BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

STUDY ORIENTED PROJECT

(MATH F266)

FINAL REPORT

Sequence Models for Earthquake Prediction Using Spatio-Temporal Data

Author

ANKIT SONTHALIA
2017B4A70468P

Supervisor

DR. SUMANTA PASARI

June 29, 2020

EARTHQUAKE PREDICTION USING DEEP LEARNING TECHNIQUES ON SPATIO-TEMPORALLY ARRANGED DATA

Ankit Sonthalia
M.Sc. (Hons) Mathematics
B.E. (Hons) Computer Science (2022)
Birla Institute of Technology
and Science, Pilani
Rajasthan - 333031, India
f20170468@pilani.bits-pilani.ac.in

Sumanta Pasari
Assistant Professor
Department of Mathematics
Birla Institute of Technology
and Science, Pilani
Rajasthan - 333031, India
sumanta.pasari@pilani.bits-pilani.ac.in

June 29, 2020

ABSTRACT

This paper explores the problem of developing an early warning system for earthquakes. The problem is not a new one, and has been attempted for a few decades now, but has seldom been solved effectively. In this paper, we build upon a previously presented idea [1], adding modifications and refinements to the original technique. The central idea is that earthquakes are spatio-temporally correlated events, i.e., the occurrence of earthquakes in a given region within a given time frame depends on previous occurrences both in the region itself and surrounding regions. We find that the method works well for light earthquakes, but performance degrades for moderate earthquakes and the method fails completely for strong earthquakes.

***K*eywords** earthquakes · spatio-temporally · deep learning · LSTM

1 Introduction

Earthquakes pose a major threat to human life and property because of their far-reaching damages and unpredictability. They arrive unannounced, leaving buildings collapsed to their foundations in their wake and claiming lives to the tune of hundreds or even thousands, depending on the magnitude.

A study on potential techniques for predicting and subsequently avoiding earthquakes is, thus, an important one. Earlier works have attempted to solve this problem in a plethora of ways ranging from precursor analysis to artificial neural networks. Much of the success achieved so far, if any, in this study, can be accredited to the latter part of the spectrum: artificial neural networks. The underlying reason is that earthquakes are inherently heavily complicated and must believably depend on several hundred different parameters. Since identifying these parameters would be a difficult feat, any human-designed model of such dependencies should not be expected to yield very promising results.

The need for automating the process of exploiting intricate data dependencies is a far more general one, and largely presents itself to the mainstream world in the form of face recognition, translation and transliteration, self-driving cars. The approach taken by an artificial neural network towards attempting this task is making repeated passes through given data, attempting to "solve" several "questions" and then comparing its answers with actual ground truth values. Based on the correctness of its answers (or lack thereof), the network then makes necessary adjustments to the weights used for calculating these answers.

Predicting future earthquakes given a history of past earthquakes fits the job description of a neural network. Dependencies are learnt by the network itself, without assistance from a human. A neural network, then, identifies and exploits complex relationships that we otherwise could not, and can thus yield far better results.

In the current paper, our main contributions will be as follows:

1. We test the model presented in [1] on another dataset.
2. We develop and make available, code for possible future works [1].
3. We consider a refinement to the prediction system in order to enhance accuracy and obtain results that are more relevant.

2 Related work

In this section we list in detail the relevant works which were referred while writing this paper.

Wang et al. [1] used spatio-temporal segmentation to predict the occurrence of earthquakes in Mainland China. The idea for this paper borrows largely from [1] and has been described more in the next section.

Rasel et al. [2] used a similar idea on five-dimensional earthquake data from the Bangladesh Meteorological Department. Although the authors reported very high accuracy, they concluded that the results would be much more reliable if they had been based on other factors apart from historical data. Most of the earthquakes were largely in the 4.0 - 5.0 magnitude bracket on the Richter Scale, which helped the results significantly.

Behrich et al. [3] used LSTM for predicting magnitude, location and time of the earthquake, and achieved satisfactory results.

Gonzale et al. [4] used recurrent neural networks for maximum earthquake magnitude prediction using hourly earthquake data.

Florido et al. [5] have conducted a survey of the methodologies and successes involved in the use of artificial neural networks to predict earthquakes

Bhatia et al. [6] have attempted earthquake prediction using artiticial neural networks and have contended that the results, while satisfactory, definitely contained room for improvement.

Vardaan et al. [7] have attempted earthquake prediction using LSTMs and established the superiority of LSTM over Feed Forward Neural Networks (FFNNs).

3 System Model and Data

While the problem formulation has originally been described in [1] and [2], we have described it here for the sake of completion. The description will be followed by a brief description of the LSTM.

3.1 Spatio-Temporal Formulation

Given a rectangular area of interest designated by its latitude and longitude boundaries, we divide up the area into several grids or **subregions**. This constitutes the *spatial* part of *spatio-temporal*. Given s such subregions, we then have a multi-hot vector x with s components, where the component x_i is 1 if the subregion i has recorded an event; zero otherwise. The vector x is then, an *image* of the area of interest during a given interval of time.

For the *temporal* part of **spatio-temporal** segmentation, we then divide up the timeline of interest into T intervals of a chosen length. For the interval t, x^t is then the multi-hot vector describing an image of the region of interest during the interval t.

The set of all these multi-hot vectors can be described as a **map** f from the set A of all natural numbers from 1 through T, to the set B of all multi-hot vectors x^1 through x^T, where

$$f(t) = x^t. \tag{1}$$

The map f, for all intents and purposes, forms a **time series**. We then formulate the problem before us as follows:

Given an area of interest and its corresponding time series, we need a model solution to provide us with an extension of the time series, i.e., given T intervals, we need to predict x^{T+1}, x^{T+2} and so on.

For the purpose of effectively training our model solutions, we will not be feeding the entire time series at once. Instead, we will feed in b images at a time, where b is the size of the **lookback window** and can be tuned as a hyperparameter.

[1]The code is available here

The task of the model solution will be to predict the multi-hot vector immediately after any particular series of b vectors. Shitting the window one vector at a time, we will have obtained N examples in total, where

$$N = T - b + 1. \tag{2}$$

3.2 Network Decomposition

The dataset was divided into three parts, depending upon the magnitudes of the earthquakes. Each of the neural networks was subsequently trained separately on these three parts.

1. Light earthquakes (magnitude 4.0 - 5.0): These are the easisest to predict.
2. Moderate earthquakes (magnitude 5.0 - 6.0): These are relatively less abundant and harder to predict.
3. Strong and Major earthquakes (magnitude 6.0 - 8.0): These are the least abundant kind of earthquakes and prediction is both an extremely difficult and extremely fruitful task; the current model proves ill-suited for the task.

3.3 Dataset

The dataset used in the paper contains location, time and magnitude data for earthquakes that occurred in Indonesia and surrounding regions (between latitudes -10^o **N** and 5^o **N** and longitudes 115^o **E** and 130^o **E**) during the period between 1963 and 2019. 80% of the data was used for training; the rest was used for testing.

3.4 Long Short Term Memory (LSTM)

The Recurrent Neural Network was original based on David Humelhart's work in 1986 as a solution to machine learning problems that include time series data of some sort. The solution proved useful to quite some extent, but eventually its limitations became clear. The major issue was the vanishing gradients problem which prevented scaling to larger inputs. The problem was addressed with the introduction of the Long Short Term Memory (LSTM) by Sepp Hochreiter and Jürgen Schmidhuber in 2000. The LSTM, while similar in spirit to the RNN, maintains an additional cell state for capturing long term dependencies. Earthquakes often tend to be influenced by past events along and around the fault line. Sometimes these events might date back many years. Such long-term dependencies call for the special capabilities of an LSTM. This premise has been explored in [1], [2] and [7].

4 Results and Discussion

4.1 Evaluation Metrics

Model performance has been evaluated with respect to two evaluation metrics: **precision** and **recall**.

Precision. Precision refers to the reliability of the prediction system. A good prediction system will produce results that can be relied upon, without worrying about a false positive. Thus, if a prediction system with high precision tells us that an earthquake might occur within the next few weeks, we can rely on the prediction and take appropriate action instead of having to worry about evacuating a large number of people only in preparation for an event that is never going to occur. Mathematically, precision can be expressed as:

$$Precision = \frac{true\ positives}{true\ positives + false\ positives} \tag{3}$$

Recall. Recall can be expressed as:

$$Recall = \frac{true\ positives}{true\ positives + false\ negatives} \tag{4}$$

Hence, a system with good recall will correctly identify a good portion of the actual positives. In our case, then, such a system will warn us correctly about most upcoming seismic events.

F1 Score. There is often a trade-off between precision and recall, with an attempt to increase one of them leading to a decrease in the other. This calls for a simpler metric: the F1 score, defined as the harmonic mean of precision and recall, i.e.,

$$F1\ score = \frac{2\ *\ precision\ *\ recall}{precision\ +\ recall} \tag{5}$$

4.2 Results

For the purpose of spatio-temporal segmentation, the Indonesian dataset was spatially divided into 3 rows and 3 columns to form a total of 9 subregions, while the timeline was divided into 1200 intervals, making each interval about half a month long in duration. A lookback window of 50 intervals was used. The LSTM had 50 hidden parameters in a single layer, while the linear layer had 40. An input fed into the network would first pass through the LSTM ; the result would be passed through an activation layer, followed by a dense layer to generate the output in the desired dimensions, and finally, another activation layer.

Table 1: Precision, Recall and F1 Score Values for LSTM

Data	Precision	Recall	F1 Score
All magnitudes	0.913	0.921	0.917
Light (4-5)	0.898	0.927	0.912
Moderate (5-6)	0.510	0.613	0.556
Strong and Major (6-8)	0.096	0.319	0.147

As can be observed from the table,

- With an F1 score of 0.917, the model performs fairly well when fed the entire dataset (all magnitudes) at once, but this result is hardly useful. The knowledge of an imminent seismic event serves much greater use when its magnitude can be limited to a value, or a reasonably small range of values.
- When used on just light earthquakes,the model still performs almost as good with an F1 score of 0.912. This is plausibly because light earthquakes are very abundant.
- For moderate earthquakes, the F1 score decreases to 0.556 which is quite a degradation over the previous two cases. This is largely because moderate earthquakes are not as abundant.
- For strong and major earthquakes, the model, for all intents and purposes, fails with an F1 score of 0.147.
- The unsatisfactory results in the last two cases might be indicative of the fact that it takes knowledge of the occurrence of light earthquakes in order to predict that of larger earthquakes. The fact that the former were excluded, caused the prediction to fail.

5 Conclusion and Future Works

In this paper, we applied an LSTM based prediction system on the Indonesian dataset, and then further decomposed the network based on magnitude ranges in order to demonstrate how larger earthquakes are harder to predict. In the future, a prediction system can be designed to take smaller earthquakes into consideration while predicting larger earthquakes. This can be attempted with modifications to the code accompanying this paper.

References

[1] Qianlong Wang, Yifan Guo, Lixing Yu, and Pan Li. Earthquake prediction based on spatio-temporal data mining: an lstm network approach. *IEEE Transactions on Emerging Topics in Computing*, 2017.

[2] Risul Islam Rasel, Nasrin Sultana, GM Azharul Islam, Mahfuzul Islam, and Phayung Meesad. Spatio-temporal seismic data analysis for predicting earthquake: Bangladesh perspective. In *2019 Research, Invention, and Innovation Congress (RI2C)*, pages 1–5. IEEE, 2019.

[3] Asmae Berhich, Fatima-Zahra Belouadha, and Mohammed Issam Kabbaj. Lstm-based models for earthquake prediction. In *Proceedings of the 3rd International Conference on Networking, Information Systems & Security*, pages 1–7, 2020.

[4] Jesús González, Wen Yu, and Luciano Telesca. Earthquake magnitude prediction using recurrent neural networks. In *Multidisciplinary Digital Publishing Institute Proceedings*, volume 24, page 22, 2019.

[5] Emilio Florido, José L Aznarte, Antonio Morales-Esteban, and Francisco Martínez-Álvarez. Earthquake magnitude prediction based on artificial neural networks: A survey. *Croatian Operational Research Review*, pages 159–169, 2016.

[6] Anmol Bhatia, Sumanta Pasari, and Anand Mehta. Earthquake forecasting using artificial neural networks. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2018.

[7] Tanvi Bhandarkar, K Vardaan, Nikhil Satish, S Sridhar, R Sivakumar, and Snehasish Ghosh. Earthquake trend prediction using long short-term memory rnn. *International Journal of Electrical and Computer Engineering*, 9(2):1304, 2019.