Earthquake Analysis in Central Italy

Carlo Alessi, Mathilde Polizzi

Technical University of Catalunya

May 27, 2018

Abstract

During the Artificial Intelligence seminars given at the Technical University of Catalunya, different ways of using AI were explored to address environmental issues such as water management, sustainability, environmental risk assessment, modeling of industrial and natural systems. This paper aims to apply one of those techniques to an appropriate topic chosen by the students. The chosen topic is earthquake analysis in the Italian peninsula. Pattern recognition will be performed using clustering algorithms such as K-means. Earthquake forecasting will be done by means of Artificial Neural Networks.

Keywords: Earthquake prediction, Temporal pattern recognition, Clustering, Artificial Neural Networks, Time Series

I. Introduction

Earthquakes are the shaking of the surface of the earth resulting from the sudden release of energy in the earth's lithosphere that creates seismic waves [1]. An earthquake generates two different types of seismic waves: the P waves (Prime) and S waves (Second). The main waves are the P waves, which make the ground vibrate in the same direction in which they propagate. Conversely the S waves make the ground vibrate perpendicularly to their direction of travel. An earthquake is defined by several characteristics such as the epicenter and the magnitude. The epicenter is the point of the surface of the ground where the earthquake began and where the damage is the most important. The damage is related with the magnitude, which measures the energy released during an earthquake and corresponds to its strength. The magnitude is not to be confused with the intensity, which is a quantity that can be used to evaluate the effects felt on the surface. The magnitude and intensity are measured respectively with the Richter and Mercalli scale. The Richter magnitude scale allows to calculate the magnitude of an earthquake with a range from 1 to 9, and is determined from the logarithm of the amplitude of the waves. Concerning the intensity, the Mercalli intensity scale measures the effects of the earthquake with a range from I to XII.

The Italian peninsula has a long history of destructive earthquakes that dates back to at least the 62 AD, when a strong earthquake caused major damage to the city of Pompeii. It is believed that, after the strong earthquake of the 62 AD, a series of minor seisms eventually resulted in the eruption of the Mount Vesuvius in the 79 AD, which completely destroyed the city[2]. Since then there have been an uncountable number of earthquakes, especially in the central part of the peninsula, which is the focus of the study performed in this paper.

The research is based on three different approaches which are:

1. Finding temporal patterns in earthquakes time series by means of clustering algorithms
2. Predicting the location, time and magnitude of the next earthquake using Artificial Neural Networks (ANNs)
3. Predicting whether there will be an earthquake in the next five days

The outline of the paper is as follows. In section II is described the problem of earthquakes occurring in the central part of Italy, considering the implications of the geographical position as well as the economical and social effects they cause. In section III are described the essential features needed to analyze the problem. Then section IV discusses different Artificial Intelligence techniques to solve the three

[1] https://en.wikipedia.org/wiki/Earthquake

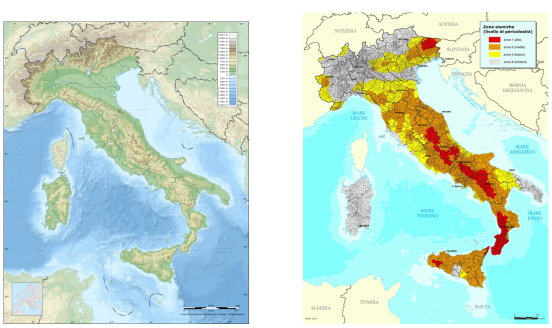
[2] https://en.wikipedia.org/wiki/62_Pompeii_earthquake

problems stated above. In section V are reported the results of a small number of experiments. In section VI there is a discussion on the advantages and limitations of our approaches. Alternative ways of study, using the techniques discussed in the seminar, are also mentioned. Finally section V concludes with a discussion and an insight for future work.

II. Environmental Problem: Analysis of Earthquakes in central Italy

Italy is one of the most seismic countries in Europe. As shown in Figure 1b [3], the zones with the highest risk are those along the landforms, especially in the Apennines, which span the entire peninsula. Moreover in Figure 1a it can bee see that most of the peninsula is covered by mountains, whereas in the the north is present a large valley.

The topography and geological conditions of the area influence the position of earthquakes epicenter. As explained by Donald L. Turcotte [1], the topography is created by a variety of tectonic processes resulting from displacements of preexisting faults. According to his study, earthquakes appear to be an example of self-organized critical behavior. This has important implications for estimating the earthquake hazard and contributes to earthquake predictions.



(a) Italy landforms. **(b)** Italy earthquake risk zones.

Figure 1: Comparison between the topographic and seismic map.

[3] http://www.eurometeo.com/english/read/doc_zone-sismiche

i. Environmental, economical and social impact of earthquakes

Earthquakes can have a wide range of consequences on the environment, economy and policies of the area affected.

Environmental issues From an environmental perspective, earthquakes can cause several damage due the formation of landslides, avalanches, soil liquefaction, tsunami and water floods in the coastal area. Landslides are particularly dangerous, as Italy has plenty of villages on and nearby the mountains. Moreover during the winter season, the risk of avalanches is high, and could cause damage in the areas where winter sports are practiced, as well as to nearby villages. Furthermore, entire structures, such as cars and buildings, could fall due to soil liquefaction.

Infrastructure issues Environmental damages are directly linked to the destruction of infrastructures in the private and public sectors. Particularly concerning could be presence of nuclear power, electrical or telecommunication stations. The stress caused by an earthquake on these in structures could lead to the leakage of chemicals into the atmosphere, which could contaminate the area affected for several years. Another source of danger are gas pipe explosions and fires in fuel stations, as well as flooding of aquariums. Beyond this, people might be injured or killed at the workplace. Moreover transports and communication links, such as harbors, highways, train railways and airports, as well as telecommunication cables, could be rendered unusable or completely destroyed. As a result, trading goods, such as food, becomes harder and people's life is disrupted. Furthermore, water pipes and electrical systems could be unusable.

Last but not least, the damage caused to buildings and homes could be considerable, since the cost of rebuilding a settlement is high. Many monuments and buildings in the Apennines were constructed in ancient times and cannot resist severe shocks. There have been recent earthquakes that destroyed entire villages. Restructuring the buildings to make them anti-seismic would cost less than rebuilding them from scratch. However the actual political instability and the general economic crisis of Italy may overshadow this topic.

Economical issues The economical impact is perhaps one of the most important in modern society, because the institutions are forced to invest resources (time, money and human) in rebuilding the destructed area. This resources could instead be spent on innovation and services. Moreover frequent earthquakes can decrease tourism.

According to Shannon Doocy and al. [4], earthquakes were responsible for 1.87 million deaths in the 20th century between 1990 and 2010. Unfortunately, with the actual technology and expertise, it is extremely challenging to predict the exact time, location and intensity of an earthquake. However, the research community is putting effort in developing reliable methods to study and possibly predict, and thus prevent, unexpected catastrophes to occur. In the next section a number of ways of study is briefly described. Some methods will be applied to study the seismic activity of Umbria, a region of central Italy.

III. Characteristics of the problem

i. Earthquake study

The study of earthquakes can help avoid or minimize catastrophes. It is based on warning systems, forecasting and prediction.

Warning systems Earthquake warning systems use devices such as accelerometers, seismometers, computers and alarms to monitor the wave projection and time lag. When an earthquake occurs, the release of P and S waves, defined earlier, are useful to locate the event. The aim is to identify the epicenter of the earthquake and warn municipalities and people about the event in due course. Yet this technique is not to be exploited in this paper and the focus is on pattern recognition and prediction.

Earthquake forecasting Earthquake forecasting is the branch of seismology science concerned with probabilistic assessment of general earthquake hazard. It includes the assessment of frequency and magnitude of damaging earthquakes in a given area over years and decades. It consists in finding trends an patterns by means of statistical techniques, which involve different measurements of an event.

Earthquake prediction Earthquake prediction aims to predict the time, location, and magnitude of future earthquakes. It used to be an immature science that focused on empirical analysis with two general approaches: either identifying distinctive precursors to earthquakes (e.g. animal behavior, radon emissions, electromagnetic anomalies), or identifying some kind of geophysical trend or pattern in seismic activity (e.g. elastics rebound, seismic gap) that might precede a large earthquake. Artificial Intelligence provides new methods to analyze and predict future earthquakes.

Features of an earthquake prediction system According to Allen Clarence [5], an earthquake prediction system must have the following features:

1. It operates on a specific location or area,
2. It covers a specific time span,
3. It deals with a specific magnitude range,
4. It must specify the probabilities associated with the predictions,
5. It must specify the confidence of its predictions.

Thereby it must specify where and when an earthquake will occur, and how sever it would be. Particularly important are also the probability of the occurrence of the predicted random event and the confidence of the predictions. The location window is usually defined in terms of latitude and longitude coordinates. The time span corresponds to the time interval of future predictions (e.g. 3 days forecast). The range of magnitudes addressed by the system depends on the objectives. Usually we are interested in predicting the catastrophic events as the minor seisms may cause little or no damage. However it could be risky to discard them from the analysis, because big shocks are often preceded and followed by a series of tiny earthquakes, called respectively foreshock and aftershock. Thus there could be interesting patterns between two major shocks. Finally the system should provide with which probability an event will occur, and whether the prediction is statistically significant.

It is worth mentioning that the predictions of such systems can fall in one of the following categories:

a. True Positive (TP): event correctly predicted
b. False Positive (FP): false alarm
c. True Negative (TN): correct prediction that an an earthquake does not occur
d. False Negative (FN): earthquake not predicted when it actually occurred

Depending on the outcome of the system, there could be several social impacts. In the case of TN, the population is safe and the only people concerned are the researches, who will need to understand the reason of such mistake. In the case of TP, hopefully the majority of the population would have evacuated, depending on the promptness of the authorities notification. Although FPs may be seen as a relief, frequent false alarms could be disruptive for society, that will be less trustful in future announcements. Thereby an earthquake prediction system should produce as few FP as possible. The worst case possible, of course, is the case of FN, that is when the system fails to detect an earthquake.

ii. Dataset description

To carry out the research it was used the open dataset from the Italian Seismological Instrumental and Parametric Data-Base (ISIDe[4]), which contains the parameters of earthquake locations performed by surveillance service of the Istituto Nazionale di Geofisica e Vulcanologia (INGV Rome).

The dataset contains events recorded by 500 stations of the National Seismic Network that occurred in Italy between 1985 and today. However our study focuses on a portion of the dataset, collected in 10 years from 01/10/2008 to 01/01/2018, on a circumference with a radius 90 Km centred at the location $42.9509°N, 12.7015°E$ [5]. The dataset is composed of 29969 earthquake observations, each with 14 attributes. Only 6 features are relevant for the study which are summarized in Table 1. The earthquakes of the dataset are visualized in Figure 2. It can be seen that, location-wise, there are 2 clusters of earthquakes. One major cluster, which almost spans cost-to-cost, and another small cluster. In particular the earthquakes with magnitude grater than 3 tend to occur on the east side. The distribution of earthquakes according to their magnitude is shown in Figure 3. The earthquake that often occur in central Italy have a magnitude lower to 3. They are considered as micro or minor earthquakes (yellow dots on Figure 2), meaning they are slightly felt and not damaging building but still recorded by seismographs. Earthquakes with magnitude greater or equal than 3 occur more rarely. This is true only in the period of time analyzed, as in 1997 there was an earthquake of magnitude 6.1 in the same region, followed and preceded by many other shocks with magnitude around 5. Including the data since 1997 would have been interesting, but our computational and memory resources did not allow such a large scale study.

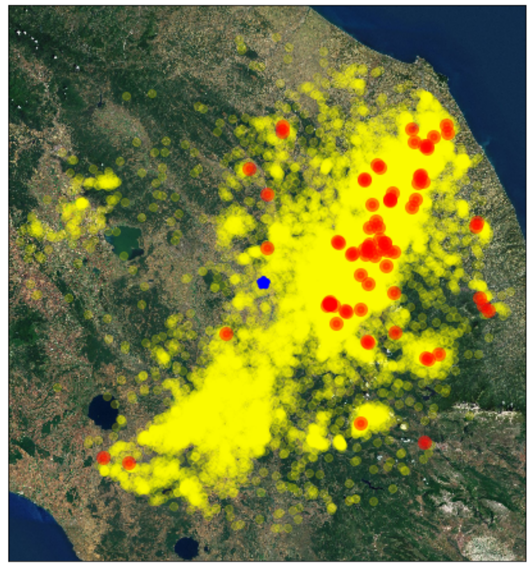


Figure 2: Plot of 29969 earthquakes occurred in 10 years near the town of Foligno (blue point). The yellow points correspond to earthquakes with magnitude less than 3. The red points are those with magnitude between 3 and 4.

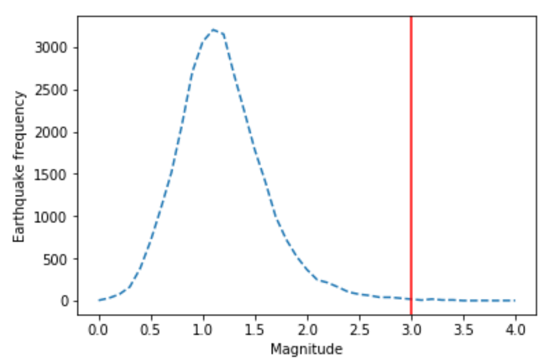


Figure 3: Frequency of earthquake against magnitude.

[4]ISIDe website: http://cnt.rm.ingv.it/en/iside

[5]These are the latitude and longitude coordinates of the town of Foligno, Umbria.

Table 1: Dataset features summary. Only the first five features are used by the machine learning algorithms.

Feature	Type	Range	Description
Time	Date-Time	[01/01/2008, 01/01/2018]	Earthquake time
Latitude	Quantitative	[42.15, 43.75]	Epicenter latitude
Longitude	Quantitative	[11.6047, 13.7917]	Epicenter longitude
Depth (Km)	Quantitative	[0.1, 70.9]	Earthquake depth
EventLocationName	Cathegorical	4267 locations	Location of epicenter

IV. Methods

i. Pattern recognition: clustering

The first adopted approach is temporal pattern recognition of seismic time series using a clustering method. In this section is first reviewed a work present in the literature that uses clustering for earthquake prediction. Finally the use of clustering to recognize patterns in earthquake time series is presented.

Morales-Esteban et Al. [7] proposed an earthquake prediction system based on K-means. Their objective was to find temporal patterns of Spanish earthquakes with magnitude greater or equal to 4.5, removing aftershocks and foreshocks from the analysis. They modeled an earthquake event using only three features (magnitude, b-value, time), without including spatial information. Although it is important to predict the location of an earthquake which was secondary in their study. This because very severe shocks would affect a large area regardless of the epicenter. In order to avoid local minima and the dependence from the starting centroid, they executed the K-means algorithm several times. Then they used the silhouette score to determine the optimal number of clusters. The results reported showed that earthquakes belonging to a particular class were preceded by a series of earthquakes of another class, or by a combination of earthquakes of different clusters. Moreover they noticed that medium-large earthquakes were preceded by a decrement in b-value. The level of significance of the results was assessed with the Wilcoxon rank-sum test.

Proposed method The clustering study is based on the K-means algorithm. The optimal number of clusters is chosen selecting a value of K at which is achieved a low Sum of Squared Errors (SSE) and a high silhouette score.

Since the the SSE monotonically decreases as we increase the value of K, we could not rely only on minimizing the SSE. The elbow method [7] was used to visually find a value of K at which the improvement of SSE by selecting K+1 would be minimal. This was achieved by plotting the SSE against the number of clusters. This method could be challenging because some plots may not have a clear elbow. The method was complemented by adding the criterion of the silhouette score.

Then the centroids of the clusters are analyzed for different interesting values of K, in order to reveal some differences or similarities between the clusters. Once the clustering is performed, the earthquakes are identified only by their label. This allowed us to plot the time series of earthquake labels, and find patterns on the occurrence of different earthquake classes overtime.

ii. Prediction using artificial neural networks

This section briefly reviews an approach presented in the literature to predict earthquakes using Artificial Neural Networks (ANNs). Finally our method will be presented.

Reyes and Al. [6] proposed an ANN-based system to predict earthquakes in Chile. They used 4 standard ANNs trained with backpropagation, each specialized in a different region divided in meshes with dimensions varying from $0.5° \times 0.5°$ to $1° \times 1°$. All networks shared the same architecture, with 7 input neurons, one hidden layer of 15 neurons, and 1 output neuron. The output of the network predicted the maximum earthquake magnitude that would occur in the next 5 days (output was 0 if no earthquake occurred). The predictions also provided:

a. The probability that an earthquake with magnitude larger than a threshold T happens.

b. The probability that an earthquake with magnitude within an interval occurs.

They engineered input vectors of length 7 combining different domain knowledge, such as the Gutenberg-Richter's law, the Bath's law, and the Omori-Utsu's law. Moreover, they adjusted the threshold T to obtain as few False Positives as possible.

Proposed method The method proposed in this paper is based on Long-Short Term Memory (LSTM) networks, which are able to deal with the sequential nature of the data. The aim is to predict whether an earthquake with magnitude grater or equal than 3 will occur in the next five days, based on the latest 50 observations. The input of the network is thus a sequence of 50 vectors, each composed of the 5 features described in Table 1. The output of the network are two numbers, which respectively correspond to the probability that an earthquake will occur or not. The complete network architecture is summarized in Table 2. The dataset is split in 3 disjoint partitions, forming the training, validation and test set. Each partition accounts for respectively 70%, 20% and 10% of the data. The data is normalized between 0 and 1 as a preprocessing step.

A similar network architecture was also used to predict the next element in the time series, based on the latest observations.

Table 2: Recurrent neural network architecture. The activation function of LSTM and Dense layers are respectively the hyperbolic tangent and ReLU.

Layer	Detail	Layer	Detail
1. LSTM	input: 50×5 units: 200 dropout: 0.1	4. Dense	units: 512 dropout: 0.5
2. LSTM	units: 100 dropout:0.1	5. Dense	units: 256 dropout: 0.5
3. LSTM	units: 50 dropout: 0.1	6. Dense	units: 2

V. Results

In this section are reported the results of the exploratory data analysis, temporal pattern recognition and prediction.

i. Data understanding

Before performing any machine learning, a basic descriptive statistics analysis was done in order to better understand the dataset at hand. Figure 10 on page 10 shows an overview of the analysis. From the box plot in Figure 10a it can be seen that the dataset is plenty of outliers, which were not removed from the study. The features distribution and histogram respectively in Figure 10b and Figure 10c show that Depth and Magnitude have a unimodal distribution. Interestingly the magnitude distribution clearly resembles a normal distribution. Conversely there is no clear pattern for the latitude and longitude, which would have been the most important thing. In Figure 10d is shown the correlation analysis computing the Pearson coefficient. The result is that only latitude and longitude are moderately negatively correlated ($r = -0.64$), whereas there is no correlation between the rest of the features. Finally Figure 10e visualizes individual features as time series. In can be seen that series are chaotic, since it is very hard to predict next value from the past observation, although there are some regular peaks in the *depth* time series.

ii. Clustering results

First step was to find the best number of clusters using two different approaches, elbow and silhouette. As seen on the plot Figure 4, the elbow method is not the most relevant since the curve does not present an abrupt variation of SSE followed by minimal SSE improvements. This is why we used in complement the silhouette method.

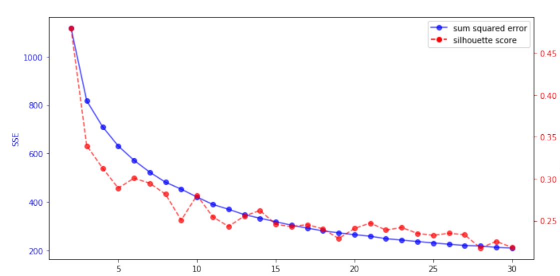


Figure 4: Elbow and Silhouette comparison

Morales-Esteban [7] decide the optimal number of clusters where the index reaches the maximum value. In our case, deciding this optimal number was more complex and it was decided to select three different values of K (number of clusters) which are

6, 10 and 14, corresponding to small spikes of the silhouette curve (red) along the elbow curve (blue) and then we perform the analysis in the further discussion.

iii. Prediction results

Given the time and computational constraints it was used only a portion of 3000 samples for training, and a portion of 1000 samples for validation. The training and validation set are continuous in time, meaning that the first example of the validation set was observed after the last example of the training set. Figure 5 shows the learning curve of the neural network. It can be seen that the learning process is fairly good in the early stages of training, achieving training and validation accuracy up to respectively 93% and 88%. However, after about 20 epochs of training, the learning curve had sharp swings. Eventually the optimization diverged into poor local minima, which the network could not escape, as there learning curve formed a plateau.

Figure 6 shows the learning curve of the network for the time series forecasting task. It can be seen that although the training loss slowly decreased as the training progressed, the validation loss remained almost stuck.

VI. Discussion

i. Clustering

Next step of the clustering method consists in analyzing the centroids of the clusters, their size and also the series of earthquake types. The analysis was carried out considering three different values of K, in order to discover temporal patterns that allow forecasting earthquakes.

The comparison of the clusters centroids for each value of K is shown in Figure 7. It can be noticed that for $k = 6$, the depth attribute is not relevant since all clusters have the same depth on average. For $k = 10$ clusters, one of the clusters has a different depth on average. About the coordinates (Latitude and Longitude attributes), all clusters have very different centroids meaning the clusters are spread in different location in center Italy.

It can be noted on Figure 8 that the size of each cluster is very heterogeneous even when increasing the number of clusters K.

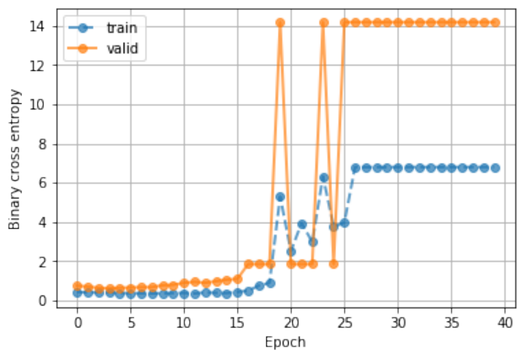
Earthquake series reveal the type (which cluster does the earthquake belong to) of the latest 200 earthquakes that occurred. It is tricky to draw any kind of clear pattern in the three cases analyzed. When having a look at Figure 9b, earthquake of type 9 seem to occur why an evident frequency (more or less the same distance between the dots that seem to get wider along time). In between those type of earthquake occur a large number of type 4 earthquakes and even more type 2 earthquakes. Extending the sample to more than 200 earthquakes could help reveal more significant shapes and infer a better pattern recognition.

ii. Prediction

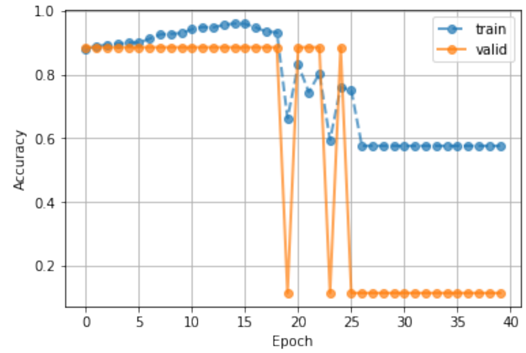
It was found that training a neural network on earthquake time series is an extremely challenging task. Even though acceptable performance was obtained after few epochs of training, only a small partition of the dataset was used. Moreover it was not straightforward to select a good network architecture, as several problems such as overfitting or instability often occurred. We believe that restricting the area of the region covered by the network, and specializing different networks in narrower regions, can potentially boost the performance. Furthermore, future work would focus on maximizing other metrics, such as precision and recall, which minimize respectively the number of false positives and false negatives.

iii. Other applicable methods

Other techniques that have been presented in the series of Artificial Intelligence seminars could have been applied to analyze the problem addressed in this paper. Density-based clustering, such as the DBSCAN algorithm [8], would be the first candidate, as it is natural to cluster earthquakes taking into account their spatial density. The challenges that would be addressed would be the selection of the optimal ϵ parameter, for which we would use an approach similar to the elbow method discussed before. Another possibility would be to use hierarchical clustering. Design choices would include selecting the linkage method (min, max, average, Wards) and the cutoff distance in the dendograms. Further possibilities would explore the usage of different neural models, such as standard feed-forward neural networks, or Neuro-Fuzzy Adaptive Network (ANFIS) which is able to reproduce highly nonlinear time series.



(a) loss



(b) accuracy

Figure 5: Learning curve of the neural network.

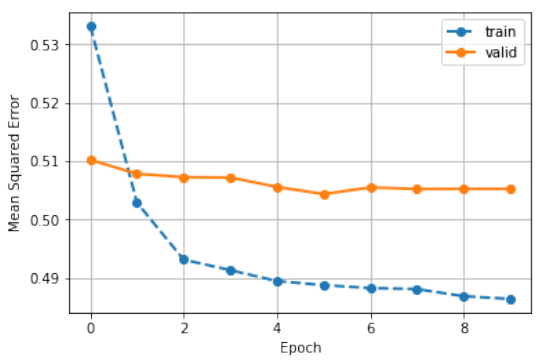
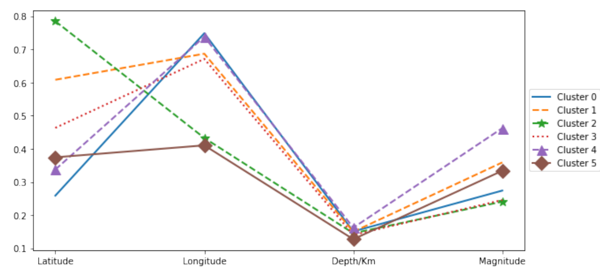


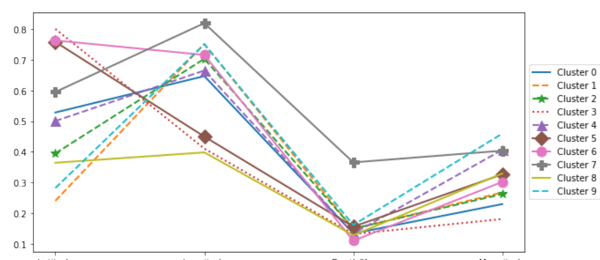
Figure 6: Learning curve of the network for the task of time series forecasting.

VII. Conclusions

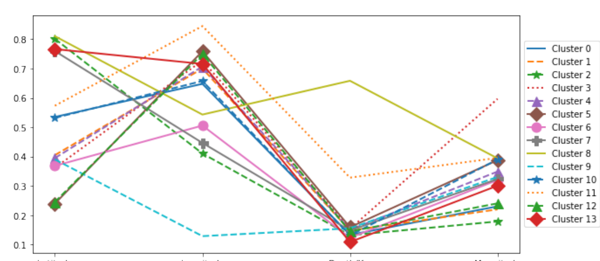
In this paper we presented described how earthquakes affect the area where they occur from different perspective (environment, infrastructures, policies, social, economy). We then proposed two methods for earthquake temporal pattern recognition and prediction, using clustering and neural models. The methods were applied to a dataset from the Italian Seismological Instrumental and Parametric Data-Base (ISIDe). Although there were not remarkable results, we have been able to study and analyze the seismic activity of the Umbria region, which is an area of interest for us. Future work would explore the effects of injecting domain knowledge, such as the Gutenberg-Richter's law, the Bath's law, and the Omori-Utsu's law, to create more complete datasets. Furthermore, targeted experiments would try to



(a) 6 clusters.

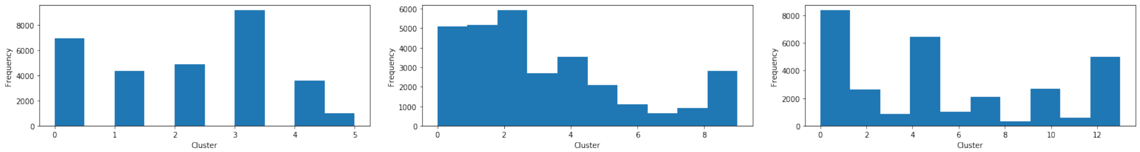


(b) 10 clusters.



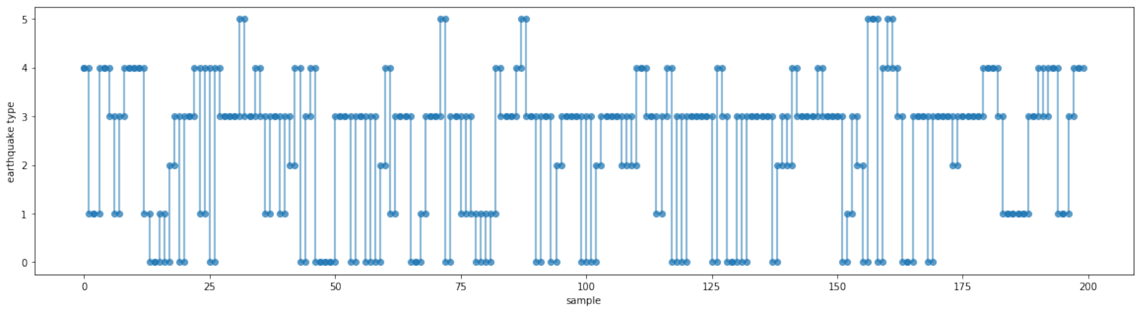
(c) 14 clusters.

Figure 7: Analysis of centroids.

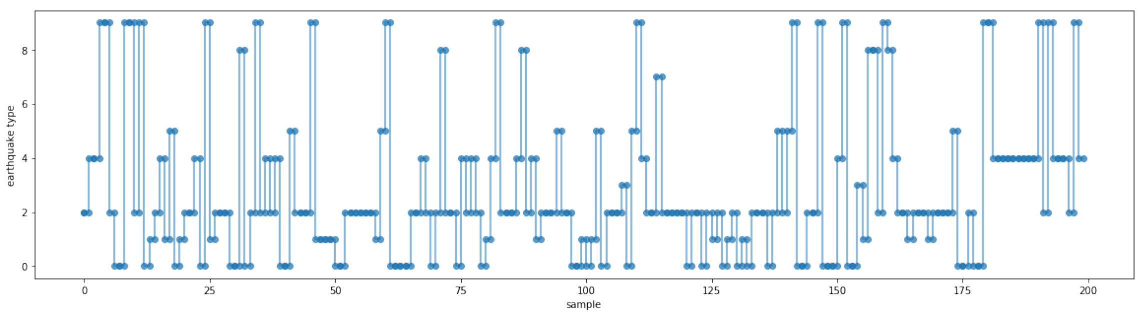


(a) Size 6 clusters. **(b)** Size 10 clusters. **(c)** Size 14 clusters.

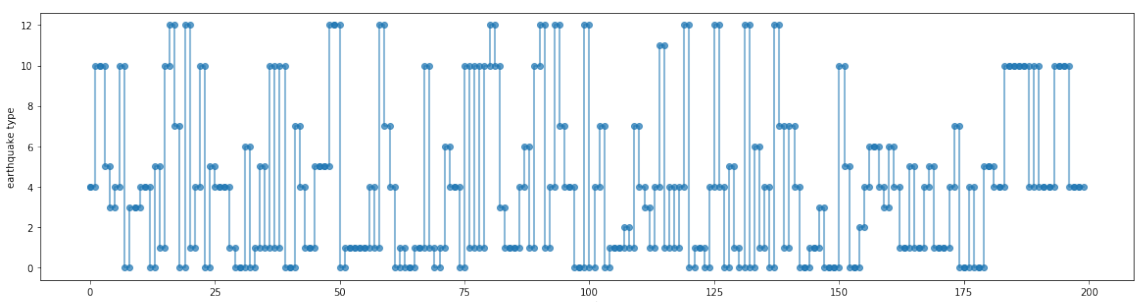
Figure 8: Clusters size comparison.



(a) Series for 6 clusters.



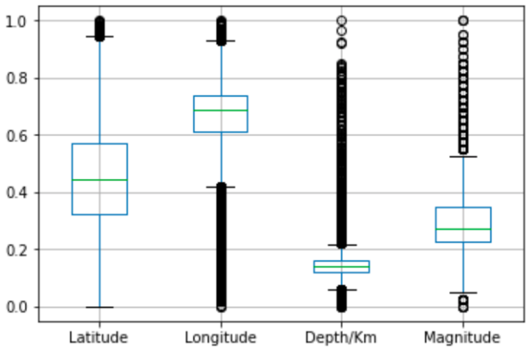
(b) Series for 10 clusters.



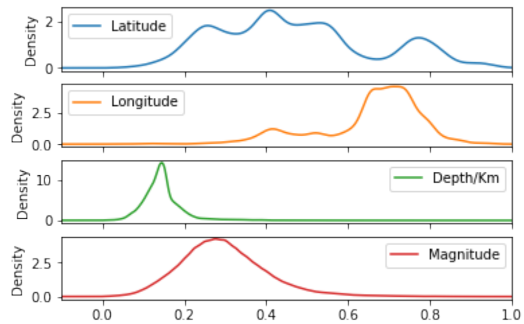
(c) Series for 14 clusters.

Figure 9: Series of earthquakes types.

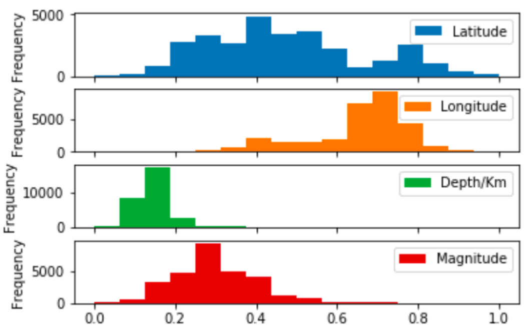
uncover whether there are complex temporal dependencies between earthquakes of different classes. Using clustering for earthquake pattern recognition, and neural networks for earthquake prediction seems to be the research direction to follow to prevent catastrophes. However it is still an open challenge and new ideas are needed.



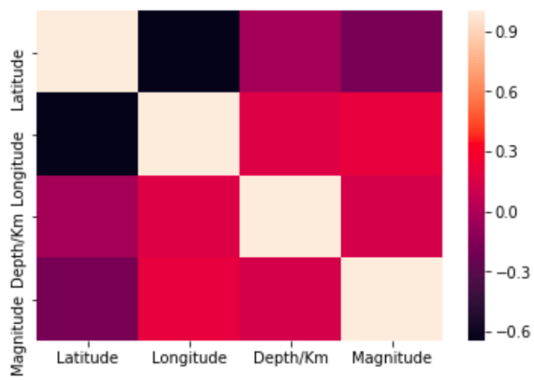
(a) Boxplot



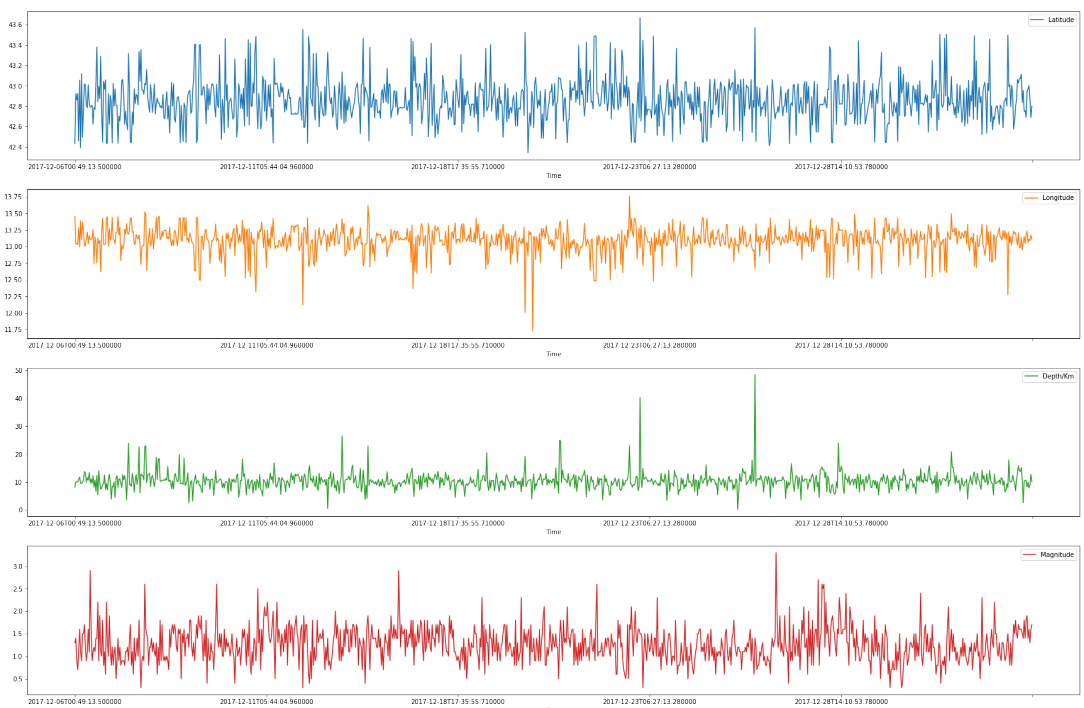
(b) Distribution



(c) Histogram



(d) Correlation



(e) Time series of each feature. The plot corresponds to the last 1000 earthquakes of the dataset, occurred in less than one month.

Figure 10: Data overview.

References

[1] Donal L. Turcotte. *Scaling in geography: Landforms and earthquakes Physics: The Opening to Complexity (1995) Vol 92, pp. 6697-6704*

[2] Yasuhiro Mitani, Fawu Wang, Austin Chukwueloka Okeke, Wenhao Qi (2012). Dynamic Analysis of Earthquake Amplification Effect of Slopes in Different Topographic and Geological Conditions by Using ABAQUS, 469-490

[3] National Academy of Sciences (1992). *The Economic Consequences of a Catastrophic Earthquake., 4 differential impact of earthquake events: 112–126.*

*[4] Shannon Doocy, Amy Daniels, Catherine Packer, Anna Dick, Thomas D. Kirsch (2013).*The Human Impact of Earthquakes: a Historical Review of Events 1980-2009 and Systematic Literature Review.

[5] Allen, Clarence R. Responsibilities in earthquake prediction: to the seismological society of America, delivered in Edmonton, Alberta, may 12, 1976. Bulletin of the Seismological Society of America 66.6 (1976): 2069-2074.

[6] Reyes, Jorge, A. Morales-Esteban, and Francisco Martinez-Alvarez. *Neural networks to predict earthquakes in Chile. Applied Soft Computing 13.2 (2013): 1314-1328.*

[7] Morales-Esteban, A., et al. Pattern recognition to forecast seismic time series. Expert Systems with Applications 37.12 (2010): 8333-8342.

[8] Ester, Martin, et al. *A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd. Vol. 96. No. 34. 1996.*