ISSN 1870-4069

Earthquakes Insights and Predictions in Mexico Using Machine Learning

Adolfo Bustillos, Carmen Santiago, Gustavo Rubín

Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación,
Mexico

juadolfob@gmail.com, {marycarmen.santiago,
gustavo.rubin}@correo.buap.mx

Abstract. Earthquake prediction is a challenging area of research and modest efforts have been made using machine learning for this application. In this work, seismic characteristics are calculated using seismological concepts, such as the Gutenberg-Richter law, rate of seismic change, anticipation frequency, release of seismic energy, and total recurrence time. With this, an Artificial Neural Network (ANN) classification model was built using sliding windows (as opposed to fixated sets of windows seen in another studies) to calculate its instances that were then resampled as a measure to counter the unbalanced data for the purpose of obtaining earthquake predictions; the model predicts more seismic events above a threshold without significantly sacrificing the metrics.

Keywords: Earthquake prediction, machine learning, neural network.

1 Introduction

Like many natural disasters, the earthquake causes many damages, economic and human losses and injuries [1]. That is why one of the most ambitious objectives of seismology is the prediction of earthquakes in the short term. In the mid-1970s, seismologists were confident that short-term earthquake prediction would be achieved in a short period of time. This confidence arose in part as a result of the first successful prediction of a major earthquake, the 1975 M7.4 Haicheng earthquake in China. Due to this prediction, an alert was issued within the 24-hour period before the main shock, probably preventing a greater number of casualties than the 1,328 deaths that actually occurred from this event. However, the impossibility of predicting another devastating earthquake 18 months later The 1976 Tangshan earthquake M7.8 was a major setback for the earthquake prediction effort. The victims of this earthquake number in the hundreds of thousands [2]. Lomnitz provides a summary of these events, as well as other successes and failures in earthquake prediction [3].

Earthquake prediction is not a task expected to be solved soon, but efforts of improving the methodology such as the ones in this study and those in the related literature make the task closer to reality. In this study a new approach to calculate seismic indicators from a catalog is presented, with sliding windows (this approach is based on the fixed window approach that other studies use) hoping this will improve sensitivity and the quantity of instances, and it in addition gives solutions to the problem of unbalanced data. The always improving of seismic networks has permitted a huge amount of data, and thus has renewed hope that earthquake forecasting could be a feasible task, if combined with the improving of forecasting methods.

In Section 2 Related Literature and the seismic precursors using in them will be discussed, then in Section 3 the process of preparing the catalog and insights about mexican región will be described, in Section 4 there will be further details on seismic indicators and how they were computed in this research, at Section 5 multiple models of prediction, its performance and description of them will be presented, and at the end in the Conclusion Section the results will be discussed.

2 Related Literature

In the literature, there are different research methodologies, indicators and seismic precursors that have been used together with Machine Learning (ML) techniques for seismic analysis. These methodologies assume that there are variations in seismic precursors during pre-earthquake, and that there are patterns in these variations.

In 2007 and 2009, two important papers that complement each other were published by A. Panakkat and H. Adeli, first in [20] they formulated the prediction problem as a classification task, this attempted to predict the magnitude of the largest seismic event in a time window and a predefined region in the next month. They proposed eight of the so-called seismicity indicators: mathematically calculated characteristics, which can be used to assess the seismic potential of a region [16]. And later in [27] the same authors proposed the architecture of a probabilistic neural network (PNN) as a solution for the same problem that was formulated in [16] where they used the same set of seismicity indicators as input data for the network training. The model was tested with data from the Southern California seismic zone and yielded good prediction accuracies for events of magnitude 4.5 to 6.0. However, PNN did not work satisfactorily for earthquakes of magnitudes greater than 6.0.

In 2013 the article [21] was published, it used an ANN model to predict the magnitude of the earthquake in an interval or limited by a threshold during the next five days. The application of statistical tests and experiments showed the higher success rate of this method than other machine learning classifiers at the time [18]. The system is designed to provide two types of predictions: a) the probability of an earthquake occurring greater than a threshold magnitude in five days and b) the probability of a seismic event occurring within a predefined magnitude range [25].

In this study, new seismic parameters were defined based on Bath's law and Omori-Utsu's law, which describe the relationships between the main shock and aftershocks, according to their magnitude and frequency of occurrence, respectively.

In 2018, adding to the earthquake prediction problem, the earthquake prediction system (EPS) called EP-GP Boost was described in article [14]. This system is a classifier based on a combination of genetic programming (GP) and a boost algorithm called AdaBoost. A total of 50 features were calculated. The study of the applicability

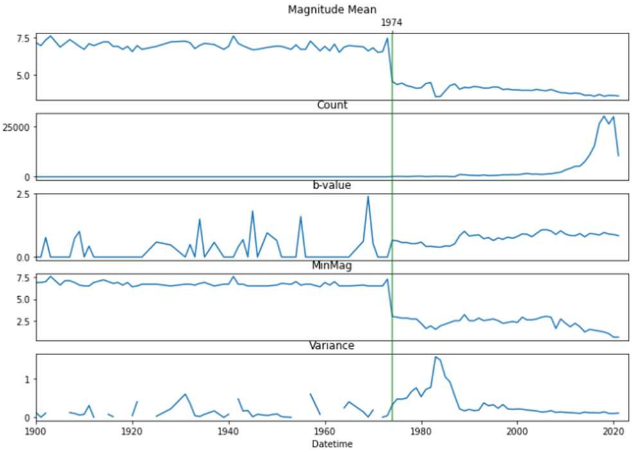

Fig. 1. Catalog perspective in Mexico, before 1974 the catalog is incomplete for magnitudes m^* > = 3.9, the gaps in variance, zeros and outliers in "b-value" are caused by lack of data. The variance graph is according to the magnitudes of the year, and the b-value graph describes the frequency of the earthquake size distribution based on the Gutenberg-Richter law.

of EP-GP Boost was carried out using data from previously used seismic zones, namely Chile, Hindukush and southern California. Experiments have shown outstanding performance in all three observed regions, both in terms of a low false alarm rate.

The best results were obtained for the Southern California region due to the completeness and quality of the corresponding earthquake catalog. However, the results for all regions show an improvement compared to previous studies [16, 18, 23].

Later the same year, [16] was written by the same authors as [14]. In this article, Asim et al. also used the approach for the use of seismicity indicators proposed in [14]. This time, 60 seismic parameters were calculated using various seismology concepts. As in their previous research, the authors intend to predict earthquakes of magnitude equal to or greater than 5.0 over the next 15 days. The system is a combination of different machine learning algorithms, and in each step, an algorithm uses the knowledge gained through learning from a previous one. They made use of mRMR, SVR, HNN and an EPSO for the same regions as [14]; it gave better results in the metrics.

In the region of Mexico there are project proposals of earthquake prediction using ML as [26] where random forest and deep learning models were proposed; also other damage prediction and modeling studies.

3 Earthquake Catalog Preparation

This data used in this study is based on earthquake catalogs. The catalogs were downloaded from the SSN (Servicio Nacional Sismológico) using the period from January 1900 to May 2021, the catalog contains 204806 seismic events.; They are available publicly, after discarding irrelevant columns and eliminating events with incomplete information as part of the cleaning and preprocessing we were left with

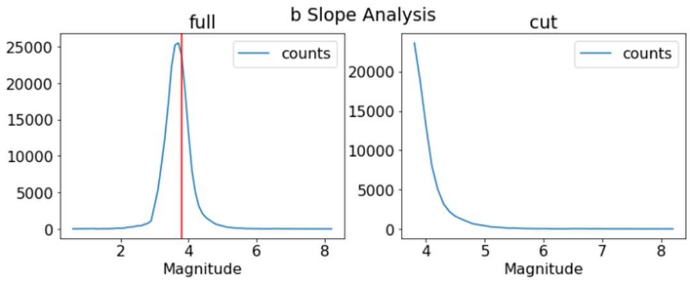

Fig. 2. Gutenberg-Richter Law applied in Mexico, from 1974 and for m^* > = 3.9 with a = 8.326 and b = 0.998, the y-axis contains the number of events, and the x-axis the magnitude of th events.

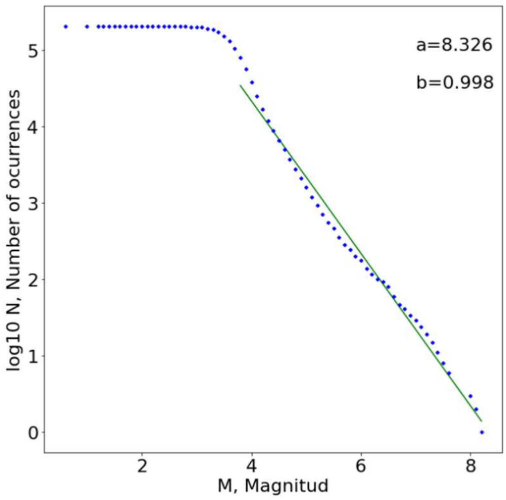

Fig. 3. Gutenberg-Richter Law applied in Mexico, starting in 1974 for m^*> = 3.9 with a = 8.326 and b = 0.998, it is calculated using the method of "linear least square regression" (lsq).

204806 seismic events and with their Date, Magnitude, Latitude and Longitude features.

3.1 Mexico Catalog Overview

The magnitude of completeness is defined as the lowest magnitude in which all earthquakes are successfully detected within a region and time period [5, 6], it is essential to analyze the seismicity of a region [6]. The magnitude of completeness varies in time and space and depends on many factors that affect the detection capacity of a seismological network, such as: the density and distribution of seismic stations, the type of instrumentation used, the efficiency of informing data from stations to the processing center, earthquake detection practices and procedures, among others [6, 7, 8, 9].

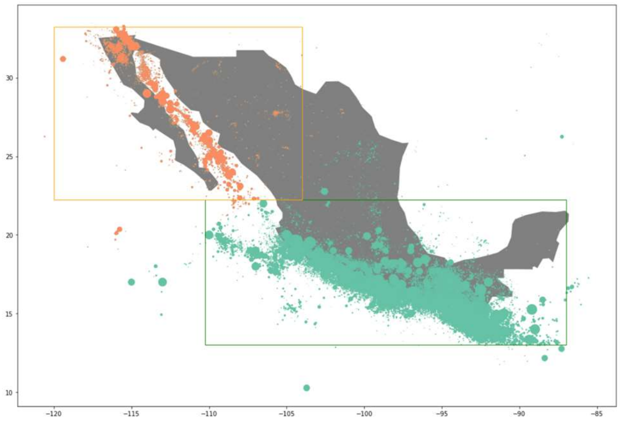

Fig. 4. Catalog division by KNN in function to its geometry, and selection of regions.

To satisfy the completeness of a region, the geometric extension is delimited, a temporal threshold and a cut-off magnitude is defined, the events that occurred below the thresholds and outside the boundaries of the region will not be used for the analysis, if the the catalog contains all the seismic events that occurred in a region, of magnitudes greater than or equal to as of a certain date, then the catalog is considered to be complete.

The seismic catalog of Mexico is probably complete for m^*> = 7.7 as of 1846, m^*> = 7.0 as of 1856 and m^*> = 4.3 as of 1969 [10]. But the quality of the catalog increases dramatically from 1974, as shown in Fig. 1 and completeness is satisfied with m^*> = 4.1, as shown in Fig. 2. This is the minimum date we will use to analyze the Mexico seismic catalog. The Gutenberg-Richter law describes the frequency of the earthquake size distribution [11], the law is described with the following expression:

$$log_{10}N = a - bM \tag{1}$$

where a represents the total seismicity rate of the region, b is the relative distribution of the size of earthquakes and N is the number of earthquakes [12]. High b values mean a predominance of small earthquakes; conversely, a low b-value means that large earthquakes dominate smaller earthquakes.

Variations in the b-values both spatially and temporally are generally considered as clues for the precursors of large earthquakes [13]. To determine the cut-off magnitude in this study, “b-slope analysis” was used, the number of events for each magnitude are counted and the point where the curve deviates from the exponential behavior is selected, this is called the cut-off magnitude m^* and only events greater than or equal to this magnitude are considered for the analysis. When applying this analysis in the Mexican catalog for events from 1974 onwards, the resulting cut-off magnitude is m^* > = 3.9.

When applying the Gutenberg-Richter law in the Mexican catalog for m^* >= 3.9 from 1974, a value for a = 8.326 and b = 0.998 resulted, it is visualized in Fig. 3.

3.2 Geometric Extension of Catalog

The geometric extension of the catalog is from latitude 33.5 to 10.3 and longitude -120.5 to -85.5, that is, it only contains events within these limits, it covers the entire Mexican territory and the events are shown in Fig. 4, which is too extensive compared to other regions of analysis in related literature such as [14, 15, 16].

To select a smaller region, the catalog was separated into two parts using the KNN algorithm according to its geometry. Fig. 4, this measure also increases the correlation of the data; After a longitude and latitude range was arbitrarily selected to formalize the space of the analysis region, this facilitates replications of the analysis and comparison with other analyzes.

The resulting division between the northwest region and the southern region is consistent with the boundaries of the interacting tectonic plates in Mexico. The northwest region is represented in Fig. 4 in the upper part. covers the North Pacific, Gulf of California and Northwest Mexico, this encompasses longitudes from 120 to -104 and latitudes from 33.25 to 22.25, it contains 12,184 events where a = 7.313 and b = 0.988, and the adjacent plates are the North American Plate and the Pacific Plate. The southern region is represented in Fig. 4 in the lower part, it encompasses the tropical south Pacific and the west, center and southeast of Mexico, this encompasses longitudes from -110.25 to -87.5 and latitudes from 22.25 to 13, it contains 192622 events where a = 8.221 and b=0.987 and the adjoining plates are the North American Plate, Rivera Plate, Cocos Plate and Caribbean Plate.

3.3 Temporal Extension of the Catalog

The model is sensitive to the frequency with which events are reported, so it is necessary that the frequency of each year is similar. The largest number of reported events are in the most recent years (its general catalog quality also increases), and they become regular frequencies after 2010 for the northwest region and after 2017 for the southern regions as seen in Fig. 5.

For the Northwest region 2010 was used for the cutoff year, as quality improves for more recent dates, resulting in higher concentration and quality of events, and for the southern region, 2017 was used as the limit year, as of 2010 the frequency of reported events increases but until 2017 the frequency of events stops growing and stabilizes, in addition, the largest number of events with a greater quality is concentrated here.

3.4 Cut-Off Magnitude

The cutoff magnitude corresponds to the minimum magnitude in the catalog above which the earthquake catalog is considered complete, that is, there are no missing seismic events, and it has to be calculated for each region and time range. The value of the magnitude of cut depends on the integrity of the catalog, which itself depends on the instrumentation. Better instrumentation in a region leads to better catalog integrity with a low cutoff magnitude.

To evaluate the cut-off magnitude of this catalog we used an analysis of the Gutenberg-Richter curve [17]. The point where the curve deviates from the exponential

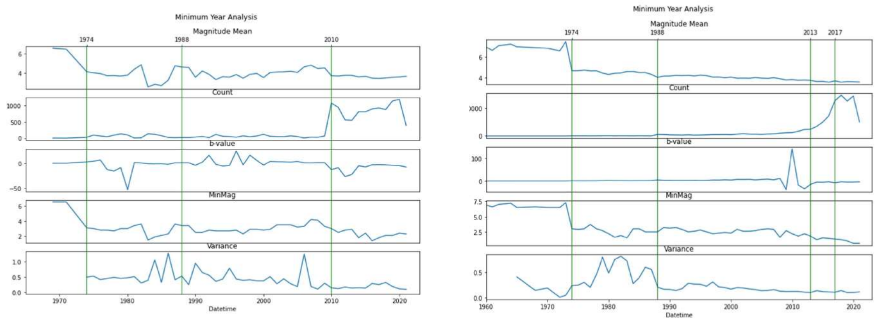

Fig. 5. Minimum year analysis for the northwest region on the left side and for the southern region on the right side.

behavior is selected as the cutoff quantity. All events below the cut-off magnitude are discarded and are not used in the analysis.

The magnitudes of earthquakes and frequencies of occurrences of a region are plotted as shown in figure Fig. 6. In the left sub-figure of each figure the red line marks the selected cut-off magnitude, in both northwest regions and south, with cut-off magnitude, m^*> = 3.9.

The resulting curves after cutting are shown in the right sub-figure of each figure; they follow an exponential behavior, which ensures that each catalog is complete at its respective cutting magnitude.

4 Calculation of Seismic Parameters

Indicators are the most important part of a classification problem [16]. Two types of indicators are calculated, parametric and non-parametric indicators that in several cases have multiple values based on different variations of a parameter, these relationships are called indicators or seismic parameters, the relationships that are used in this model are the geophysical and seismological parameters relevant for the prediction of earthquakes available in contemporary literature, these criteria are taken in order to retain the maximum information available on the internal geological state of the soil.

Time series prediction can be framed as a supervised learning problem. A sliding or moving window is a technique used to frame a set of time series data.

To calculate the metrics for this model, a multivariate of size $n \times 2$ moving window that looks at the date and magnitude of each catalog event slides through the catalog events and pulls subsets of the catalog of size $n \times 2$ at all positions in the window. The resulting $E - n + 1$ subsets will be subsets of size $n \times 2$, where E is the total number of events in the catalog.

The indicators are calculated for each subset generated, this results in $E - n + 1$ subsets of size $1 \times I$ where I is the number of indicators and labels (which are the indicators to be predicted) calculated.

The model will use the indicators to predict the labels, in this study the target parameters are called labels, which are the indicators that are sought to be predicted. To calculate the indicators and labels of new seismic events, it is not necessary to wait for

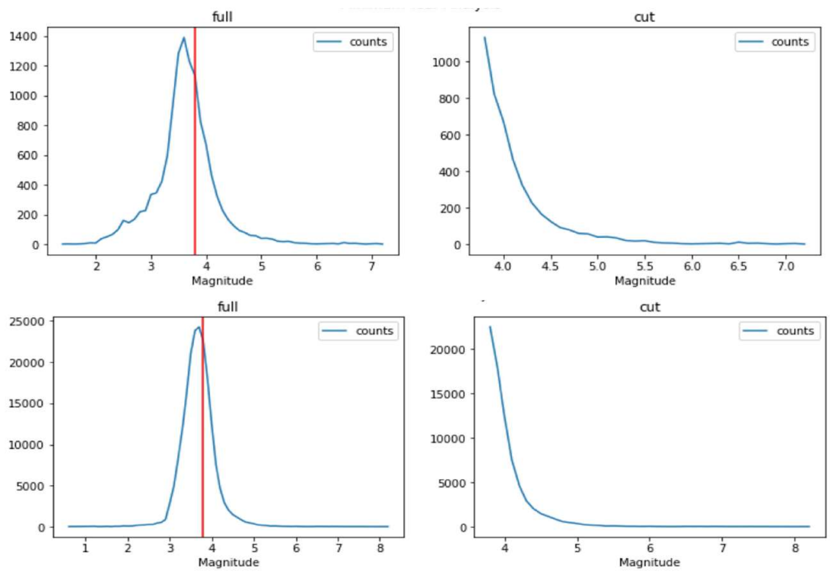

Fig. 6. Analysis of the Gutenberg-Richter curve for the northwestern region on the upper side and for the southern region on the lower side.

other new n events to occur, since the new J events and the latest n-j events can be used. In this study, the earthquake prediction problem is modeled as a binary classification problem, which seeks to predict whether a seismic event greater than or equal to the magnitude m will occur in the next d days. When using classification, the result is one class, among a limited number of classes.

With classes we refer to arbitrary categories according to the type of problem, in this study the classes are class 1 and class 0, if the model predicts that a seismic event greater than m will occur in the next d days the class will be 1, otherwise If the model predicts that a seismic event of magnitude greater than m will not occur in the next d days then the class will be 0. Although this classifier model can be trained to predict different magnitudes, there are models that use “Hybrid Machine Learning” techniques to predict the magnitude [18].

4.1 Parameters and Definitions

There are four parameters that must be determined to calculate the seismic indicators that the model is going to use.

First the parameter n is the window size, it describes the number of window events, the selected value is arbitrary and can be adjusted to obtain better results, normally in the related literature it is between 50 and 100, the size selected in this study is 100; The parameter m is the Magnitude threshold, this is the magnitude that will function as a threshold for classification, and several parameters and labels depend on this variable, the value normally used is between 5 and 6, the magnitude selected in this study is 5.5 (accommodating ourselves to the definition of a magnitude of "great earthquake" of [25]), we can't choose a very high value because the precision decreases in high

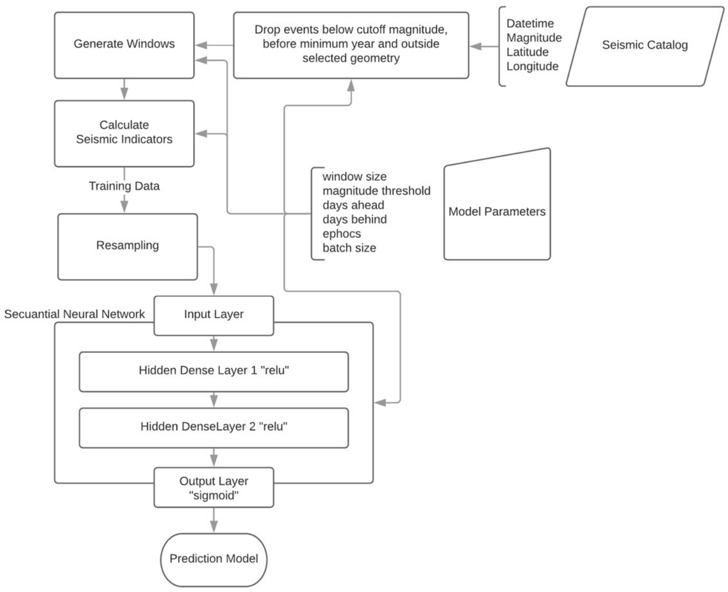

Fig. 7. It shows the structure of the prediction model. There are different variations of the model, where the Resampling stage is not used or is replaced by a Weighting stage, these variations are explained later.

magnitudes [25]; The parameter d_a is days ahead, this value can change depending on how many days ahead you want to predict the next seismic event, the value normally used is between 7 and 15, the value chosen in this study is 14. And the parameter d_b is days behind, the value normally used is 7, it is used to calculate the seismic indicators, the value chosen in this study is 7.

4.2 Indicators

The indicators used here are all those proposed in [20] and [21], plus a selection of the indicators proposed in [14] and [16].

The selected indicators are: date T_1 this is the date of the first seismic event of the window where T_i is the date of the i-th event of the window; Last date T_n last day of the window; Time elapsed T_Θ between the first and the last T_i of a window proposed in [20]; Events Time Difference Mean μ proposed in [20]; Standard deviation of the mean of the difference in elapsed time between events c [20]; Value a and b calculated numerically proposed in [20] using two different methods, one method represents linear least squares regression (lsq), and the other represents Maximum Maximum (mlk) as proposed in [14][16]; Mean magnitude M_{mean} proposed in [20]; Max Magnitude $M_{max\ observed}$ proposed in [20]; Expected Magnitude $M_{max\ expected}$ proposed in [20]; Magnitude Deficit ΔM proposed in [20]; Standard Deviation of the b-value σb proposed in [21]; Mean Square Deviation η proposed in [20]; Square Root Rate of Energy $dE\frac{1}{2}$ proposed in [20], if the release of energy stops, the phenomenon is known

Table 1. Result of the evaluation metrics.

Métricas	Baseline	Weighted	Resampled
TP	2043	2695	3000
FP	455	1617	946
TN	8221	7059	7730
FN	2170	1518	1213
Accuracy	0.79	0.75	0.83
Precision	0.81	0.62	0.76
Recall	0.48	0.63	0.71
F1 score	0.60	0.63	0.73
NPV	0.79	0.82	0.86
R score	0.43	0.45	0.60
MCC	0.51	0.45	0.61
FPR	0.94	0.81	0.89

as quiescence, which can be a precursor to a large seismic event [16]; Seismic Rate Change B β proposed in [10]; Seismic Rate Change Z z proposed in [19] but the mathematical expression used is in [28]; Maximum Magnitude in the last d days $x6$ prior to the last window event proposed in [21] ;and the Probability of event greater than m $x7$ proposed in [21].

Every indicator that made use of a and b values as parameters has two variations, one using a_{lsq} and b_{lsq} values and the other using a_{mlk} and b_{mlk} variation, there are 23 seismic indicators calculated in total.

4.3 Labels

It is still very ambitious to use this type of analysis to predict a magnitude (ie y1), in this study it is analyzed whether a magnitude is above or below a threshold magnitude (ie y2), y2 is 1 if the magnitude is above the threshold and 0 if the magnitude is below the threshold.

The precision of the analysis is lower with high magnitudes, since it is more difficult to predict events with higher magnitudes.

The proposed label is $y2$ but we calculate first $y1$ the Maximum magnitude in the next d days is calculated, this when $T \in [T_n, Tn + d_a)$ and $d_a = days\ to\ look\ ahead$; and then use $y1$ to calculate $y2$ the occurrence event greater than m is calculated using the equation 2:

$$y_2 = \begin{cases} 1 & \text{if } y1 >= m \\ 0 & \text{if } y1 < m \end{cases} \quad (2)$$

5 Results

To make the prediction, a neural network was used and the 23 calculated indicators were introduced. The structure of the model can be seen in Fig. 7.

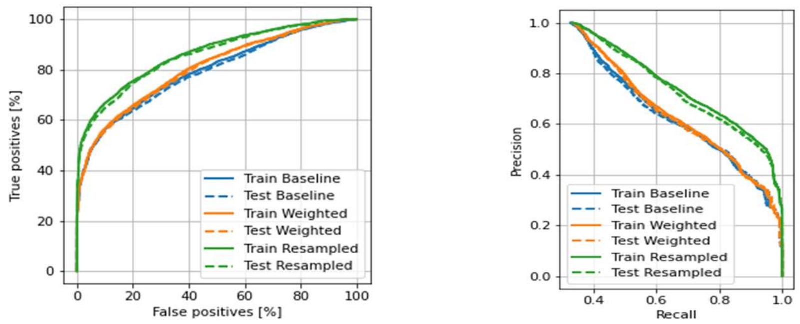

Fig. 8. The AUC curves on the left and the PRC curves on the right of the different models.

5.1 Neural Network

In this model, a sequential neural network was used. The neural network consists of two dense layers (deeply connected, meaning that each neuron in the dense layer receives information from all the neurons in its previous layer) with 23 neurons each hidden with “relu” activation, and one output layer with one neuron and sigmoid activation. There are 552 parameters for the first two layers and 24 for the third, adding up to a total of 1128 parameters.

To train a model on unbalanced data, it is fitted with a lot size large enough to ensure that each lot has a decent chance of containing some positive samples. Class 1 samples are called positive samples and class 0 samples are negative samples.

5.2 Metrics

Model performance was measured using the test sample obtained from 30% of the data, and training was performed on the training sample obtained from 70% of the data (20% of this sample was also used for validation), with the metrics, TP, FP, TN, FN, Accuracy, Precision, Recall, False Positive Ratio (FPR), Negative Predictive Value (NPV), MCC, F1 score and R score, Area Under the Curve (AUC), Precision Recall Area (PRC).

It is often possible to calibrate the parameters of a model and improve the results of one metric at the expense of another. The model has 3 variations that outperform each other in different metrics. Those models performed better on different metrics. The metrics of the models are shown in Table 1 and the models are explained later.

The purpose of analyzing the results through these mentioned metrics is that each performance metric highlights a certain aspect of the results. Therefore, the purpose is to highlight all the merits and demerits of the results obtained through the different variations of the proposed prediction models.

In the Baseline model the indicators are entered without modification, and there are no special parameters for the model. All models were trained using 200 epochs.

In the Weighted model the indicators are introduced without modifying, but in the model the weight parameter is added so that classes 1 and 0 have different weight when classifying. The objective is to identify earthquakes above the threshold, but according to the Gutenberg–Richter law we will not find many of those class 1 samples compared to those of class 0, for this reason we must give more statistical value to class 1. With these Class weights, accuracy, and precision are lower because there are more FP, but recall and AUC are higher because the model also found more TP.

Despite having a lower precision, this model has a higher recall (and identifies more earthquakes above the threshold). To calculate the weights for the class 0 it was used the expression $weight\_0 = (1/neg) * (total/2.0)$ resulting in a weight value of 0.74, and for class 1 the following $weight\_1 = (1/pos) * (total/2.0)$ resulting in a weight value of 1.54 In the Resampled model the indicators are modified, and the parameters were also modified.

The approach is related to the previous one, but instead of calculating weights for each class, a resampling of the data set is performed on the minority class and thus balancing the data set, the result will be approximately 50% of data class 1 and 50 % of data class 0. The number of steps per epoch was modified, this is the number of batches needed to see each negative example once. With this, instead of class 1 data being displayed in one batch with a large weight, it is displayed in many different batches with a small weight. This smoother gradient signal makes training the model easier.

The PRC and AUC curves of every model and their variations can be seen in Fig. 8.

Metrics from different studies are not directly comparable, in contrast to other studies here we use a sliding window of a fixed size to calculate the instances that will be introduced to the model, other studies divide the dataset in chunks of a fixed size to calculate the instances, this results in almost as half as instances and less sensitivity to the indicator changes.

Facing the problem as a classification problem using time series and a sliding window, is the model proposed, and treating it also as a problem with unbalanced data gave better results after using the weighted model and in greater degree the resampled model, in comparison to the baseline model.

There are downsides from using the resampling model, first it takes more time and memory to train, second as it resamples the data, it runs the risk of having an overfitting problem.

6 Conclusion

In this interdisciplinary analysis, earthquake prediction has been performed through the interaction of earthquake precursors and computational Machine Learning techniques. The Resampled model with a performance from 0.60 to 0.89 in different metrics, is the most suitable model for this type of prediction since it predicts more seismic events above the threshold without significantly sacrificing the other metrics, giving the most desirable results of all three models.

The performance of the resampled model in the company of using sliding windows made this comparable with that of other techniques in the other studies. In the future more progress will have been made and more data will be available, machine learning has proved to be very useful in many fields, and many efforts are being made to adequate the use of machine learning for earthquake prediction.

References

1. Asim, K. M., Martínez-Álvarez, F., Basit, A., Iqbal, T.: Earthquake magnitude prediction in Hindukush region using machine learning techniques. Hazards, vol. 85, no. 1, pp 471–486 (2017) doi: 10.1007/s11069-016-2579-3
2. Cicerone, R. D., Britton, J., Ebel, J. E.: A systematic compilation of earthquake precursors. Tectonophysics, vol. 476, no. 3–4, pp 371–396 (2009) doi: 10.1016 /j.tecto.2009.06.008
3. Lomnitz, C.: Fundamentals of earthquake prediction. John Wiley & Sons, New York, pp 326 (1994)
4. Géli, L., Piau, J. M., Dziak, R., Maury, M., Fitzenz, D., Coutellier, Q., Henry, P.: Seismic precursors linked to super-critical fluids at oceanic transform faults. Nature Geoscience (2014) doi: 10.1038/ngeo2244
5. Mignan, A., Woessner, J.: Estimating the magnitude of completeness for earthquake catalogs. Community Online Resource for Statistical Seismicity Analysis, no. 4, pp 1–45 (2012)
6. Arroyo, M., Godínez, K., Linkimer, L.: Completitud del catálogo de la red sismológica nacional de Costa Rica. Boletín de Geología, Universidad Industrial de Santander Bucaramanga, Colombia, vol. 39, no. 3, pp. 87–98 (2017)
7. Chouliaras, G.: Investigating the earthquake catalog of the National Observatory of Athens. Natural Hazards and Earth System Sciences, vol. 9, pp 905–912 (2009)
8. Mignan, A., Woessner, J.: Estimating the magnitude of completeness for earthquake catalogs. Community Online Resource for Statistical Seismicity Analysis, vol. 4, pp 1–45 (2012)
9. Wiemer, S., Wyss, M.: Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the Western United States, and Japan. Bulletin of the Seismological Society of America, vol. 90, no. 4, pp 859–869 (2000)
10. Rudolf-Navarro, A., Muñoz-Diosdado, A., Angulo-Brown, F.: Seismic quiescence patterns as possible precursors of great earthquakes in Mexico (2010)
11. Gutenberg, B., Richter, C.: Frequency of earthquake in California. Bulletin of the Seismological Society of America, vol. 34, no. 4, pp 185–188 (1944)
12. Wu, Y. M., Chen, S. K., Huang, T. C., Huang, H. H., Chao, W. A., Koulakov, I.: Relationship between earthquake b-values and crustal stresses in a young orogenic belt. Geophysical Research Letters, vol. 45, pp 1832–1837 (2018) doi: 10.1002 /2017GL076694
13. Smith, W. D.: The b-value as an earthquake precursor. Nature, vol. 289, no. 5794, pp 136–139 (1981). doi: 10.1038/289136a0
14. Asim, K. M., Adnan-Idris, I. T., Martínez-Álvarez, F.: Seismic indicators based earthquake predictor system using genetic programming and AdaBoost classification. Soil Dynamics and Earthquake Engineering, vol. 111, pp 1–7 (2018) doi: 10.1016/j.soildyn.2018.04.020
15. Asim, K. M., Martínez-Álvarez, F., Basit, A., Iqbal, T.: Earthquake magnitude prediction in Hindukush region using machine learning techniques. Natural Hazards, vol. 85, pp 471–486 (2016)
16. Asim, K. M., Idris, A., Iqbal, T., Martínez-Álvarez, F.: Earthquake prediction model using support vector regressor and hybrid neural networks. PLoS ONE, vol. 13, no. 7 (2018) doi 10.1371/journal.pone.0199004
17. Salam, M., Ibrahim, L., Abdelminaam, D.: Earthquake prediction using hybrid machine learning techniques. International Journal of Advanced Computer Science and Applications, vol. 12, no. 15. (2021) doi: 10.14569/IJACSA.2021.0120578
18. Marsan, D., Wyss, M.: Seismicity rate changes. Community Online Resource for Statistical Seismicity Analysis (2011) doi: 10.5078/corssa-25837590
19. Panakkat, A., Adeli, H.: Neural network models for earthquake magnitude prediction using multiple seismicity indicators. International Journal of Neural Systems, vol. 17, no. 01, pp 13–33 (2007) doi: 10.1142/S0129065707000890

20. Reyes, J., Morales-Esteban, A., Martínez-Álvarez, F.: Neural networks to predict earthquakes in Chile. Applied Soft Computing, vol. 13, no. 2, pp 1314–1328 (2013) doi 10.1016/j.asoc.2012.10.014
21. Matthews, M., Reasenberg, P.: Statistical methods for investigating quiescence and other temporal seismicity patterns. Pure and Applied Geophysics, vol. 126, pp 357–372 (1988)
22. Habermann, R.: Precursory seismic quiescence: Past, present, and future. Pure and Applied Geophysics, vol. 126, pp 279–318 (1998)
23. Morales-Esteban, A., Martínez-Álvarez, F., Reyes, J.: Earthquake prediction in seismogenic areas of the Iberian Peninsula based on computational intelligence. Tectonophysics, vol. 593, pp 121–134 (2013) doi: 10.1016/j.tecto.2013.02.036
24. Galkina, A., Grafeeva, N.: Machine learning methods for earthquake prediction: A survey. In: The Fourth Conference on Software Engineering and Information Management, Saint Petersburg, Russia (2019)
25. Telesca, L., Wen, Y.: Development of machine learning-based innovative methods of seismic forecasting in Italy and Mexico. Development of machine learning-based innovative methods of seismic forecasting in Italy and Mexico, Consiglio Nazionale delle Ricerche (2021)
26. Adeli, H., Panakkat, A.: A probabilistic neural network for earthquake magnitude prediction. Neural Networks, vol. 22, no. 7, pp 1018–1024 (2009) doi: 10.1016/ j.neunet.2009.05.003
27. Habermann, R. E., Wyss, M.: Reply [to "Comment on Habermann's Method for detecting seismicity rate changes"]. J Geophys Res, vol. 92, no. B9, pp 9446–9450 (1987) doi: 10.1029/JB092iB09p09446