Natural Disaster Prediction using LSTM

Rohit Ramesh Lambade
Electronics and Telecommunication
Mumbai, India
rohit.lambade@spit.ac.in

Ravish Rakesh Shah
Electronics and Telecommunication
Mumbai, India
ravish.shah@spit.ac.in

Rahul Rajaram Mane
Electronics and Telecommunication
Mumbai, India
rahul.mane@spit.ac.in

Dr.Sujata Kulkarni
Electronics and Telecommunication
Associate Professor
Mumbai, India
sujata_kulkarni@spit.ac.in

***Abstract*—We live in an era of technology, people still face abrupt changes in the climate which leads to natural calamities like flood and earthquake[1]. Such uncertainties that would strike people, which will lead to the economic downfall in their present and future. This can be updated to the people present in that region just prior 30 days of the occurrence of the calamity. Artificial Intelligence has the capability to solve such issues in its unique approaches. The solution developed by us enables people to view the percentage of flood[2] and earthquake striking the particular region where he or she lives. We are deploying a website for the users where they would be getting warning of getting struck by an earthquake or flood. This is completely handled by the cloud. The future which is going to be predicted is completely dependent on the past behavior of those natural calamities which are considering each and every factor affecting the environmental disturbance. The deployed software will automatically get updated after every month with its predictions for that month. The website will be monitored by the team members to answer the questions of the user. The prediction might be sometimes inaccurate, but has the capability to give its highest percent of accuracy. The major aim of the website is to safeguard the lives of people from getting drowned by such uncertain calamities, so that they can shift to the place where those calamities does not strike.**

***Index Terms*—Machine Learning, Prediction, Natural Calamity, Website.**

I. INTRODUCTION

A natural disaster is a consequence when a natural hazard affects humans and/or the built environment. Human vulnerability, and lack of appropriate emergency management, leads to financial, environmental, or human impact. The resulting loss depends on the capacity of the population to support or resist the disaster: their resilience.

Various natural disaster like earthquakes,volcanic eruptions, landslides, cyclones and floods kill thousands of people and destroy billions of dollars of property each year.

According to previous study India is one of the country which is affected by both earthquake[3][4] and flood disaster from the year 1994 to 2013, includes 6,873 natural disasters all over the world,which claimed 1.35 million lives or almost 68,000 lives on average each year, and per Annam affected 218 million people on an average during this 20-year period. at present people get updated about disaster via news or website like ”Windy.com” which tells then how much rainfall is occurred. in which area Water-Logged is there and in which direction wind is flowing etc.to overcome the following issue we need to predict the future occurrence of disaster and with proper use of machine learning algorithm we can achieve that The main aim of the project is to predict the natural calamities like flood and earthquakes and display the data being predicted regarding the natural calamity on the website so that it safeguards the life of people nearby.

II. LITERATURE SURVEY

Artificial Neural Network (ANN) is widely used for prediction, clustering, and pattern recognition based on past and present training data in which accuracy is obtained to a huge extent. The prediction is limited to few factors like wavelet data, rainfall data and underground plates data for flood and earthquake respectively. There is no website of prediction available as such for these two natural calamities at one place. As live updates can be provided by social media, news channels, etc. But the website for prediction is not yet been used widely. The current predictions that are available are approx 83% and below. Web scraping was only used for extracting limited data from the website, whereas many more information can be extracted from the site. The updates of the prediction currently are slow, regardless of better accuracy. If some extra efficient parameters are introduced in the system accuracy and correct prediction can be obtained to a huge extend.

III. WORKING

When a person in a particular division in Maharashtra feels he or she is facing abrupt changes in nature. They can check the website to get updated regarding the prediction of the natural calamity going to strike the particular month. Once the user is on the website he or she has to click on the region and then will get the predicted information of natural calamity.

The decision displayed is pre-loaded by the machine itself depending on the past few environmental information. The user will be able to see the percentage of possibility of flood[5] and earthquake is their region and it the percentage exceeds the predefined limit, then that would beep an alarm message on the website for that particular region. Once visited website, the user will also be exposed to the recent disasters striking near by their region, which would also be the alarm to get safeguarded.

IV. ALGORITHM

Sensors for creating dataset. The amount of rainfall and magnitude of earthquake in the divisions of Maharashtra is measured using rainfall sensors and seismographs respectively. These sensors update the data to the website of government of Maharashtra (Regional Meteorological Centre, Mumbai) everyday.

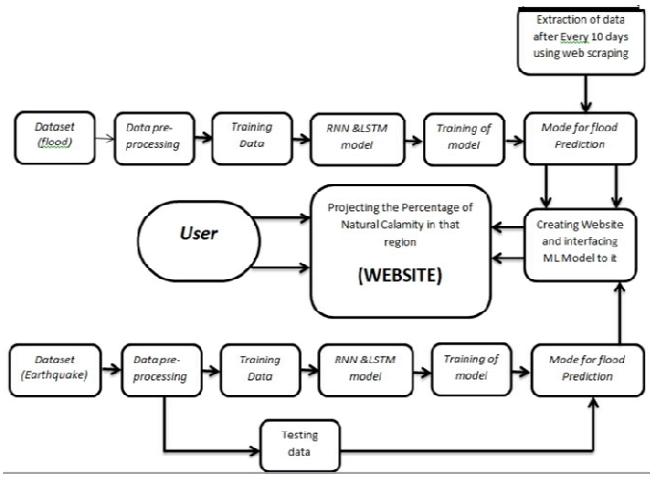

Fig. 1: Detailed Block Diagram

Getting data update from gov website by web scraping. The dataset used for testing the model and predicting natural calamity is updated on a daily basis. The dataset is updated using web scraping. Dataset used for testing is of past 30 months and after every month the testing dataset is updated and data from testing dataset is send to training dataset. For better accuracy the model is created after every 3 months and using training data.

Preprocessing of Dataset : The dataset[6] sent by government consists is for entire India having parameters as mentioned :- SUBDIVISION: ANDAMAN NICOBAR ISLANDS; YEAR: 1901; JAN: 49.2; FEB: 87.1;...; DEC: 33.6 for Flood prediction and Year: 2017; Month: 2; Day: 27; TimeUTC: 04:21:42; Mag: 4.6; Lat: 27.35; Lon: 85.96; Depth km: 25.7; Region: PAKISTAN; IRIS ID: 10004460; Timestamp: 1488169302 for Earthquake prediction.

As the project focuses more on regions of Maharashtra, the dataset is preprocessed with data of different divisions such as Konkan and Goa; Madhya Maharashtra; Marathvada; Vidharbha; Chhattisgarh and regions like Andaman and Nicobar Islands and Lakshadweep are eliminated from the dataset as shown in Fig 1. As LSTM needs input in columns, the row of months is converted in columns and various divisions of Maharashtra are specified in different columns. For earthquake LSTM needs magnitude, Depth km, Lat, Lon, Timestamp and regions for output shown in Fig 2.

	YEAR/MONTH	RAINFALL_KONKAN & GOA	RAINFALL_MADHYA MAHARASHTRA	RAINFALL_MARATHVADA	RAINFALL_VIDHARBA	RAINFALL_CHHATTISGARH
0	JAN1901	5.6	18.8	15.8	36.8	46.9
1	FEB1901	0.1	0.6	3.3	39.9	116.5
2	MAR1901	0.4	7.7	32.1	30.9	27.8
3	APR1901	35.7	36.6	48.5	26.1	5.5
4	MAY1901	19.9	30.4	26.5	7.3	18.4
...	...	...	...	...	...	...
1375	AUG2015	377.3	67.2	112.2	288.9	286.6
1376	SEP2015	240.9	146.6	154.3	167.5	216.9
1377	OCT2015	91.4	48.3	19.3	7.0	17.7
1378	NOV2015	27.3	16.2	4.8	0.0	0.6
1379	DEC2015	0.0	0.1	0.0	0.2	1.5

1380 rows × 6 columns

Fig. 2: Prepossessed dataset for Flood

```
In [30]: df.head()
```

Out[30]:

	Mag	Depth km	Lat	Lon	Timestamp	Region_AFGHANISTAN	Region_AFGHANISTAN-TAJIKISTAN BORD REG.
0	0.253455	-0.914457	-0.642393	0.668328	1.833269	0.0	0.0
1	0.494023	-1.134402	-0.620598	0.689347	1.832979	0.0	0.0
2	0.734591	-1.134402	-0.216660	0.428110	1.829495	0.0	0.0
3	-0.949383	0.181064	0.679850	-0.805955	1.824413	0.0	1.0
4	-0.468248	0.515350	0.650790	-0.929063	1.808172	0.0	0.0

Fig. 3: Prepossessed dataset for Earthquake

Model Creation : For Creation of Model the algorithm used is LSTM, it has long term dependency and preserves the vanishing and exploding gradients of the past cell. Functions used are Sigmoid: Used to range the input data between 0 to 1 and Tangent: Used to range the input data between -0.5 to 0.5.

The basic functioning of LSTM is based on Gates and Cell State. Types of gates :

1. Input gate

$$i_t = \sigma(W_i[h_t - 1, x_t] + b_i)$$

2. Forget gate

$$f_t = \sigma(W_f[h_t - 1, x_t] + b_f)$$

3. Output gate

$$o_t = \sigma(W_o[h_t - 1, x_t] + b_o)$$

where:-

$i_t \rightarrow$ *represents input gate.*

$f_t \rightarrow$ *represents forget gate.*

$o_t \rightarrow$ *represents output gate.*

$\sigma \rightarrow$ *represents sigmoid function.*

$W_x \rightarrow$ *weight for the respective gate*(x) *neurons.*

$h_t - 1 \rightarrow$ *outputofthepreviousLSTMblock*$(attimestamp - 1)$.

$x_t \rightarrow$ *input at current timestamp.*

$b_x \rightarrow$ *biases for the respective gates*(x).

The Forget gate is used to check the data to be kept or deleted from the prior steps. The input gate decides what information is relevant to be added from the current step. The output gate is used to state what the next hidden state should be. While Training [7] the model the number of epochs used are 50 it is adjusted according to time required to train the model and to increase the accuracy of the model. The dataset is segregated in the order of 75:25 as for ratio above and below this the model appends wrong prediction.

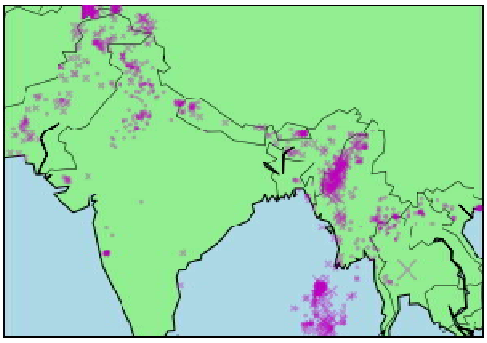

Fig. 4: Training Data representation of Earthquake

Sending data to the Server : The created model is given the dataset of few prior months so that it learns and predicts the output for next immediate month. The generated outputs values and images are send to backend server generated by DJANGO and then those output are processed and displayed on the website. in figure

Deployed Website : Website deployed is much user friendly as it allows the user to explore the recent cities stroked by natural calamity, contact details of important agencies in that region and also to select if one wants to check flood or earthquake. On clicking the particular division of Maharashtra user would get exposed to the predictions of flood and earthquake in the particular region user has selected.

V. RESULT

The output in Fig 4 shows real and predicted rainfall data each peak shows January to December rainfall reading in one year there are total 19 peak are plotted as India has maritime climate. This gives us an accuracy of about 89.4%. We trained the model for last 19 years of rainfall data, Fig 4 shows that is December 2015 rainfall reading and 23.21 is predicted January 2016 rainfall reading. Values of prediction are shown in Fig 5 Here the data below 50 is neglected as that does not affect any environmental hazard.

The output is fig 7 resembles total r2 score is 0.858496044395 r2 scores for individual variables are 'Timestamp': 0.66200000000000003, 'Depth km': 0.34399999999999997, 'Mag': 0.64400000000000002, 'Lat': 0.44700000000000001, 'Lon': 0.60399999999999998

As we can see values in fig 8(a) and fig 8(b) has a lot of variation as during the month September its difficult to predict using ML Algorithms as its variation is due to external factors

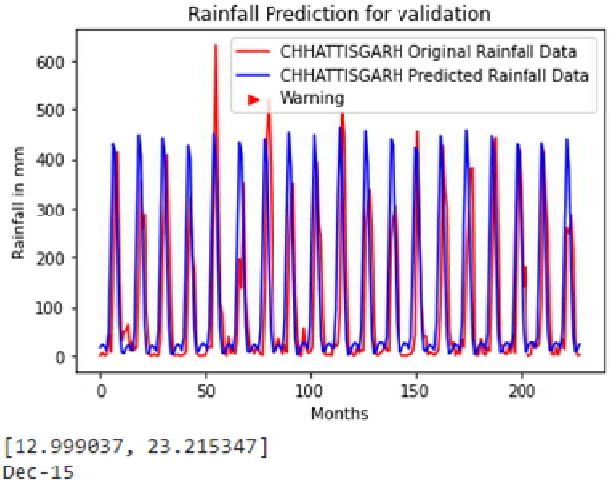

Fig. 5: Graphical representation of Real and predicted Rainfall

[257.6],	[196.45781],
[248.6],	[395.41446],
[286.6],	[397.58136],
[216.9],	[223.68234],
[17.7],	[76.84234],
[0.6],	[20.811031],
[1.5]]	[11.840351],

(a) Real Value (b) Predicted Value

Fig. 6: Reading to resemble the graph of fig 5

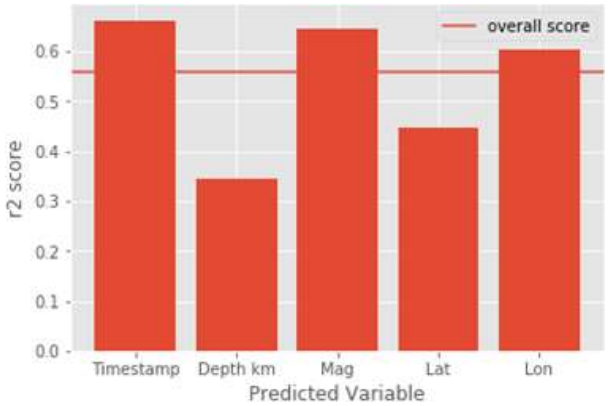

Fig. 7: Earthquake Predicted r2 Scores

(Global Warming, Pollution etc) and it increases or decreases year by year depending on climatic conditions.

VI. CONCLUSION

Nowadays machine learning is getting implemented in every sector for easy computational processing and high accurate outputs. As we know year 2020 is full of disasters having cyclones, corona pandemic, earthquakes, floods and many more hazardous calamities, our project is a small contribution

to the society where we have created a website using machine learning for finding a non-linear output which is the prediction of natural calamities (i.e Earthquakes, floods) which can occur anywhere at any time. The main pillar of our project was the dataset itself. Better quality data would give better accuracy. We have taken dataset from official government site having past data of flood and earthquake from 1905 to 2015 but the main challenge we faced was that having a good quality dataset does not mean 100 percent accurate output as year by year climate changes, global warming and many other external factors are affecting rainfall, temperature, sea-level of earth ‘s surface. We cannot win over climate change but have increase the accuracy of the model by increasing the number of epochs of entire training dataset. Increasing the epoch size, we have easily overcome mean square error (MSE). Thus, the accurate output which we have got using machine learning is deployed on our website which is user-friendly and predicting output (i.e... the occurrence of earthquake or flood at any place and time) one month prior. Thus, our project can be a game changer by saving many people ‘s life by predicting a natural calamity before its occurrence.

VII. FUTURE SCOPE

Accurate prediction of calamities can be a turning point in every human’s life. We have successfully predicted floods and earthquakes occurrences one month prior. Further development in this field could help a lot to the society. Accurate Predicting cyclones, volcanic eruptions, hurricanes, tornadoes, tsunamis would save people ‘s life to a great extent, precautionary measures can be taken to avoid that calamity. Official application for predicting disasters occurring all around the world is the most awaited thing nowadays which humans are waiting for.

VIII. REFERENCES

[1] Dr.P.Vishnu Raja, Dr.K.Sangeetha , S.Sibikrishna, C.Swetha M.Vijayakumar,“Earthquake prediction using Machine Learning using Support vector machine algorithm.” ,International Journal Of Advanced Science and Technology, Vol.29, No.3s, pp. 997-1003, 2020.

[2] Nadia Zehra “Prediction analysis of floods using machine learning algorithms (narx and svm).”,International Journal of Sciences: Basic and Applied Research, Vol. 49, No. 2, pp 24-34,2020

[3] Anmol Gaba, Arnab Jana, Rahul Subramaniam, Yash Agrawal, Merin Meleet , “Analysis and prediction of earthquake impact-a machine learning approach.”,4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), 2019.

[4] Roxane Mallouhy, Chady Abou Jaoude, Christophe Guyeux, Abdallah Makhoul, “Major earthquake event prediction using various machine learning algorithm.”, International Conference on Information and Communication Technologies for Disaster Management (ICT-DM),2019.

[5] Suresh Sankaranarayanan Malavika Prabhakar, Sreesta Satish; Prerna Jain, Anjali Ramprasad; Aiswarya Krishnan ,“Flood prediction based on weather parameters using deep learning.” .Journal of Water and Climate Change,2019.

[6] Awal Rais Sanubari, Purba Daru Kusuma, Casi Setianingsih “Flood modelling and prediction using artificial neural network.” IEEE International Conference on Internet of Things and Intelligence System (IoTaIS), 2018.

[7] Ashif Panakkat, Hojjat Adeli, “ Neural network models for earthquake magnitude prediction using multiple seismicity indicators.”, International Journal of Neural Systems, Vol. 17, No. 1,pp 13–33,2017.

IX. AUTHORS PROFILE

Rohit Ramesh Lambade is currently a final year student in the Electronics and Telecommunications Department of SPIT, Mumbai. His research interests lie in Machine learning, Computer Vision, Data Analytics and Embedded Systems.

Rahul Rajaram Mane is currently a final year student in the Electronics and Telecommunications Department of SPIT, Mumbai. His research interests lie in Machine learning, Computer Vision, Data Analytics and Embedded Systems.

Ravish Rakesh Shah is currently a final year student in the Electronics and Telecommunications Department of SPIT, Mumbai. His research interests lie in Machine learning, Computer Vision, Data Analytics and Embedded Systems.

Dr. Sujata Kulkarni holds PhD in Electronics from RSTMNU Nagpur and has work experience of more than 5 years in the industry. She has been in the teaching profession for over 13 years, of which she has been associated with Bhartiya Vidya Bhavans’ Sardar Patel Institute of Technology for more than 5 years.