Performance Evaluation of LSTM Model for Earthquake Prediction

[1]Ravi Sharma, [2]A.J. Singh

[1]Research Scholar, [2]Professor
Department of Computer Science,
Himachal Pradesh University, Shimla, India

Abstract : : In these days Artificial Intelligence plays a very important role in our life. Deep Learning is one of the topics in which researchers and academician taking an interest. Computer Vision, Natural Language Processing are some of the applications where deep learning performance is very remarkable. There is a different type of models in Deep Learning Autoencoder, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), etc. CNN models are best suited for Image classification problem and RNN model is for Time Series Data. Earthquake prediction is a challenging task because there is no state-of-the-art method to solve this problem. In this research, we use Long Short-Term Memory (LSTM) a type of RNN model to predict earthquake and we will find significant improvement from the existing approach.

Keywords: Deep Learning, CNN, RNN, LSTM, Earthquake Prediction.

I. INTRODUCTION

Earthquake prediction is an unsolvable and unpredictable natural disaster. Most of the time Earthquake occur without warning and do not allow much time for people to react. Therefore, they can cause serious injuries and loss of life and destroy buildings and infrastructure, leading to great economy loss. Predicting an earthquake is a challenging task because there is no state-of-the-art method to solve this problem, that’s why traditional mathematical and statistical methods cannot analyze well in this process. Deep learning methods such as RNNs are able to find the nonlinear correlations in data. LSTM is a type of RNNs method that is used to analyze time series data to make a prediction. In this proposed work we were using LSTM for earthquake prediction.

Deep Learning is a subfield of Machine Learning. It uses multi-layered artificial neural networks to deliver state-of-the-art accuracy in tasks such as object detection, speech recognition, language translation, and others. Deep Learning is different from machine learning in a way how features extracted. In traditional machine learning, feature extraction is done manually whereas in deep learning it is selected automatically. In Deep Learning there is an input layer, hidden layer, and output layer. There is only one input layer and output layer but there can be as many numbers of the hidden layer that is why a word “Deep’ is used in this type of neural network model. There are different types of model in Deep learning like Autoencoder, Deep Belief Network, Convolutional Neural Network, and Recurrent Neural Network. In this study, we use CNN and RNN model for earthquake prediction.

Convolutional Neural Network is a type of artificial neural network consists of convolutional layers, pooling layers, and fully connected layers. In this convolutional layer is used for feature extraction, polling layer is used for feature mapping. In the field of image classification, CNN performance is very good

RNN is a kind of artificial neural network apart from having the structure of the feed forward neural network, there exists a directed cycle in Recurrent Neural Network. This type of structure allows information to be circulated in the network so that the output of each time is not only related to the input at present but related to the input at previous timestamps. RNN is used in many applications like Speech Recognition, Robot Control, Music Generation, Time series prediction

Long Short-Term Memory is a type of Recurrent Neural Network. The main drawback of the RNN model is that it cannot hold the data for a long duration, due to this limitation we cannot predict the upcoming event accurately, to overcome this shortcoming of RNN model we use LSTM. They are used to avoid long term dependency problem All types of recurrent neural networks have a chain of repeating modules of the neural network. In standard RNNs, this repeating module will have a simple structure, such as a single tanh layer. LSTMs also have chain-like structure, but the repeating module has a different structure [1]. A typical LSTM network is a set of different memory blocks called cells. These memory blocks used for remembering things and data transfer to this memory is done through three major mechanisms, called gates. The function of each gate is discussed below

- **Input Gate:** This gate is responsible for the addition of new data to the cell state.
- **Forget Gate:** This gate used for removing information from the cell state. The data that is no longer required to understand things or the information that is of less importance is removed.
- **Output Gate:** The output gate selects useful information from the current cell state and showing it as an output

Earthquakes are one of the most destructive yet unavoidable natural disasters around the world, with a great physical and economic impact in the population. Due nonlinear correlations among earthquake occurrences and also their occurrence depend on a multitude of variables that in most cases are yet unidentified. Therefore, having a better understanding of the occurrence of each seismic event, and estimating the seismic hazard risk, would represent an invaluable tool for improving earthquake prediction. Predicting an earthquake is a challenging task because there is no state-of-the-art method to solve this problem, that’s why traditional mathematical and statistical methods cannot analyze well in this process

II. RELATED WORK

In this section, we introduce some of the related works that have been done in the area of earthquake prediction

Hayakawa [4] and Jiang [9] predict earthquakes base on precursor signals studies. They take the electromagnetic signals as the precursor of significant earthquakes. Hayakawa et al. [5] study the abnormal behavior of animals about 10 days before earthquakes for earthquake prediction. it is very difficult to draw conclusions on theses precursor signals due to very limited data. Besides, these precursor signals alone usually cannot lead to satisfactory prediction results.

Qianlong Wang et al. [7] uses the spatio temporal data for earthquake prediction. They use the Deep Learning approach LSTM. In this study, they applied their research on one-dimensional data and two-dimensional data and they found that giving two-dimensional data as an input we will get a more accurate result in earth quake prediction. Francisco Plaza et al. [8] use a deep neural network to asses seismic hazard in Chile. For this research, they use historic earthquake data from 2012 to 2018 to predict future earthquake events. They use a deep feedforward artificial neural networks (DFANNs) and long short-term memory (LSTM) for ground intensity function estimation. Both deep learning approaches are compared to find the best model.

Mustafa Al Ibrahim et al. [6] made an earthquake warning system. They use a deep neural network to recognize and predict an earthquake. For this research, they take three deep learning models: one dimensional CNN, two dimensional CNN, and an RNN model. In earthquake recognition, all three models give a good result, but in earthquake prediction, none of the models are able to perform very well. They found that in earthquake prediction there is a little bit of improvement from the random guessing. Because there is no predefined method to predict an earthquake that is why we cannot predict an earthquake accurately.

III. DATASETS AND FEATURES

In this study, we use dataset available on GitHub [2]. For this dataset author [2] uses Continuous seismic waveform data that is available through many publicly funded seismic station networks. Because there is no proven method for identifying earthquake precursors, there is little information about how to create an optimal dataset for the earthquake prediction problem. They also decided to use a balanced dataset of positive and negative samples based on the conclusions of Buda et al. [3] who found that training a CNN under sampled negative samples yielded the best result.

The dataset is comprised of 1) for each seismic station, the query the catalog for all earthquakes above a minimum(greater than 3) magnitude 2) for each earthquake, they estimate the time of arrival of the seismic wave 3) they then download the seismic waveform for a specified period of time around the arrival time of the earthquake; and 4) then they download a random seismic waveform to create a balanced dataset of positive and negative samples.. These samples are retrieved from 46 stations from the Berkeley Geysers Network. In this study, we use 100 seconds of seismic data to predict an earthquake.

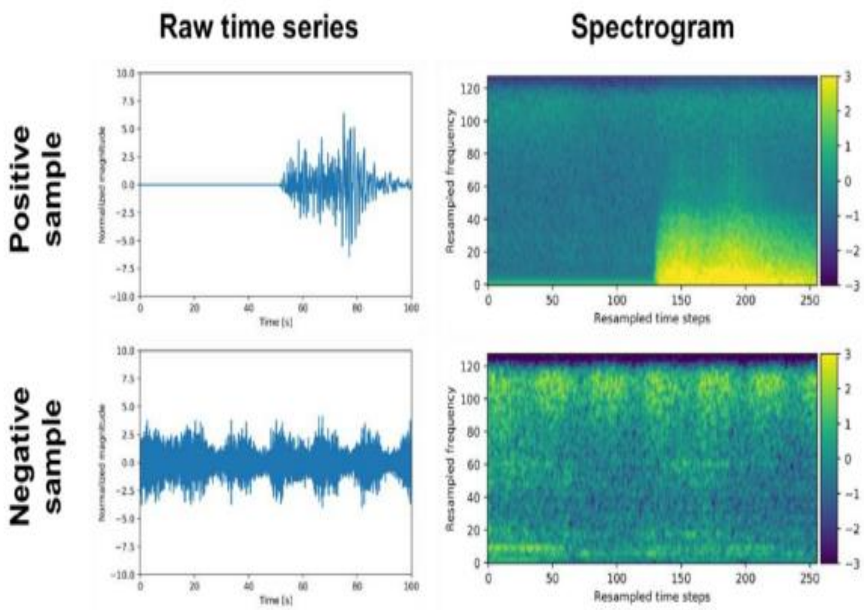

Fig. 1: Raw Data and Spectrogram. A positive sample represent there is earthquake during the period and negative sample represent there is no earthquake [6]

IV. METHODS

An RNN neural network architectures were developed using the Keras library for Earthquake prediction problems. The first two RNN model consists of a 1 convolution layer followed by two LSTM layers, other two RNN model consists of a 1 convolution layer followed by three LSTM layers and also uses the spectrogram of the seismic waveform as the input variable. In all models, binary cross entropy is used for the loss.

The dataset is first to split into training/test sets using 90% and 10%, and then 10% of the training set is used as a validation set. Hyperparameters are tuned by computing accuracy on the validation set. For example, the optimum values of dropout rate and the number of epochs were determined such that the model does not overfit the training data.

V. RESULTS

In this study, we tested the same dataset on four different LSTM model that have a different number of LSTM layers and dropout rate. For each model, there is a confusion matrix. On the basis of confusion matrix accuracy is calculated.

Model I

There are 1 CNN layer and 2 LSTM layers are used with the dropout rate of 0.3, 0.2, 0.2 respectively

No. of Epochs	FN	FP	TN	TP	Accuracy
25	0.28947368	0.13157895	0.44736842	0.13157895	0.421052631579
50	0.36842105	0.26315789	0.21052632	0.15789474	0.526315789474
75	0.18421053	0.23684211	0.28947368	0.28947368	0.473684210526
100	0.44736842	0.18421053	0.34210526	0.02631579	0.473684210526

Table 1: Confusion matrix and Accuracy for Model I

Model II

There are 1 CNN layer and 2 LSTM layers are used with the dropout rate of 0.3, 0.4, 0.4 respectively

No. of Epochs	FN	FP	TN	TP	Accuracy
25	0.31578947	0.10526316	0.36842105	0.21052632	0.526315789474
50	0.21052632	0.21052632	0.18421053	**0.39473684**	**0.605263157895**
75	0.05263158	0.36842105	0.13157895	0.44736842	0.5
100	0.28947368	0.13157895	0.34210526	0.23684211	0.526315789474

Table 2: Confusion matrix and Accuracy for Model II

Model III

There are 1 CNN layer and 3 LSTM layers are used with the dropout rate of 0.3, 0.4, 0.4, 0.4 respectively

No. of Epochs	FN	FP	TN	TP	Accuracy
25	0.26315789	0.15789474	0.39473684	0.18421053	0.447368421053
50	0.23684211	0.18421053	0.31578947	0.26315789	0.5
75	0.23684211	0.18421053	0.28947368	0.28947368	0.526315789474
100	0.26315789	0.15789474	0.31578947	0.26315789	0.526315789474

Table 3: Confusion matrix and Accuracy for Model III

Model IV

There are 1 CNN layer and 3 LSTM layers are used with the dropout rate of 0.3, 0.3, 0.3, 0.3 respectively

No. of Epochs	FN	FP	TN	TP	Accuracy
25	0.44736842	0.18421053	0.28947368	0.07894737	0.526315789474
50	0.44736842	0.18421053	0.13157895	0.23684211	0.684210526316
75	0.34210526	0.28947368	0.28947368	0.07894737	0.421052631579
100	0.34210526	0.28947368	0.18421053	0.18421053	0.526315789474

Table 4: Confusion matrix and Accuracy for Model IV

From the above results it is clear that Model II has the best result when the no. of epochs is 50. We will get accuracy of 61 percent with true positive (TP) of 40 percent. In model IV we were able to reach the accuracy of 68 percent but true positive value is very less as compared to the Model II. All four model have a good result when the number of epochs is 50 it means accuracy degrade when we train our model more and less than 50 epochs. Figure 2. shows comparison of graphical representation for accuracy of different model.

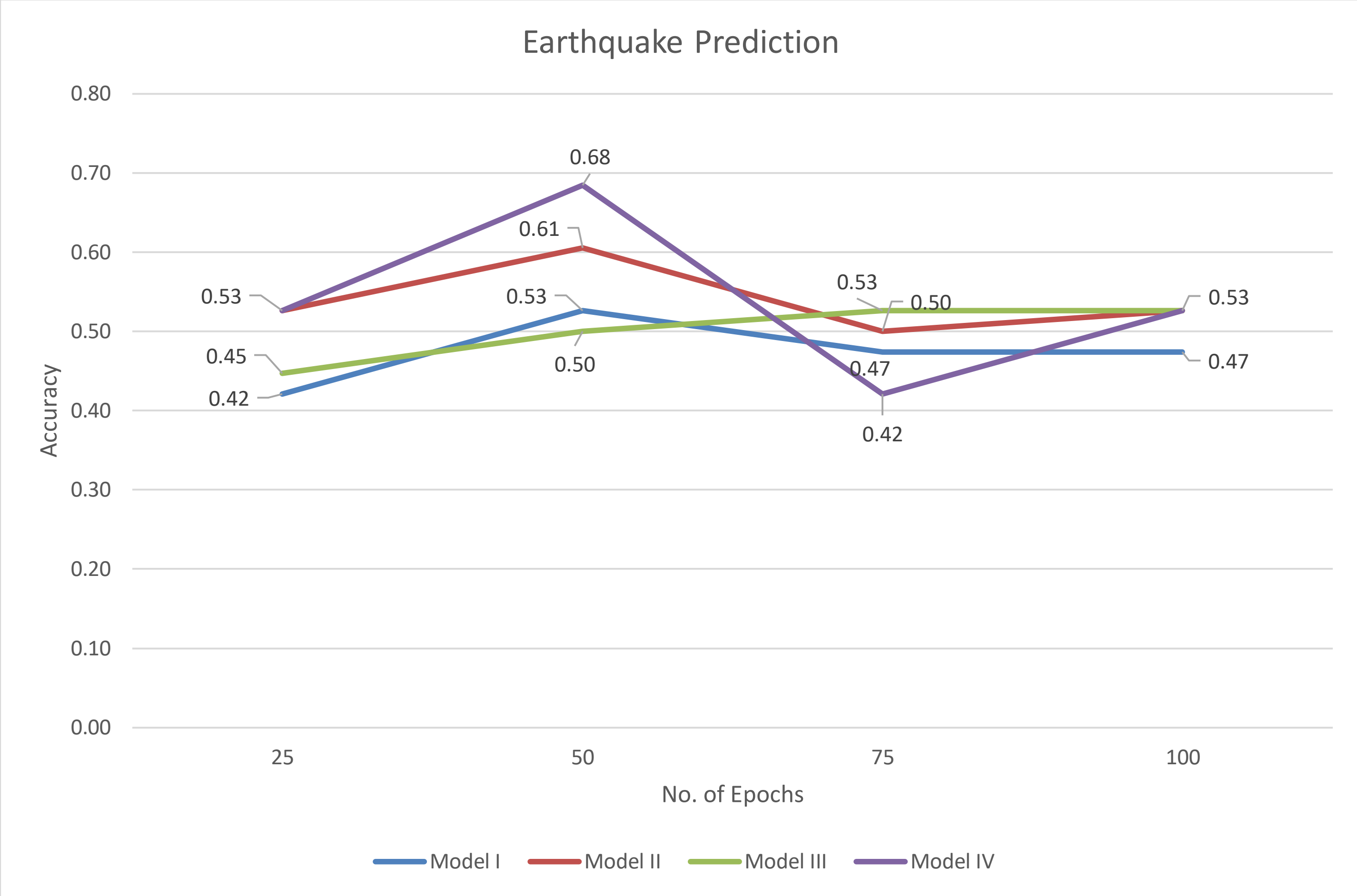

Fig. 2: Comparison of Different Model Accuracy

VI. CONCLUSION AND FUTURE SCOPE

In this study we use Long Short-Term Memory (LSTM) a deep neural network approach for earthquake prediction. We compare four model to predict the accuracy and we found that in our research model II has the best result. Although our results show improvement over previous result, but there is only little improvement from the existence approach, possible issue for this problem is, there is no state-of-the-art method to predict earthquake, and the network architecture are not suitable for the setup of our problem.

Predicting earthquake before they occur is still a very challenging task. Future work may include use of bigger dataset, to find parameters that will help in earthquake prediction and finding the relationship between warning time and prediction accuracy.

REFERENCES

[1] https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[2] https://github.com/MosGeo/TerraeMotus/tree/master/Datasets

[3] Buda, Mateusz, Atsuto Maki, and Maciej A. Mazurowski. "A systematic study of the class imbalance problem in convolutional neural networks." *Neural Networks* 106 (2018): 249-259.

[4] M. Hayakawa. Earthquake prediction with electromagnetic phenomena. In THE IRAGO CONFERENCE 2015: 360 Degree Outlook on Critical Scientific and Technological Challenges for a Sustainable Society, volume 1709, page 020002. AIP Publishing, 2016.

[5] M. Hayakawa, H. Yamauchi, N. Ohtani, M. Ohta, S. Tosa, T. Asano, A. Schekotov, J. Izutsu, S. M. Potirakis, and K. Eftaxias. On the precursory abnormal animal behavior and electromagnetic effects for the kobe earthquake (m˜ 6) on april 12, 2013. Open Journal of Earthquake Research, 5(03):165, 2016.

[6] Mustafa Al Ibrahim, Jihoon Park and Noah Athens. “Earthquake warning system: Detecting earthquake precursor signals using deep neural networks.” Stanford University. 2018

[7] Wang, Qianlong, et al. "Earthquake prediction based on spatio-temporal data mining: an LSTM network approach." IEEE Transactions on Emerging Topics in Computing (2017).

[8] Plaza, Francisco, Rodrigo Salas, and Orietta Nicolis. "Assessing Seismic Hazard in Chile Using Deep Neural Networks." Natural Hazards. IntechOpen, 2019.

[9] M. Jiang. Easily magnetic anomalies earthquake prediction. In MATEC Web of Conferences, volume 63, page 01020. EDP Sciences, 2016