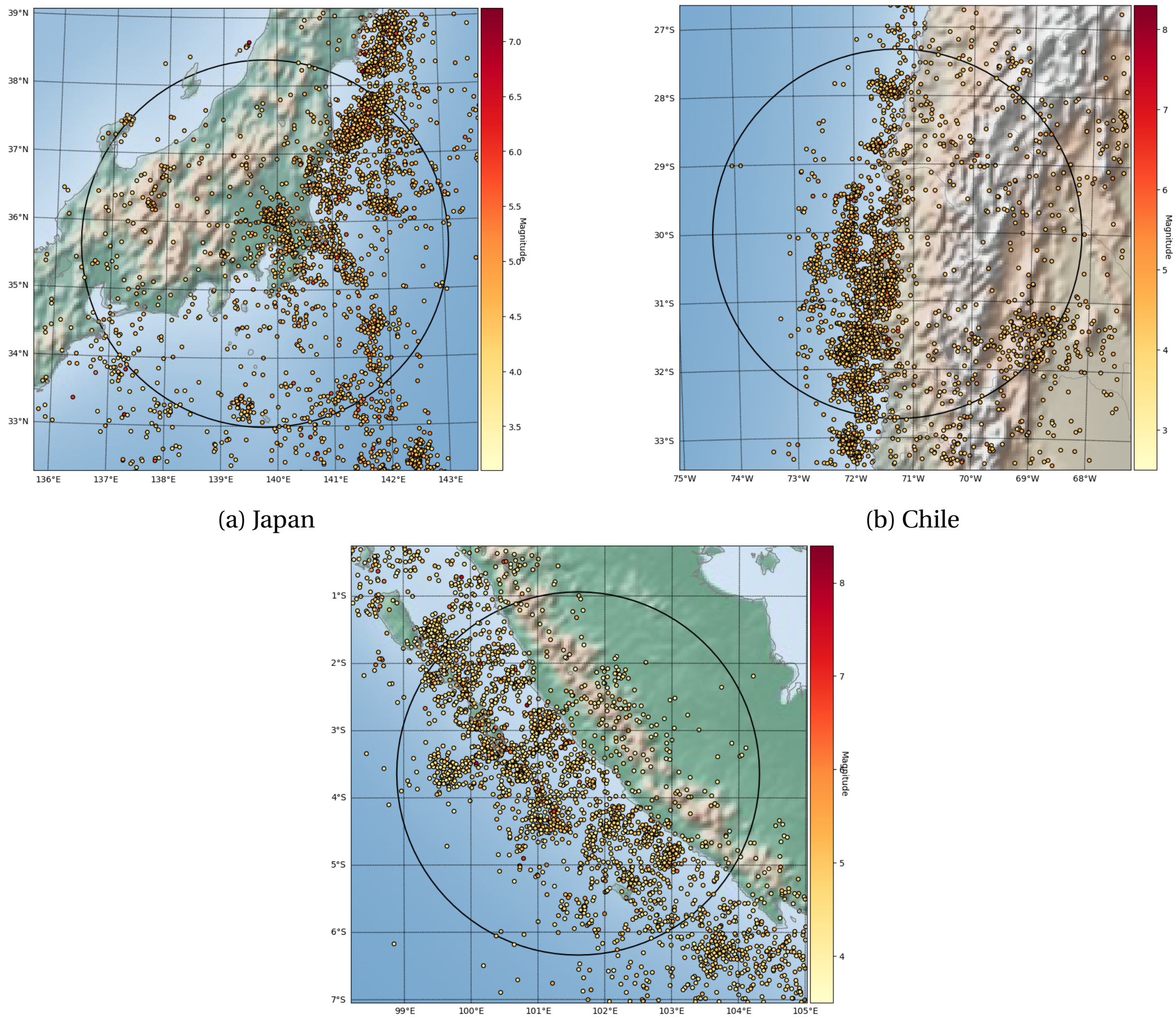
- `'depthError'` : The estimated error of the depth.
- `'magError'` : The estimated error of the magnitude.

These error estimates, while they may be valuable in other analyses or approaches, are secondary to our main objective which is to learn seismic activity patterns. As a consequence, they won't be used in the context of the project.

4.0.2 Regions

Given the comprehensive nature of our dataset which encompasses global seismic activity, we need to focus on specific regions, ideally seismic ones. By doing this, we can gather more relevant and accurate information for our research purposes. Moreover, these seismic regions have been extensively researched in the literature and it provides an opportunity for comparative analysis. One more reason is that some seismic parameters like b-value are highly dependent on the region.

To facilitate the region selection process, we employed a technique described in Asencio et al. (2017) [4]. It consists in selecting a central point defined by latitude and longitude coordinates and setting a radius, measured in kilometers, around this point. This way, we can extract all events for which the coordinates fall within the defined circular boundary. The radius value was set to be large enough to encapsulate the seismic cluster. We can get a detailed overview on the selected regions and the distribution of earthquake events on them on Figure 4.1.



(a) Japan

(b) Chile

(c) Indonesia

Figure 4.1: Extracted regions. Each point represents an individual earthquake event. The dark circle indicates the boundaries of the selected region.

Depending on time periods, the data could be very sparse or very dense. Moreover, the range of magnitudes represented could suddenly change over time.

For reasons that will be explained later, we extracted 3 regions' data on very specific time ranges and made 2 versions for each.

In the first version, we extracted all the data that we have from 1st January of 1973 to 31st December of respectively 2010, 2012 and 2023 and we kept only events of magnitude $M \geq 5.0$. This way of filtering of data, by keeping only high magnitudes that we're interested in, has already been used in Berhich et al. [1].

In the second version, we kept all the magnitudes but we select a shorter time range, from the 1st January of respectively 2012, 2017 and 2011 to 27-02-2023.

We give more details in section 5.1 about why we selected these specific time ranges.

Following table 4.1 summarizes the extracted datasets.

Table 4.1: Datasets Used For Our Models. Datasets of version has a restricted magnitude range, containing only events of magnitude $M \geq 5.0$.

Region	Latitude	Longitude	Radius (km)	Version	Mag Range	# Entries	Time Period
Japan	35.683	139.774	300	1	5.0 – 7.9	802	1973-2010
				2	3.1 – 7.9	2,961	2012-2023
Chile	-30.022	-71.277	300	1	5.0 – 8.3	467	1973-2012
				2	2.8 – 8.3	1,097	2017-2023
Indonesia	-3.641	101.601	300	1	5.0 – 8.4	975	1973-2023
				2	4.0 – 7.8	1,115	2011-2023

Chapter 5

Method

5.1 Preprocessing

The data is already cleaned in a standard format by USGS. For some regions, there are inconsistencies like sudden changes in distribution or range of magnitudes.

On the magnitude plot for South Italy on figure 5.1,we can observe that the data for the time period from 2009 becomes scarce in comparison with past values. Moreover, from 2002 to 2009 there is a sudden increase in the number of earthquake events with magnitudes $2.5 \geq M \geq 3.5$ This can be attributed to a shift in seismic activity; however, the primary reason is likely to be a change in the way of how data has been collected by USGS for this specific time period.

We can use data augmentation to handle this kind of data that becomes sparse over time. However, this wouldn't preserve data integrity and cause a lack of diversity that wouldn't truly capture the seismic patterns of the region. For this reason, we only selected seismic regions for which we have consistent and complete data. Moreover, the time periods have been selected such that the cumulative count plots form almost a straight line. Figure 5.2 shows the difference between a suitable dataset to train models — Japan — and a non-suitable one.

5.1.1 Normalization

Except for time-related features used in LSTM models, all features have been normalized using StandardScaler from sklearn library, once the features have been computed. The normalisation is done at the feature level, except for raw features (magnitude, depth) for which it's done at the dataset level.

To encode time differences in hours between two consecutive events, a MinMaxScaler has been

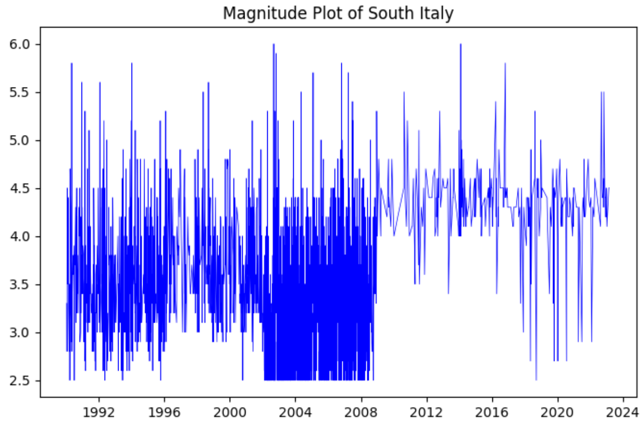
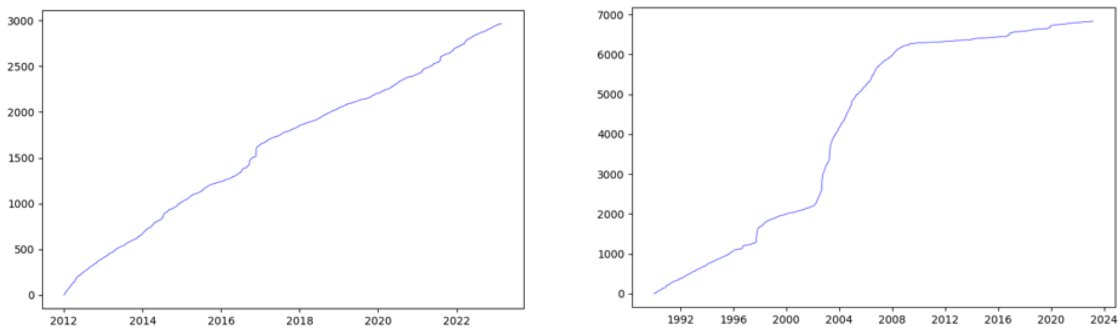


Figure 5.1: Example of temporal sparsity in the dataset



(a) Japan

(b) Italy

Figure 5.2: Cumulative Count Plots for Japan_1 and Italy datasets

used.

5.2 Feature extraction

On our baseline MLP model, features are dependent on b-value from the Gutenberg–Richter law :

$$\log_{10} N(M) = a - bM \tag{5.1}$$

Based on the Gutenber-Richter law, the key features are calculated using the last 50 recorded earthquakes as follows:

$$b_i = \frac{\log(e)}{(1/50) \cdot \sum_{j=0}^{49} (M_{i-j} - 3)} \tag{5.2}$$

$$\Delta b_1^i = b_i - b_{i-4} \equiv x_1^i \tag{5.3}$$

$$\Delta b_2^i = b_{i-4} - b_{i-8} \equiv x_2^i \tag{5.4}$$

$$\Delta b_3^i = b_{i-8} - b_{i-12} \equiv x_3^i \tag{5.5}$$

$$\Delta b_4^i = b_{i-12} - b_{i-16} \equiv x_4^i \tag{5.6}$$

$$\Delta b_5^i = b_{i-16} - b_{i-20} \equiv x_5^i \tag{5.7}$$

The sixth input variable, x_6^i, is the maximum magnitude Ms from the quakes recorded during the last week in the area analyzed. The use of this information as input is to indirectly provide the ANN with the required information to model Omori/Utsu and Bath's laws on aftershocks.

$$x_6^i = \max\{Ms\}, \quad \text{when } t \in [-7, 0) \tag{5.8}$$

where the time t is measured in days.

5.3 Models Architecture

Model	Layers (units)	Activation function	Epochs	Optimizer	Learning Rate	Batch size	Loss function
MLP Predict if next big	Dense(30), Dropout(0.6), Dense(20), Dense(2)	ReLU	2500	Adagrad	0.001	32	Binary Cross-Entropy
LSTM Predict Time Until Next	LSTM (50), Dense(1)	Linear	250	Adam	0.001	32	Mean squared error
LSTM Predict If Next Big	LSTM (50), Dense(1)	Linear	250	Adam	0.001	32	Binary Cross-Entropy

Models' architectures are described on table 5.3.

"MLP Predict if next big" model uses Softmax function on the last layer to make the classifications.

It should be noted that for both LSTM models, early stopping has been applied. For "LSTM Predict Time Until Next" model trained on Japan, Chile, Indonesia datasets, the model stopped respectively after 144, 250, and 94 epochs. Similarly, for "LSTM Predict If Next Big", the model stopped respectively after 43, 91 and 23 epochs.

5.4 Objectives, Datasets & Features Used and Performance Assessment

Table 5.2: Model features, labels and scores used to assess the performance

Type	Model	Features	Predicted Label	Score Used
Regression	LSTM Predict Time Until Next	Time elapsed from last event, Magnitude, Longitude, Latitude, Depth	Time until next earthquake of magnitude $M \geq 5.0$	MSE, MAE, R^2
Classification	MLP Predict Big In Window	7 features from Asencio et al. [4]	If the next earthquake within a time window of 7 days is of magnitude $M \geq \text{mean}(M) + 0.6 \cdot \text{std}(M)$	Accuracy, AUC, F1, Precision, Recall
	LSTM Predict If Next Big	Time elapsed from last event, Magnitude, Longitude, Latitude, Depth	If the next earthquake is of magnitude $M \geq 5.0$	Accuracy, AUC, F1, Precision, Recall

On Table 5.2, you'll find the features used, the predicted labels and the type of scores used to assess the models' performance. "LSTM Predict Time Until Next" is the only model that has used Version 1 of datasets containing only magnitudes $M \geq 5.0$. The two other models has used the Version 2.

For LSTM models, sequence size that was performing best, found empirically, was 10. This means that each input vector processed by the model is composed of 10 consecutive data points from our dataset.

One main issue with the "MLP Predict Big In Window" models' label is that it can only predict if an earthquake will happen in the following 7 days. If an earthquake happens later, it is completely missed by this approach.

5.5 Training

5.5.1 Dataset Split

For all models, the dataset has been split in 70% training, 10% validation, and 20% test sets using fixed partitioning. Figure 5.3 illustrates this split.

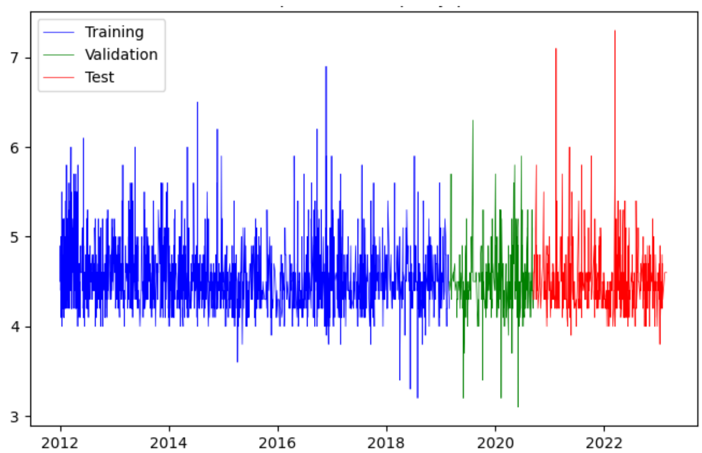


Figure 5.3: Example of a dataset split (Japan_2)

5.5.2 Hyperparameters

Hyperparameters have been choosen using K-Fold Cross Validation with $K = 5$ without shuffling based on the Japan dataset.

A different set of hyperparameters have been chosen for each model using this method.

The other parameters such as the number of layers, the number of neurons by layer or the optimizer have been chosen empirically.

Chapter 6

Results

Table 6.1: Regression models' performance

Model	Region	Train			Test		
		MSE	MAE	R^2	MSE	MAE	R^2
LSTM Predict Time Until Next	Japan	0.0482	0.1866	0.0691	0.0640	0.1851	0.1657
	Chile	0.0343	0.1134	0.1012	0.0232	0.1111	0.0482
	Indonesia	0.0895	0.2136	0.15031	0.0899	0.2136	0.0680

Table 6.1 shows the results for the regression model. Early stopping has been used for the training of each of the 3 regions. While MSE and MAE scores are close between train and test sets' , the R^2 on test set shows that the model doesn't generalize well on unseen data except for the Japan dataset.

Table 6.2: Classification models' performance

Model	Region	Test				
		Acc	F1	ROC	Precision	Recall
MLP Predict Big In Window	Japan	0.516	0.668	0.587	0.514	0.951
	Chile	0.562	0.275	0.537	0.271	0.280
	Indonesia	0.615	0.392	0.569	0.396	0.388
LSTM Predict If Next Big	Japan	0.804	0.891	0.691	0.804	1.0
	Chile	0.812	0.890	0.691	0.786	0.98
	Indonesia	0.793	0.885	0.677	0.793	1.0

Table 6.2 shows the results for the classification models. Accuracy, F1, ROC_AUC, Precision and Recall scores are given.

We can notice that LSTM performs well on predicting if next earthquake will be big, i.e. of magnitude $M \geq 5.0$ but not very well on predicting *when* it will arrive. This is the case for all 3 regions. This could be explained by the fact that a restricted dataset has been used for this model, on a longer time range however.

Respective to features, results show that those derived from GR and Bath laws have limited power in comparison with raw features used along with advanced models such as LSTM.

6.1 Metrics

We propose a simple evaluation metric for predicting the next significant earthquake that accounts for the average change over time, which we refer to as *average_delta*. This metric offers a baseline that the prediction model should aim to outperform.

Given a series of timestamps $\boldsymbol{y}_{\text{actual}} = [y_1, y_2, \ldots, y_N]$ of N points, where each $y_i \in [-1, 1]$ represents a timestamp (scaled using MinMaxScaler from the exact date times), we first compute the average delta Δ as follows:

$$\Delta = \frac{1}{N-1} \sum_{i=1}^{N-1} (y_{i+1} - y_i) \tag{6.1}$$

Next, we compute two Mean Squared Error (MSE) measures, one for the model's predictions $\boldsymbol{y}_{\text{pred}} = [y'_1, y'_2, \ldots, y'_N]$ and one for the actual values, calculated as follows:

$$\text{MSE}_{\text{pred}} = \frac{1}{N} \sum_{i=2}^{N} ((y'_i - y_i)^2) \tag{6.2}$$

$$\text{MSE}_{\text{actual}} = \frac{1}{N-1} \sum_{i=1}^{N-1} ((y_i + \Delta - y_{i+1})^2) \tag{6.3}$$

The MSE for the model's predictions should ideally be lower than the MSE for the actual values :

$$\text{MSE}_{\text{pred}} < \text{MSE}_{\text{actual}} \tag{6.4}$$

This indicates that the model's predictions are better than simply predicting the next timestamp based on the average change in the timestamp series. Conversely, if $\text{MSE}_{\text{pred}} > \text{MSE}_{\text{actual}}$, this suggests that the predictions are worse than the average.

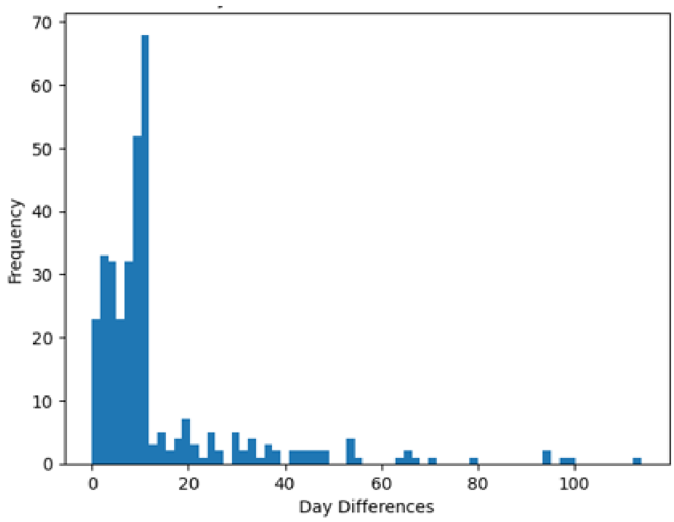


Figure 6.1: Absolute Difference in Days Between Predicted and Actual Dates of Earthquakes

The relation 6.4 was satisfied by the model "LSTM Predict Time Until Next" but only with some regions and specific time ranges. On figure 6.1, we can observe the absolute difference in days between predicted and actual earthquake dates on the Japan_1 dataset, with an actual average difference of 10.2 days.

Chapter 7

Conclusion

The application of deep learning models for earthquake prediction, as explored in this project, has revealed intriguing potential but also significant limitations. The predictive power of features derived from GR Law and Bath Law was limited, echoing the findings of Mignan et al. [8] LSTM model has shown important predictive power in classifying if the next earthquake was going to be big, demonstrating the potential use of such models in this task. However, the LSTM model was less successful in predicting the temporal occurrence of these events.

Given these findings, it is clear that the development of successful predictive in the future will need a more comprehensive approach. Future models should aim to integrate a larger set of inputs. This could be satellite images, fault measurements, and other geophysical data not integrated in current models.

Additionally, a more global model, taking into account seismic activity across various regions, may offer improved prediction accuracy. In this direction, an avenue of exploration that could be promising is the application of graph neural networks.

To finalize, while deep learning offers significant potential in improving our ability to predict earthquakes, this project has also emphasized the complex nature inherent to seismic activities.

Bibliography

[1] Berhich et al. “Soil Dynamics and Earthquake Engineering”. In: *Soil Dynamics and Earthquake Engineering* 165 (2023), p. 107663. DOI: 10.1016/j.soildyn.2022.107663. URL: https://doi.org/10.1016/j.soildyn.2022.107663.

[2] M. Moustra et al. “Expert Systems with Applications”. In: *Expert Systems with Applications* 38.12 (2011), pp. 15032–15039. DOI: 10.1016/j.eswa.2011.05.043. URL: https://doi.org/10.1016/j.eswa.2011.05.043.

[3] F. Martínez-Álvarez et al. “Applied Soft Computing”. In: *Applied Soft Computing* 13 (2013), pp. 1314–1328. DOI: 10.1016/j.asoc.2012.10.014. URL: https://doi.org/10.1016/j.asoc.2012.10.014.

[4] G. Asencio-Cortés et al. “Neural Computing and Applications”. In: *Neural Computing and Applications* 28.5 (2017), pp. 1043–1055. DOI: 10.1007/s00521-015-2121-7. URL: https://doi.org/10.1007/s00521-015-2121-7.

[5] K. M. Asim and Martínez-Álvarez et al. “Earthquake magnitude prediction in Hindukush region using machine learning techniques”. In: *Natural Hazards* 85.1 (2017), pp. 471–486. DOI: 10.1007/s11069-016-2579-3. URL: https://doi.org/10.1007/s11069-016-2579-3.

[6] DeVries et al. “Deep learning of aftershock patterns following large earthquakes”. In: *Nature* 560.7720 (2018), pp. 632–634. DOI: 10.1038/s41586-018-0438-y. URL: https://doi.org/10.1038/s41586-018-0438-y.

[7] González et al. “Earthquake Magnitude Prediction Using Recurrent Neural Networks”. In: *Proceedings* 24.1 (2019), p. 22. DOI: 10.3390/IECG2019-06213. URL: https://www.mdpi.com/2504-3900/24/1/22.

[8] Arnaud Mignan and Marco Broccardo. “Neural Network Applications in Earthquake Prediction (1994–2019): Meta-Analytic and Statistical Insights on Their Limitations”. In: *Seismological Research Letters* 91.4 (2020), pp. 2330–2342. DOI: 10.1785/0220200021. URL: https://doi.org/10.1785/0220200021.

[9] Banna et al. “Attention-Based Bi-Directional Long-Short Term Memory Network for Earthquake Prediction”. In: *IEEE Access* 9 (2021), pp. 56589–56603. DOI: 10.1109/ACCESS.2021.3071400. URL: https://doi.org/10.1109/ACCESS.2021.3071400.