*17th World Conference on Earthquake Engineering, 17WCEE*

*Sendai, Japan - September 13th to 18th 2020*

# FORECASTING OF EARTHQUAKE USING RECURRENT NEURAL NETWORK

S. Nayan(1), B. Prakke(2), A. Bhalkikar(3), P. Ramancharla(4)

(1) *MS Scholar, Earthquake Engineering Research Centre, International Institute of Information Technology, Hyderabad, India, swapnil.nayan@research.iiit.ac.in*
(2) *PhD Scholar, Earthquake Engineering Research Centre, International Institute of Information Technology, Hyderabad, India, bharat.prakke@research.iiit.ac.in*
(3) *PhD Scholar, Earthquake Engineering Research Centre, International Institute of Information Technology, Hyderabad, India, aniket.bhalkikar@research.iiit.ac.in*
(2) *Professor and Head, Earthquake Engineering Research Centre, International Institute of Information Technology, Hyderabad, India, ramancharla@iiit.ac.in*

## Abstract

Estimating the time, location and the size of an earthquake event using neural networks is an emerging technology. Present study aims to predict the near future earthquake events in each of the regions partitioned as clusters of epi-central locations of earthquakes related to Indian subcontinent. K-means clustering algorithm is applied to develop a catalogue of earthquake epicenters. Two approaches are considered for the purpose of forecasting. In the first approach regions obtained as clusters are considered as the area sources. The distribution of time series of (temporal distribution) events are divided into five time periods viz., 15 days, 30 days, 60 days, 90 days and 180 days. The accuracy of forecasting a magnitude for the next time period is studied using application of Recurrent Neural Network In the second approach the seismic data of each cluster is further grouped into three categories of magnitude viz., minor (≤ 4 Mw), moderate (between 4 Mw and 6 Mw) and strong (> 6 Mw). Accuracy of forecasting for each of the group belonging to the clusters is computed considering the same five time periods. The epicentral locations represented as latitude and longitude is the inputs for recurrent neural network. The results show that accuracy of second approach is significantly more than first approach. Finally, an attempt has been made to compare the accuracy level of Recurrent Neural Network method with the accuracy level of Gutenberg Richter law in predicting the earthquakes for future 60 days. It has been observed that Gutenberg-Richter law poorly forecasts near future earthquake.

*Keywords: Recurrent Neural Network; Earthquake forecasting, Clustering.*

## 1. Introduction

Forecasting a catastrophic disaster should answer three important questions i.e., where (location), when (time) and how big (size) to prevent loss of life and to minimize infrastructure damage. Of all the natural disasters earthquake prediction and forecasting has remained an unsolved problem. Studying the physical phenomena and establishing relationship among the related parameters that causes earthquakes is a method used for forecasting in the past [1]. Another approach is the statistical study of the past earthquake data which is used to forecast the occurrence of next earthquake event [1]. The Gutenberg-Richter recurrence law depicts the relationship between magnitude and frequency of occurrence of earthquakes. However, Omori-Utsu law, which is time dependent short-term model, establishes relationship between decrease of aftershock activities with time, is the most often used statistical model for forecasting. In past, innumerable contributions have been made by the researchers from various disciplines like seismology, geology, geophysics, engineering, mathematics, computer science etc., because of the complexity of this problem which demands the efforts of interdisciplinary research. New models of computing based on neural networks have evolved, which extracts the relevant features from the input data and perform pattern recognition tasks by learning from examples [2]. This new model has shown promising progress by adding a new dimension to solve research problems in various fields such as seismology, geophysics, earthquake engineering, signal processing, neuroscience etc.

## 2. Data Set

The reliability of records of historical seismic data has improved from the past 200 years. The earthquake data representing the seismic activities of Indian Subcontinent are collected from the open catalogues of United States Geological Survey (USGS), Indian Meteorological Department (IMD), National Centers for Environmental Information (NCEI) and International Seismological Centre (ISC). The dataset consists of 8683 Seismic events in Indian Subcontinent for a time duration of 200 years i.e., from 1$^{st}$ January 1818 to 31$^{st}$ December 2017. The parameters considered in the data set are location, moment magnitudes and the time of occurrence of earthquake. The distribution of data based on magnitude is shown in the Table 1.

Table 1 – Number of events based on Magnitude

| Number of Events | Magnitude |
|---|---|
| 86 | <3 $M_w$ |
| 2303 | 3 – 4 $M_w$ |
| 4110 | 4 – 5 $M_w$ |
| 1854 | 5 – 6 $M_w$ |
| 285 | 6 – 7 $M_w$ |
| 42 | 7 – 8 $M_w$ |
| 5 | > 8 $M_w$ |

It can be observed that about 95% of the recorded earthquakes have a magnitude ranging between 3 to 6. This distribution poses a difficulty in predicting very low magnitude earthquakes as well as severe earthquakes. To combat this difficulty the data set is segregated into three categories namely 'Minor Magnitude', 'Moderate Magnitude' and 'Strong Magnitude'. The range of Magnitudes in each category is shown in Table 2.

Table 2 – Earthquake magnitude categories

| Group | Range |
|---|---|
| Minor Magnitude | ≤ 4 $M_w$ |
| Moderate Magnitude | 4 $M_w$ – 6 $M_w$ |
| Strong Magnitude | > 6 $M_w$ |


 

Table 3 shows the number of earthquake events in each of the minor, moderate and strong magnitude categories. It can be observed that strong magnitude events are still considerably low.

Table3 – Number of events in each Category

| Number of Events | Category |
|---|---|
| 2389 | Minor Magnitude |
| 5964 | Moderate Magnitude |
| 332 | Strong Magnitude |

## 3. Methodology

Broadly the seismicity of India can be divided into four groups namely, the Himalayan region, Andaman Nicobar, Kutch region and the Peninsular India [3]. The Himalayan region is tectonically most active with inter-plate collision whereas, Kutch region and Peninsular India experiences occasional moderate earthquakes due to intra-plate seismic activities. A catalogue of earthquake epicenters is partitioned using application of K-Means algorithm [4]. K-means algorithm partitions data into number of clusters with *Euclidean* distance as the goodness measure. Choosing the optimum number of clusters is the foremost task. To achieve this, the K-means algorithm is repeated number of times with different number of clusters and the Within Cluster Sum of Squares (WCSS) is calculated for each repetition. A graph of WCSS vs. Number of clusters Fig.1 depicts the optimum number of clusters to be chosen. Fig.1 shows that as number of clusters increases, the value of WCSS decreases. The number of clusters at a point in Fig.1 after which there is no significant change in the value of WCSS is considered. K-Means ++ is used for choosing initial centroids. Therefore, optimum number of clusters obtained from the graph is five.

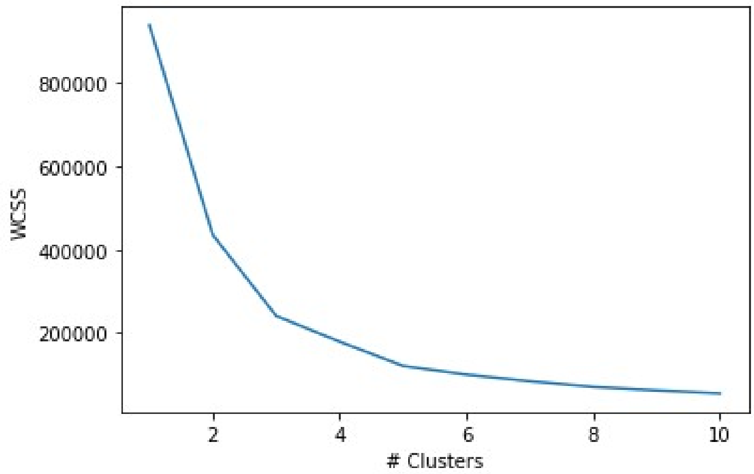


Fig. 1 – Number of Clusters vs. Within Clusters Sum of Squares (WCSS)

The spatial distribution of the earthquake events in each of the five Clusters is shown in Fig.2. Cluster 1 (i.e., *Blue circles*) represents the peninsular India and the Kutch region. Clusters 2 (i.e., *Magenta circles*) and Cluster 5 (i.e., *Cyan circles*) represent the Himalayan region, Cluster 3 (i.e., *Red circles*) represents the Andaman Nicobar regions and Cluster 4 (i.e., *Green circles*) represents North Eastern part of India. Table 4 shows the total number of events in each cluster along with the latitude and longitude range values.



 

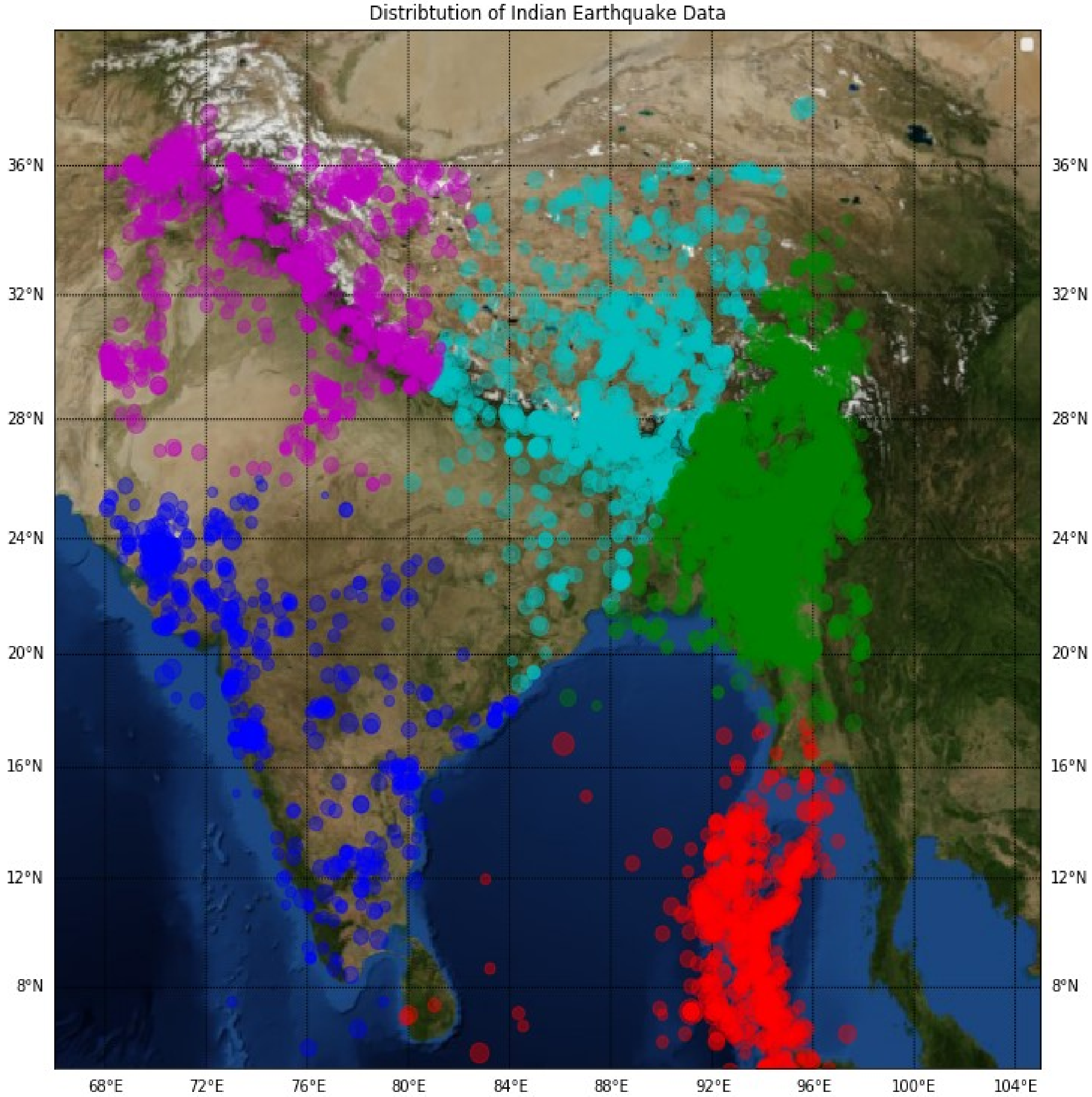


Fig. 2 – Distribution of Earthquake Data into Clusters

Table 4 – Details of Each Cluster

| Clusters | Number of Events | Latitude Range | Longitude Range |
|---|---|---|---|
| Cluster 1 | 617 | $05.80^0$ N to $26.03^0$ N | $68.07^0$ E to $84.03^0$ E |
| Cluster 2 | 1638 | $19.00^0$ N to $37.80^0$ N | $80.10^0$ E to $95.70^0$ E |
| Cluster 3 | 845 | $04.01^0$ N to $17.47^0$ N | $79.05^0$ E to $97.30^0$ E |
| Cluster 4 | 4414 | $17.62^0$ N to $34.38^0$ N | $86.28^0$ E to $97.99^0$ E |
| Cluster 5 | 1172 | $25.85^0$ N to $37.60^0$ N | $68.01^0$ E to $84.42^0$ E |



 

3.1 Recurrent Neural Network

Recurrent Neural Network (RNN) is a type of Artificial Neural Network (ANN) which can exhibit temporal dynamic behavior. While ANN algorithms provide solution to problems based on the designed architecture and gaining experience from learning. But the dynamic nature of changing of data with time is special feature of RNN algorithm over ANN algorithms. This makes RNN an ideal tool to solve time series relate problems.

RNN algorithm use long short-term memory (LSTM) in order to connect previous information with the current task. This gives RNN algorithm the ability to decide what amount of previous information that is required to carry out the task at hand. In case of Earthquake prediction, an earthquake of magnitude 4 may come every month but an earthquake of magnitude 7 may only come once every 10 years. This means for prediction of 4 magnitude earthquake, data of previous few months is enough to predict the next earthquake while the prediction of 7 magnitude earthquakes may require hundreds of years of data.

In the present study, the data has been fitted with RNN by considering varying amounts of previous data to make the future prediction. Each of the 15, 30, 60, 90 and 180 days prediction was made by considering previous 6, 12, 24, 30, 40, 60 and 80 day time steps.

For making predictions, RNN always considers the timely pattern of previous data. For example, to make prediction of earthquake in next 30 days, while considering 12 timesteps of previous 30-day period to make future prediction means a total of 360 days of previous data is considered to make the next prediction.

3.2 Methods to create sub regions

Two different approaches to apply Recurrent Neural Network (RNN) have been explored in this paper. In each of the two approaches, the data is divided into training data which is 90% of total data and testing data which is 10% of the total data. Training data is used to fit RNN while the testing data which is previously unknown by the RNN is used to make predictions. The attempt of prediction made covers three important aspects as discussed in section 1 i.e., location, size and time. Each cluster represents the location, magnitude represents size. For predicting time, a duration of 15-days, 30-days, 60-days, 90-days and 180-days is considered. These time durations only are considered as it gave more accurate prediction result on the subset data of time durations of 7-days, 15-days, 20-days, 30-days, 40-days, 50-days, 60-days, 70-days, 80-days, 90-days, 120-days, 150-days, 180-days, 210-days, 240-days, 270-days, 300-days, 330-days, 360-days and 400-days.

3.2.1 Approach 1

In the first approach, RNN has been applied to each of the five clusters obtained from K-Mean clustering. In this approach algorithm predicts only maximum magnitude of earthquake, which is likely to occur in future 15 days, 30 days, 60 days, 90 days and 180 days.

To achieve this, the raw data is divided into intervals of 15, 30, 60, 90 and 180 days for prediction in respective days. Taking example of 15-day prediction, the maximum magnitude of earthquake that occurs in every 15 days interval is considered for training the Recurrent Neural Network. If an earthquake does not occur in one of the 15 days duration, then the maximum magnitude of earthquake for that period is considered as zero. To compare the Magnitudes obtained from predictions against the real values, Mean Absolute Error (MAE) and Accuracy is calculated. Mean absolute error can be calculated as Eq. (1).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - x| \tag{1}$$

Here n is the total number of observations, $x_i$ is the actual value and x is the predicted value. The accuracy for the prediction is obtained by converting each observation from actual dataset and predicted values into the three groups of Minor Magnitude, Moderate Magnitude and Strong Magnitude earthquakes as described in Table 2. Hence an earthquake of magnitude 4.3 will belong into group 2 i.e., Moderate Magnitude. If the groups of both the actual value and predicted value are the same, then it is considered as correct prediction.

$$Accuracy = \frac{Total\ Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \tag{2}$$

3.2.2 Approach 2

In the second approach, the data from each cluster has been further classified into Minor Magnitude, Moderate Magnitude, and Strong Magnitude defined in Table 2. This allows us to not only predict the earthquake with the maximum magnitude that is likely to occur in the next time instance but also earthquakes of Minor and Moderate Magnitudes. This creates total fifteen groups. Each of this group is further divided into time intervals of 15, 30, 60, 90 and 180 days in a similar way as in Approach 1. RNN is then fitted to each of the 15 groups to obtain the predictions.

Mean absolute error and Accuracy is calculated to check the error in predicted Magnitude and actual Magnitude. However, the accuracy in this approach is calculated by converting each occurrence of earthquake in a cluster as 1 and the non-occurrence as 0. Similarly, the predicted value is converted to 1 or 0 based on whether earthquake occurred in the next time step or not respectively. This is done since the data is already segregated based on the magnitudes and hence only occurrence or non-occurrence is considered.

## 4. Results and Observations

The prediction accuracy in percentage and mean absolute error obtained for all the time intervals in each cluster using Approach 1 is shown in Table 4 and Table 5 respectively.

Table 4 – Prediction Accuracy (%) of Five Clusters in Approach 1

| **Clusters** | **Time Period** | | | | |
|---|---|---|---|---|---|
| | **15 Days** | **30 Days** | **60 Days** | **90 Days** | **180 Days** |
| Cluster 1 | 84.04 | 76.06 | 64.79 | 56.34 | 45.07 |
| Cluster 2 | 56.25 | 62.92 | 75.83 | 82.50 | 22.50 |
| Cluster 3 | 51.21 | 33.33 | 31.73 | 31.88 | 17.14 |
| Cluster 4 | 60.77 | 73.48 | 82.42 | 75.41 | 61.29 |
| Cluster 5 | 43.01 | 40.32 | 49.46 | 66.13 | 58.06 |

From Table 4, it can be observed that the accuracy of prediction is the lowest for Cluster 3 i.e. Andaman and Nicobar Islands region. This is also evident from the high MAE for Cluster 3 as seen in Table 5. From Table 5 it can also be observed that Cluster 5 i.e. North and North-Western Regions of India has MAE that is comparable to the MAEs of Cluster 1, Cluster 2 and Cluster 4, yet the accuracy is low. This can be attributed to the fact that accuracy calculation applies the rules of segregating the data into Minor, Moderate and Strong magnitude earthquakes rigorously.

Similarly, the prediction accuracy in percentage and mean absolute error obtained for all the time intervals in each cluster using Approach 2 is shown in Table 6 and Table 7 respectively.



 

Table 5 – Mean Absolute Error of Five Clusters in Approach 1

| **Clusters** | **Time Period** | | | | |
|---|---|---|---|---|---|
| | **15 Days** | **30 Days** | **60 Days** | **90 Days** | **180 Days** |
| Cluster 1 | 1.69 | 1.78 | 1.96 | 2.47 | 2.81 |
| Cluster 2 | 1.39 | 1.58 | 0.93 | 0.73 | 1.09 |
| Cluster 3 | 2.73 | 2.10 | 2.38 | 2.20 | 2.97 |
| Cluster 4 | 2.62 | 1.04 | 0.92 | 0.75 | 0.93 |
| Cluster 5 | 1.17 | 1.68 | 1.23 | 1.23 | 1.41 |

Table 6 – Prediction Accuracy (%) of Five Clusters in Approach 2

| **Clusters** | **Magnitude Class** | **Time Period** | | | | |
|---|---|---|---|---|---|---|
| | | **15 Days** | **30 Days** | **60 Days** | **90 Days** | **180 Days** |
| Cluster 1 | Minor | 100 | 100 | 100 | 100 | 100 |
| | Moderate | 85.98 | 76.15 | 61.67 | 48.75 | 37.5 |
| | Strong | 98.92 | 98.08 | 96.15 | 94.96 | 92.86 |
| Cluster 2 | Minor | 100 | 63.75 | 100 | 100 | 100 |
| | Moderate | 58.66 | 61.90 | 79.31 | 89.61 | NA |
| | Strong | 97.11 | 94.23 | 89.42 | 84.43 | 74.29 |
| Cluster 3 | Minor | 100 | 100 | 100 | 100 | 100 |
| | Moderate | 52.17 | 32.85 | 83.65 | 15.94 | 94.29 |
| | Strong | 95.65 | 91.30 | 84.06 | 76.09 | 65.22 |
| Cluster 4 | Minor | 100 | 100 | 100 | 100 | NA |
| | Moderate | 67.13 | 80.11 | 94.51 | 95.08 | 100 |
| | Strong | 94.90 | 90.48 | 82.43 | 75.51 | 56 |
| Cluster 5 | Minor | 100 | 100 | 100 | NA | 23.33 |
| | Moderate | 44.62 | 23.12 | 92.47 | 95.16 | 100 |
| | Strong | 92.82 | 85.71 | 77.36 | 71.43 | 55.56 |

From Table 6, it can be observed that every Earthquakes of Minor Magnitude earthquake was correctly predicted using Approach 2 and hence the accuracy of prediction is 100%. Whereas an accuracy of greater than 85% was observed in predicting Moderate Magnitude earthquakes in each of the Clusters in at least one of the time periods considered i.e. 15, 30, 60, 90 and 180 days. Approach 2 was even able to predict the Strong Magnitude earthquake with a high accuracy of more than 90% for each of the five clusters for at least one of the time periods considered. Moreover, Table 7 shows that MAE is comparatively very high in case of predicting minor magnitude earthquake and is very low for strong magnitude earthquakes. The MAEs of 15 Day period has least value compared to other time periods along all the magnitude classes. The MAE value shows gradual increase in trend from 15-day period to 180-day period. This indicates that there is a better prediction for 15-day time period in case of using RNN.



 

Table 7 – Mean Absolute Error of Five Clusters in Approach 2

| Clusters | Magnitude Class | Time Period: 15 Days | 30 Days | 60 Days | 90 Days | 180 Days |
|---|---|---|---|---|---|---|
| Cluster 1 | Minor | 0.86 | 1.38 | 2.06 | 2.42 | 2.95 |
| | Moderate | 0.67 | 1.15 | 1.85 | 2.49 | 3.09 |
| | Strong | 0.08 | 0.13 | 0.27 | 0.34 | 0.50 |
| Cluster 2 | Minor | 1.28 | 1.38 | 2.74 | 3.20 | 3.21 |
| | Moderate | 1.94 | 2.06 | 1.35 | 1.31 | 0.78 |
| | Strong | 0.20 | 0.41 | 0.75 | 1.10 | 1.83 |
| Cluster 3 | Minor | 0.82 | 1.27 | 0.86 | 1.42 | 1.25 |
| | Moderate | 2.39 | 3.41 | 2.37 | 1.83 | 0.97 |
| | Strong | 0.30 | 0.60 | 1.10 | 1.65 | 2.43 |
| Cluster 4 | Minor | 1.18 | 1.04 | 2.82 | 1.42 | 1.48 |
| | Moderate | 1.77 | 1.27 | 0.93 | 0.79 | 0.37 |
| | Strong | 0.35 | 0.65 | 1.20 | 1.68 | 3 |
| Cluster 5 | Minor | 1.88 | 1.53 | 0.82 | 0.64 | 0.73 |
| | Moderate | 2.76 | 1.53 | 0.73 | 2.06 | 0.92 |
| | Strong | 0.48 | 0.96 | 1.52 | 1.97 | 3.04 |

Further to check the accuracy level of RNN, an attempt has been made to compare the results with the accuracy level of Gutenberg-Richter recurrence law. For this purpose, the earthquake data from Cluster 4 is used and Gutenberg-Richter recurrence law is applied to predict earthquakes in future 60 days. The result is shown in Table 8.

Table 8 – Prediction Accuracy (%) of Gutenberg Richter Relation for future 60 Days

| Magnitude Class | Accuracy (%) | Remarks |
|---|---|---|
| Minor | 4.17 | *Always No* |
| Moderate | 100 | *Always Yes* |
| Strong | 91.67 | *Always No* |

It is observed that for the near future event, the relation results in zero prediction of events (*Always No)* for minor and strong magnitude earthquakes while for moderate magnitude earthquakes a single event (*Always Yes*) is being predicted for every 60-day duration over five years (i.e., from 2013 to 2017). This can also be seen by the fact that accuracy of prediction from Gutenberg-Richter recurrence law for moderate and strong magnitude earthquakes surpasses that accuracy from RNN as there are only two strong magnitude earthquakes in the five years and only one moderate magnitude earthquake for every 60 day time duration for the five years.



 

## 5. Conclusion

Efforts made to forecast earthquakes using neural networks have been discussed in the paper. Application of Recurrent Neural Network to the grouped data of both the approaches discussed in the paper have resulted in mixed success. The reason being the compelling difference in the level of accuracy achieved in both the approaches. The results of Approach 2 are significantly more accurate than the results of Approach 1. The reason being that the data in Approach 2 is properly segregated and is range bounded by minor, medium and strong category earthquake data. This is attributed to the fact that the Recurrent Neural Network (RNN) fits more accurately to the range bounded data rather than haphazard data. A comparison of RNN and Gutenberg-Richter recurrence law shows that Gutenberg-Richter law poorly forecasts near future earthquake. Developing a specific neural network architecture or a model with a strong basis based on features from the direction of multiple disciplines is an area to be explored.

## 6. References

[1] Leonid F. K., George V. C., John O. R. Jr., Bernard E. (2000): *Progress in Developing a Forecasting System, Preferential Precursors, and Monitoring Network*. Gas Migration Events Proceeding Earthquakes 2000, pp 209-219.

[2] B. Yegnanarayana (1998): Artificial Neural Network. *Book*, Prentice Hall India Learning Private Limited.

[3] Pankaj A., Manish S. (2015): Earthquake Resistant Design of Structures. *Book*, Prentice Hall India Learning Private Limited.

[4] Ashif P., Hojjat A. (2008): Recent Efforts in Earthquake Prediction (1990-2007). *Natural Hazards*, **9** (2), 70-80.



 