

# EARTHQUAKE FORECASTING USING ARTIFICIAL NEURAL NETWORKS

Anmol Bhatia<sup>1</sup>, Sumanta Pasari<sup>1,\*</sup>, Anand Mehta<sup>2</sup>

<sup>1</sup> Dept. of Mathematics, Birla Institute of Technology and Science, Pilani, Rajasthan, India  
(f2014745, sumanta.pasari)@pilani.bits-pilani.ac.in

<sup>2</sup> Dept. of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India  
anand.me@ktr.srmuniv.ac.in

Commission V, WG V

**KEY WORDS:** Earthquake Forecasting, Neural Networks, Multilayer Perceptron, Recurrent Neural Networks, LSTM Networks

## ABSTRACT:

Earthquake is one of the most devastating natural calamities that takes thousands of lives and leaves millions more homeless and deprives them of the basic necessities. Earthquake forecasting can minimize the death count and economic loss encountered by the affected region to a great extent. This study presents an earthquake forecasting system by using Artificial Neural Networks (ANN). Two different techniques are used with the first focusing on the accuracy evaluation of multilayer perceptron using different inputs and different set of hyper-parameters. The limitation of earthquake data in the first experiment led us to explore another technique, known as nowcasting of earthquakes. The nowcasting technique determines the current progression of earthquake cycle of higher magnitude earthquakes by taking into account the number of smaller earthquake events in the same region. To implement the nowcasting method, a Long Short Term Memory (LSTM) neural network architecture is considered because such networks are one of the most recent and promising developments in the time-series analysis. Results of different experiments are discussed along with their consequences.

## 1. INTRODUCTION

Earthquake occurs due to the relative movement of the tectonic plates that make up the Earth's crust. The most of the damage although occurs at places located along plate boundaries, the stable continental regions also occasionally experience disastrous events. The stress caused by this movement travels large distances and therefore places at larger distances from the plate boundary may also suffer.

Among several natural disasters, earthquake is the most crucial one that causes damage within a few minutes. While the primary effect of earthquake includes intense ground shaking, building collapsing and land splitting, the secondary effect may involve landslides, land-subsidence, fire, gas leakage, electric grid blackouts and tsunami. A recent survey from United States Geological Survey (USGS) shows that the last decade experienced approximately 450,000 deaths due to earthquakes. It not only breaks the backbone of the socio-economic ecosystem of a nation, but also ignites the lack of earthquake hazard preparation. This threat cannot be averted by mankind, but if properly analyzed, the damage can be substantially minimized.

Artificial Neural Networks (ANN) are increasingly used in predicting and classifying tasks because of their ability to capture the inherent complex relationship of a process with the set of inputs (Lakshmi and Tiwari, 2006; Madahizadeh and Allamehzadeh, 2009; Alarifi et al., 2012; Niksarlioglu and Kulahci, 2013; Reyes et al., 2013; Sriram et al., 2013; Zamani and Sorbi, 2013; Amar et al., 2014; Florido et al., 2016; Kurach and Pawlowski, 2016; Narayanakumar and Raja, 2016; Asencio-Cortés, et al. 2017; Perol et al., 2017). The ANN modeling requires finding two important factors: set of inputs and set of hyper-parameters. Therefore, the performance of ANN is evaluated based on different inputs and also for different set of hyper-parameters given to the

network (Lakshmi and Tiwari, 2006). For this particular task, there are so many factors involved in the process that other model based approaches cannot accommodate as accurately as neural network does (Perol et al., 2017).

In literature, there are very limited studies available that specifically compares the performance of different neural networks on the basis of different set of inputs and the number of hidden layers (Lakshmi and Tiwari, 2006; Reyes et al., 2013; Asencio-Cortés et al., 2017; Perol et al., 2017). This study therefore is an attempt to address that gap to an extent. For this purpose, the study presents a systematic comparison of different neural network architectures with different hyper-parameter and different set of inputs.

Moreover, the application of neural networks in the field of nowcasting earthquakes is a developing area, and virtually no literature is available for determining the natural time statistics for seismic hazard analysis. This study considers Long Short Term Memory (LSTM) architecture, along with the different set of hyper-parameters to obtain the least error in prediction (Wang et al., 2017).

The neural network models developed in this study can prove beneficial to the community because it can be used to create an early-warning alarm system so that the loss is minimized (Reyes et al., 2013). The following section lists out the efforts done in order to achieve the objective.

## 2. METHODOLOGY

This study uses two different techniques for earthquake forecasting and analysis. First technique compares the performance of multilayer perceptron based on different set of inputs and hyper-parameters. Later, the accuracy of nowcasting technique was evaluated using recurrent neural networks, namely the LSTM neural networks (Moustra et al., 2011; Kurach and Pawlowski, 2016;

\*Corresponding author

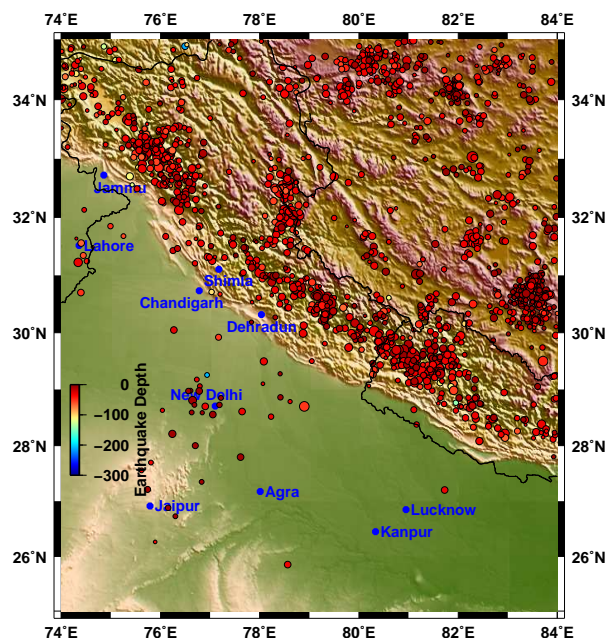


Figure 1. The selected region for Experiment 1

Wang et al., 2017). Detailed discussion of these two methods is provided below.

## 2.1 Multilayer perceptron

**2.1.1 Inputs** A multilayer perceptron is a class of feedforward ANN that uses backpropagation for training. In this method, the experiments are first conducted to find the set of inputs that predict the magnitude of earthquakes with highest accuracy. To forecast the earthquakes, the data points are divided into four classes based on their magnitude values. The data points in the catalog which had magnitude in the range of 3.0 to 4.0 were considered to lie in one class, and similarly for 4.0 to 5.0, 5.0 to 6.0 and 6.0 to 7.0.

Initially, inputs given to multilayer perceptron were time difference (in minutes) between subsequent events, latitude, longitude and depth. Later on, as suggested in Reyes et al., (2013), seven new inputs are included in addition to the inputs used in the above experiment. These inputs are:

- $b_i = \frac{\log(e)}{\sum_{j=0}^{49} M_{i-j-3}}$
- $\Delta b_{1i} = b_i - b_{i-4}$
- $\Delta b_{2i} = b_{i-4} - b_{i-8}$
- $\Delta b_{2i} = b_{i-8} - b_{i-12}$
- $\Delta b_{2i} = b_{i-12} - b_{i-16}$
- $\Delta b_{2i} = b_{i-16} - b_{i-20}$
- $\max\{M_t\}$  where  $t \in [-7, 0]$  and  $M_t$  is the maximum magnitude on  $t^{\text{th}}$  day.

**2.1.2 Hyper-parameters** After obtaining the set of inputs, experiments were conducted to find an optimal set of hyper-parameters such as the number of layers, the number of neurons in each layer and various other attributes like loss function and activation functions of different layers.

## 2.2 Nowcasting using recurrent neural networks

After achieving the optimal set of inputs and hyper-parameters, another model is formulated to analyze the time of occurrence

of earthquake using nowcasting techniques. Nowcasting is a surrogate method to find the current progression of occurrence of large earthquakes using the count of small events that occur between two large earthquakes. The definition of large magnitude earthquakes changes throughout the study (Rundle et al., 2016). We use different threshold magnitude in different experiments to define large earthquake events. For instance, at the beginning, we considered a threshold magnitude of 5.0 for large event. Later, the magnitude threshold was changed to 6.0 to carry out the experiment. It may be noted that the homogeneity in magnitudes in the nowcasting approach is not an important issue (Rundle et al., 2016).

## 3. DATASET AND SOFTWARE

For this study, the earthquake data is obtained from the global public seismic catalogs, such as USGS National Earthquake Information Centre (NEIC), Advanced National Seismic System (ANSS) and International Seismological Centre (ISC). The data in catalog consists of parameters including latitude, longitude, time, date, depth, magnitude, azimuthal gap, horizontal distance from the epicenter to the nearest station, the root-mean-square (RMS) travel time residual, in seconds, using all weights. The RMS provides a measure of the fit of the observed arrival times to the predicted arrival times for this location. All the earthquake occurrences from 1975 to 2018 were selected for the experiment.

The models in the present study are trained using the "TensorFlow" library of Python. Another open-source high-level neural network library "Keras" is also used as a wrapper for TensorFlow to conduct the desired experiments.

## 4. RESULTS

As suggested in the methodology section, different experiments were conducted as narrated below.

### 4.1 Experiment 1 - Comparison of different set of inputs

The information of Himalayan region between the longitudes  $74^{\circ}\text{E}$  to  $84^{\circ}\text{E}$  and latitudes between  $25^{\circ}\text{N}$  to  $35^{\circ}\text{N}$  was used and earthquakes with magnitude greater than 3.0 were considered. The selected region and the earthquakes are highlighted in Figure 1. The first input set comprises of

- time difference in minutes with the previous earthquake
- latitude
- longitude
- depth.

The other set of inputs are

- $b_i$
- $\Delta b_{1i}$
- $\Delta b_{2i}$
- $\Delta b_{2i}$
- $\Delta b_{2i}$
- $\Delta b_{2i}$
- $\max\{M_t\}$  where  $t \in [-7, 0]$  and  $M_t$  is the maximum magnitude on  $t^{\text{th}}$  day

along with latitude, longitude, depth and elapsed time since the last large earthquake.

For both the set of inputs, we observed almost similar accuracy in the earthquake magnitude class prediction. The reason for such

behaviour could be because of the fact that new set of inputs was derived from the previous inputs, such as magnitude, latitude and longitude. The neural network therefore may not have benefited much from these new inputs, as it may have already captured these relationships on its own. The loss versus epoch graph for the first set of inputs is illustrated in Figure 2. Results for other set of inputs are observed to be almost same.

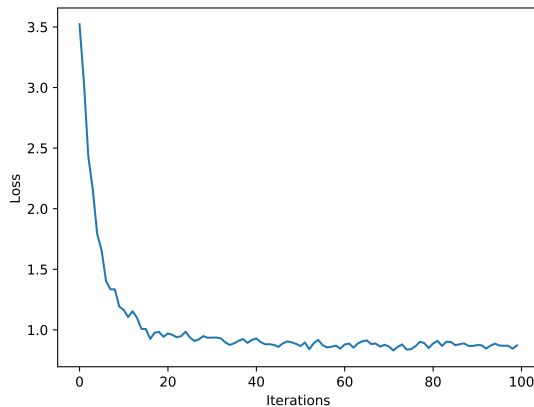


Figure 2. Loss versus iterations (epoch) graph for the Experiment 1

#### 4.2 Experiment 2 - Comparison of different set of hyper-parameters

Hyper-parameters such as number of epochs, number of hidden layers, cost function, optimizer, and the learning rate were varied for the set of inputs as given in Experiment 1 (Nair and Hinton, 2010; Maas et al., 2013). A table highlighting the differences in results from different set of input parameters is illustrated in Table 1. While the number of epochs was set to be 100, the loss function is considered as “softmax loss” with a learning rate of 0.01 for all the trials (Maas et al., 2013). Total number of cells in each layer was set to 256.

	Hidden Layers	Activation Function	Optimizer	Pre-processing	Accuracy
1	1	ReLu	Gradient-Descent	No	0.7699
2	1	ReLu	Gradient-Descent	Yes	0.7512
3	1	ReLu	Adam	Yes	0.6925
4	1	ReLu	Gradient-Descent	No	0.7699
5	2	ReLu	Gradient-Descent	No	0.7699
6	3	ReLu	Gradient-Descent	No	0.7699
7	1	Leaky ReLu	Gradient-Descent	No	0.7558

Table 1. Effect of hyper-parameters on accuracy

For more details on the above activation functions, readers may consult Nair and Hinton, (2010) and Maas et al., (2013). In addition, from Table 1, it is evident that the results are more or less consistent with the changes in hyper-parameters (Row 2–Row 7). No major improvement in accuracy is observed.

#### 4.3 Experiment 3 - Nowcasting using recurrent neural networks

Nowcasting is a method to indirectly determine the progression of large earthquakes in a defined geographic region using the count of small events that occur between subsequent large events (Rundle et al., 2016). To implement, the LSTM neural network architecture is used. The LSTM is a special kind of network which has the ability to use the previous input to predict the next value (Wang et al., 2017). Since the LSTM network is well suited for tasks involving time series analysis analogous to the data used for nowcasting, it may serve as a potential tool for this study (Lakshmi and Tiwari, 2006; Moustra et al., 2011; Wang et al., 2017). In comparison to earthquake forecasting that looks forward in time, nowcasting analyzes the present state of earthquake system by evaluating the cumulative probability for the current number of small earthquakes since the last large event in a selected region (Rundle et al., 2016). To compute the cumulative distribution of interevent counts of small events, the number of small earthquakes is tabulated to develop probability distribution function or cumulative distribution function in a defined geographical region. The potential candidate probability distributions for the underlying seismicity statistics are considered to be exponential, gamma, and Weibull distribution (Pasari, 2018). The exponential distribution is time-independent, whereas the others are time-dependent probability models. Since neural networks can handle complex relationships in a moderately simpler manner, it may be used to compute nowcast values, rather than finding a forecast value.

In this experiment, input to the neural network is in different form than the inputs used in earlier experiments. Input here is a sequence of number of small magnitude interevent earthquakes between two large magnitude earthquakes. In this experiment, earthquake recordings with magnitude greater than 5.0 were termed as large magnitude earthquakes and earthquakes between magnitudes 3.0 and 5.0, were considered as small magnitude earthquakes. The look-back hyper-parameter for the LSTM was set to 5 and the time-step was set to 1. The data points are summarized in Figure 3. The recurrent neural network used here is the LSTM network as illustrated in Figure 4.

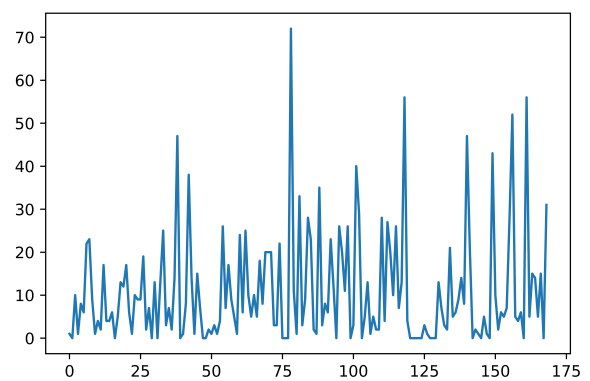


Figure 3. Time series of number of small earthquakes between subsequent large earthquakes in Experiment 3

#### 4.4 Experiment 4 - Nowcasting of large earthquakes

Since the previous dataset was relatively of small size with only 113 entries for training and 52 points for testing, a larger dataset

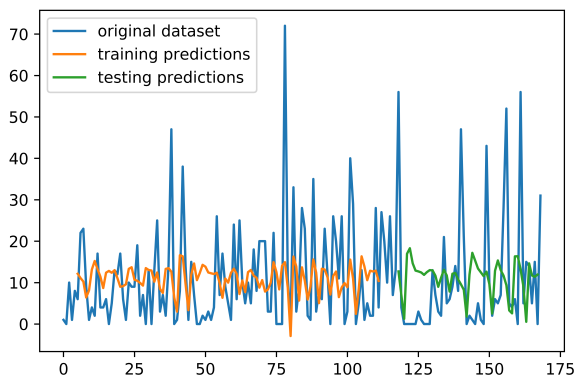


Figure 4. Time series of number of small earthquakes between two subsequent earthquakes and corresponding training and validation predictions in Experiment 3

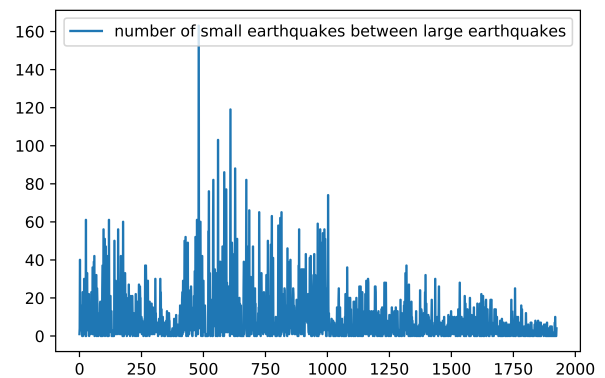


Figure 6. Time series of number of small earthquakes between two subsequent large earthquakes in Experiment 4

was used for the above purpose. This dataset with 1289 training points and 636 testing points was obtained from a larger region of latitudes between  $20^{\circ}\text{N}$  to  $40^{\circ}\text{N}$  and longitudes from  $70^{\circ}\text{E}$  to  $105^{\circ}\text{E}$  (Figure 5). The graph in Figure 6 depicts the data points in this dataset. The associated predictions for the LSTM network are given in Figure 7. After training the LSTM network, we found an RMS error of 6.16 on the test data.

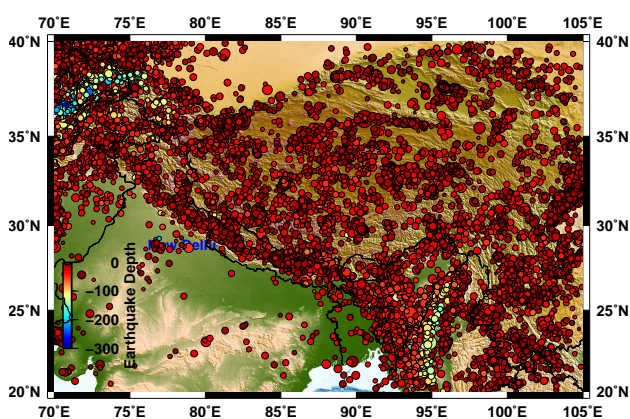


Figure 5. Selected region for Experiment 4

## 5. CONCLUSIONS

The first part of this study presents a simple multilayer perceptron based neural network model which can predict the magnitude of earthquake and its date and location. The whole magnitude range was divided into four classes: 3.0 to 4.0, 4.0 to 5.0, 5.0 to 6.0, and 6.0 to 7.0, respectively.

Besides the basic neural network discussed above, another network was trained that took into consideration the factor of the Gutenberg-Richter frequency-magnitude  $b$ -value and also the difference in  $b$ -values as mentioned earlier. Since majority of earthquake events in the present dataset corresponds to class 2 (earthquakes of having magnitude between 4.0 and 5.0), there is a significant level of class imbalance present in the dataset. The

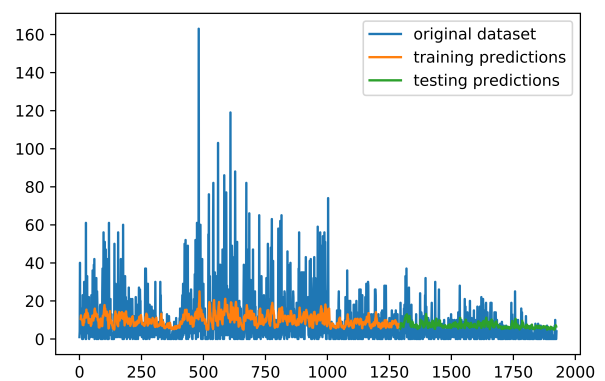


Figure 7. Time series of number of small earthquakes between two subsequent large earthquakes and corresponding training and validation predictions in Experiment 4

limitation of instrumental capability to capture lower magnitude events may have created this issue.

The USGS catalog has information about large number of earthquakes, but it lacks in the number of useful parameters that may play a role in predicting the earthquakes. Almost all the datasets corresponding to earthquake recordings have four common attributes for a record: latitude, longitude, focal-depth and magnitude. The presence of less number of attributes in the dataset and the class imbalance could also be the reason of failure of neural network in the classification task. These limitations led to use of LSTM networks and the nowcasting method as an alternative to the previous experiment. As the data used in nowcasting can be treated like a time series, the LSTM networks are one of the best options to carry out our analysis. A number of experiments were conducted and evaluated for different set of input hyper-parameters. However, since the input data for nowcasting is very unevenly distributed, the LSTM technique in the present study could not produce desirable outputs. Different other techniques therefore should be considered to compare analytical results.

To summarize, the present study, for the first time, has attempted to use the LSTM architecture to lay the foundation for estimating

different hyper-parameters to obtain the least uncertainty in nowcasting results. Preliminary results of this study are discussed. Although the method is capable to model physical complex dynamic threshold systems in an efficient manner, further efforts are required in this direction for a stringent conclusion.

### ACKNOWLEDGEMENTS

The generic mapping tool is used to prepare some figures. This work is partially supported by the ACRG grant (No: PLN/AD/2016-17/01) provided to the second author [S.P.] by BITS Pilani.

### REFERENCES

- Alarifi A., Alarifi N., Humidan S., 2012. Earthquakes magnitude prediction using artificial neural network in northern Red Sea area. *Journal of King Saud University*, 24(4), pp. 301-313.
- Amar E., Khattab T., Zada F., 2014. Intelligent earthquake prediction system based on neural network. *International Journal of Civil and Environmental Engineering*, 8(12), pp. 874-878.
- Asencio-Cortés G., Martínez-Álvarez F., Troncoso A., Morales-Esteban A., 2017. Medium-large earthquake magnitude prediction in Tokyo with artificial neural networks. *Neural Computing and Application*, 28(5), pp. 1043-1055.
- Florido E., Aznarte J., Esteaban A., 2016. Earthquake magnitude prediction based on artificial neural networks: A survey. *Croatian Operational Research Review*, 7(2), pp. 159-169.
- Kurach K. and Pawlowski K., 2016. Predicting dangerous seismic activity with recurrent neural networks. In: *Proceedings of the Federated Conference on Computer Science and Information Systems*, pp. 239-243.
- Lakshmi S. and Tiwari R., 2006. Model dissection from earthquake time series: A comparative analysis using modern non-linear forecasting and artificial neural network approaches. *Computers and Geosciences*, 35(2), pp. 191-204.
- Madahizadeh R. and Allamehzadeh M., 2009. Prediction of aftershocks distribution using artificial neural networks and its application on the May 12, 2008 Sichuan earthquake. *Journal of Seismology and Earthquake Engineering*, 11(3), pp. 112-120.
- Maas A., Hannun A., Ng A., 2013. Rectifier nonlinearities improve neural network acoustic models. In *proceedings of the 30<sup>th</sup> International Conference on Machine Learning*, Atlanta, Georgia, USA, 30(1).
- Moustra M., Avraamides M., Christodoulou C., 2011. Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals. *Expert Systems with Applications*, 38(12), pp. 15032-15039.
- Nair, V. and Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27<sup>th</sup> International Conference on Machine Learning*, pp. 807-814.
- Narayanakumar, S. and Raja, K., 2016. A BP artificial neural network model for earthquake magnitude prediction in Himalayas, India. *Circuits and Systems*, 7(11), pp. 3456-3468.
- Niksarlioglu S. and Kulahci F., 2013. An Artificial neural network model for earthquake prediction and relations between environmental parameters and earthquakes. *International Journal of Geological and Environmental Engineering*, 7(2), pp. 87-90.
- Pasari S., 2018. Stochastic modeling of earthquake interoccurrence times in northwest Himalaya and adjoining regions. *Geomatics, Natural Hazards and Risk*, 9(1), pp. 568-588.
- Perol T., Gharbi M., Denolle M., 2017. Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2), e1700578.
- Reyes J., Morales-Esteban A., Martinez-Ivarez F., 2013. Neural networks to predict earthquakes in Chile. *Applied Soft Computing*, 13(2), pp. 1314-1328.
- Rundle, B., Turcotte D., Donnellan A., Grant Ludwig L., Luginbuhl M., and Gong G. 2016. Nowcasting earthquakes. *Earth and Space Science*, 3, pp. 480-486.
- Sriram A., Rahanamayan S., Bourennani F., 2013. Artificial neural networks for earthquake anomaly detection. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 18(5), pp. 701-702.
- Wang Q., Guo Y., Yu L., Pan L., 2017. Earthquake prediction based on spatio-temporal data mining: an LSTM network approach. *Transactions on Emerging Topics in Computing*, pp. 2168-6750.
- Zamani A. and Sorbi M., 2013. Application of neural network and ANFIS model for earthquake occurrence in Iran. *Earth Science Informatics*, 6(2), pp. 71-85.