INTERNATIONAL JOURNAL OF TECHO-ENGINEERING
ISSN : 2057-5688

ANALYSIS OF BUILDINGS DAMAGE CAUSED BY EARTHQUAKES USING MACHINE LEARNING

Dr.SARANGAM KODATI
Professor, Department of CSE, Teegala Krishna Reddy Engineering College, Hyderabad, Telangana, India.
e-mail: k.sarangam@gmail.com

Vaishnavi Vangala,Raju Kandula, Saisuhas Guttula
Under Graduate Students, Department of CSE, Teegala Krishna Reddy Engineering College, Hyderabad, Telangana, India.
e-mail: vaishnavivangala279@gmail.com,rajukandula095@gmail.com ,guttulasaisuhas@gmail.com

---

ABSTRACT:

A natural disaster is considered to be an earthquake because of the devastation it causes to both living systems and man-made structures, such as homes, bungalows, and buildings, to name a few. Seismometers, which detect vibrations caused by seismic waves entering the earth's crust, are used to measure earthquakes. In this paper, earthquake-related damage was rated according to its severity using a scale from one to five. A previously acquired data set was employed, and a number of characteristics were taken into account in order to predict the damage grade of a particular building, which is associated with a 33 entirely unique identifying String.The prediction was made using an analysis of current device learning classifier methods. Datasets for earthquakes are sourced from Kaggle. Logistic Regression, Naive Bayes Classifier, Random Wooded Area Classifier, and okay-Nearest Friends were the system analysis techniques employed in this work. The most appropriate set of regulations was taken into consideration after an evaluation of a set of rigid criteria. Logistic Regression, Naive Bayes Classifier, Random Forest Classifier, and K-

Nearest Neighbors were the machine learning techniques employed in this study. The best algorithms were considered after a set of attributes were evaluated.

Keywords: K-Nearest Neighbour, Random Forest Classifier, Earthquake Datasets.

1. INTRODUCTION

An earthquake is nothing more than the Earth's crust moving. Unexpected cracks create elastic energy, which is saved in rocks that have undergone significant strain. Strength generated during an earthquake is being saved for a long time and will release in a matter of minutes or seconds. Greater elastic energy is stored as a result of rock cracks, which increases the likelihood of an earthquake. The seismic waves, which are created by an earthquake, are nothing more than elastic waves. Seismic waves are low frequency waves that are released during an earthquake and cause significant loss of human life.It causes severe injury to a wide range of civil engineering shapes. The Gorkha earthquake on April 25, 2015, is the sole inspiration for our challenge. It is also known as the "Gorkha earthquake." Nepal was destroyed by the earthquake, a natural tragedy. No less people have died as a result of the earthquake in Nepal. Nearly 9000 individuals lost their lives and 22,000 others had injuries as a result of the earthquake in 2015. Because of such natural disasters, countless countless Nepalese now lack a place to live. Due to damage from an earthquake, world heritage sites in the Kathmandu valley and some other regions are damaged. Using specific reports, we can categorise the earthquake's strength and damage cost. On this investigation, artificial intelligence is being used. The prediction process keeps going hard and fast, reducing the relative diversity of earthquakes based on the time distribution of part earthquakes. The area's daily seismic strength has decreased relative to the number of earthquakes that occur there. The damage caused by an earthquake is represented by grades. Grades show damage to the building that was caused by the earthquake. There are three levels of harm: Grade 1, Grade 2, and Grade 3 represent, respectively, mild, medium, and complete destruction [2]. The 2015 Nepal earthquakes in April and May resulted in over 9000 fatalities and almost 20000 injuries. The resulting Destruction became out of control. After the disaster, around 2 million people were left without a place to live. On April 25, an earthquake of magnitude 7.8Mw was followed by an earthquake of significance 7.5, which caused significant damage to highways, bridges, colleges, public health facilities, water and electricity networks, and other structures near people's homes.

There is currently no standard approach for predicting earthquake damage. Furthermore, there is still disagreement among scientists as to whether a solution to this issue is even attainable. However, the quick development of machine learning techniques and their successful application to a variety of issues suggest that these technologies could aid in the discovery of hidden patterns and the generation of precise forecasts. The methods listed above dealt with predicting the damage caused by earthquakes and other natural disasters. However, there aren't many methods that work similarly to anticipate an earthquake or its characteristics, like its magnitude, timing, etc.A map reduce model was used to find the locations with the most tremors in order to anticipate earthquakes using data analytics. A back propagation neural network and artificial neural networks were also used to complete the task. The latter was employed to calculate an earthquake's magnitude, whilst the former addressed its effects. Finally, weighted factor coefficients were utilised to calculate the likelihood of an earthquake.

ISSN : 2057-5688

2. Motivation

To estimate the extent of damage done to buildings after an earthquake, it is possible to use machine learning. It can aid in the distinction between safe and unsafe structures, predict damage-prone locations, and prevent deaths and injuries brought on by an earthquake's aftershock while also facilitating effective rescue operations. This is accomplished by categorising these buildings according to a degree of damage severity based on a number of features, including their age, foundation, number of floors, and material composition. Then, ward-by-ward in a district, the number of families and the likely casualties are considered. This enables the proportionate distribution of relief forces by ward and their prioritising according to the severity of the damage. These types of models could save as many lives as possible as early as possible and prove to be a useful and affordable solution. It can be further enhanced by include the

distribution of goods like food, clothing, medical care, and money in accordance with the number of fatalities among people and the degree of structural damage. This study can help architects, engineers, and city planners by extrapolating and predicting the kinds of structures that are most likely to sustain earthquake damage. It is possible to fortify structures with characteristics similar to those of more damaged structures. The research on earthquake prediction can be used in conjunction with the visualisation and classification algorithms to deliver humanitarian relief in advance so that buildings can be strengthened to sustain much less damage.

3. PROPOSED SYSTEM

Based on the F1 scores determined for each of the four algorithms previously discussed in this study, this analysis demonstrates that

ISSN : 2057-5688

the Random Forest Classifier method has the highest accuracy in forecasting the damage caused by earthquakes. It has been noted that K-Nearest Neighbors is the second most used method for predicting earthquake damage. The research finds that reinforced concrete is the material best suited to preventing damage to structures during an earthquake after analysing the materials that can do so. It is commonly known that earthquakes trigger electromagnetic pulses that induce tremors beneath the Earth's crust. Reinforced concrete adequately shields these electromagnetic pulses.Due to the low tensile strength of reinforced concrete, steel bars that are implanted in the concrete are used. Because of this, reinforced concrete has an incredible capacity to withstand natural disasters like earthquakes. This fact explains why reinforced concrete is used extensively in structures with earthquake damage grade 1 and barely at all in buildings with earthquake damage grade 5. The uses of this work can be expanded to forecast earthquake damage in regions where it is possible to get a comparable and pertinent dataset.

4. IMPLEMENTATION OF DAMAGE TO BUILDINGS CAUSEDBY EARTHQUAKES

Integration of software with libraries, databases, and other programmes may be necessary (s). The integration of software with external world elements takes place during this stage of the SDLC. Installing the software on user computers is the term for this. Software occasionally requires post-installation settings on the user's end. During implementation, software is tested for portability and adaptability, and integration-related problems are resolved. Web applications, which run on multiple computers and communicate over a network or server, are by their very nature distributed apps. Because a web browser is so simple to use as a user client, web applications may specifically be accessed using one. Web applications' success among businesses is largely due to their ability to be updated and maintained without the need to deploy and install software on potentially thousands of client PCs. Web apps are used for a variety of things, including web mail, online shopping, message boards, blogs, and online banking. Millions of individuals can access and use a single web application. Similar to desktop programmes, web applications are composed of numerous components and frequently include little programmes, some of which have user interfaces and others of which don't even need a graphical user interface (GUI). Additionally, web applications usually need a second markup or scripting language, such HTML, CSS, or a programming language like JavaScript. Additionally, a lot of applications solely employ the Java programming language, which is perfect due to its adaptability. A web application might be as basic as a website that displays the time and date or as complicated as a collection of pages where you can research and reserve the best flights, lodgings, and rental cars for your upcoming vacation. The Java EE platform includes many of the Java Platform, Standard Edition (Java SE) classes and packages in addition to the Java technologies you'll employ to build web applications. A container, or web server, that can recognise and run the classes you develop, must be placed on the server for many of these technologies to function. You can use the tools described in this article for the development and testing of these technologies, but before you deploy, make sure that the server has Java server software installed to run web applications based on Java technologies. Consult the server administrator if you don't have access to this data.

DATA EXPLORATION

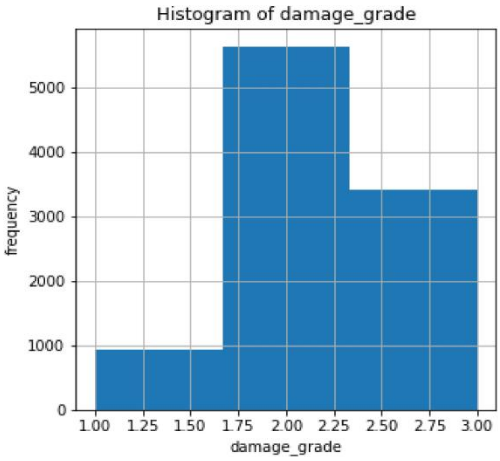


Fig 1: Histogram of Damage Grade

4.1 Random Forest Classifier:

It makes predictions entirely based on the votes cast in each of the decision bushes. We employ 750 estimators to determine the vast diversity of timber in the forested area. We can get greater accuracy by using a larger range of estimators, but doing so makes computation more difficult because of the larger data sets and the lengthy processing time required for output delivery. The criterion helps in determining the quality of each split in statistics in bushes and comes in two varieties: Gini or entropy.Since it does require us to obtain output results, we have employed the Gini impurity in this version. Because this setting makes it easier for us to determine the size of each node in all of the choice bushes, we set the min samples leaf rate to be 5. We chose a boundary of 42 for the Random state. If we provide the same academic records and all other parameters remain constant, this parameter enables us to provide the same results.

The RF is a learning technique that uses a collection of tree topologies. The RF makes use of bagging and random feature selection, two potent machine learning approaches (Breiman, 2001).Using a boot-strap sample of the training data, each tree is individually built in bagging, and the mean value of the trees' outputs is utilised to make predictions (Breiman, 1996). A modified variation of bagging is the RF. When constructing a tree, RF randomly chooses a subset of predictor variables to be split at each node rather than using all predictor variables as in DT. The following are the key steps of the RF algorithm:

1. Generate$n_t$bootstrapsamplesfromthetrainingdataset;
2. GenerateaDTfromeach bootstrapsamplebyselectingthebestsplitamong thedataset;
3. Predicttheoutputofanewdatasetbyaveragingtheaggregateofpredictionsofn_tDTs.

TABLE-1Statistical Features For Different Number Of Tress For Random Forest Algorithms

Nb of trees	Prediction	MAE	RMSE
1	74.82	0.3108	0.4313
2	74.82	0.3237	0.4379
3	76.97	0.312	0.4161
4	74.1	0.3123	0.4233
5	72.66	0.3373	0.4417
10	69.06	0.3313	0.4375

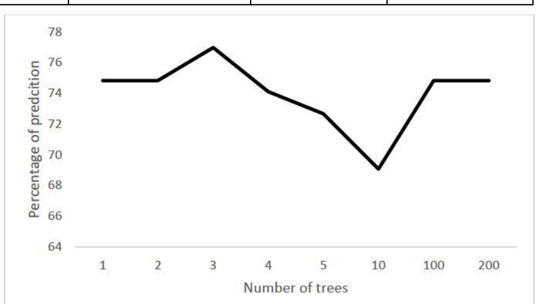


Fig2.PercentageofpredictionfordifferentnumberoftreesusingRandomForestAlgorithm

4.2 K Nearest Neighbour Algorithm

K closest Neighbour is a straightforward algorithm that sorts fresh information or cases based on similarities to the stored examples. The k-nearest neighbor's algorithm, sometimes referred to as KNN or k-NN, is a supervised learning classifier that employs proximity to create classifications or predictions about the grouping of a single data point. Although it can be applied to classification or regression issues, it is commonly employed as a classification algorithm because it relies on the idea that comparable points can be discovered close to one another.

The KNN is a non-parametric method for classifying data into groups. For a positive inte-

Number of Neighbors	Prediction	MAE	RMSE
1	69.06	0.3105	0.5545
2	64.74	0.3064	0.4958
3	75.53	0.3002	0.4565
4	74.10	0.3079	0.4541
5	74.82	0.3067	0.4503
10	74.82	0.3088	0.4527

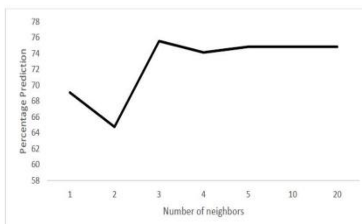
ger*k*,andanobservation*x*,KNN firstidentifies*k*pointsinthetrainingdatathatareclo-sesttox,whichisrepresentedby N_k.Hence,KNNcanbeviewedasassigningtheKNNsa weight 1/*k* and all others zero weight. The conditional probability that observation *x*belongstoclass*k*isthenestimatedas

Y=kjX=xÞ=$\frac{1}{}$X *Iðy= kÞ* ð4Þ

TheKNNassigns*x*totheclass*k* thathasthelargestprobability.As giveninEquation 4: The choice of k determines the accuracy of the model and the decision border between the classes. A classifier with a smaller k value will have low-bias output class boundaries that are less clear and high-variance output class boundaries that are less responsive to noise.

Table 2 : Statistical Features Using Knn Algorithm For Different Number Of Neighbors

Fig3.Percentageofpredictionfordifferentnumberofneighboursusingknnalgorithm



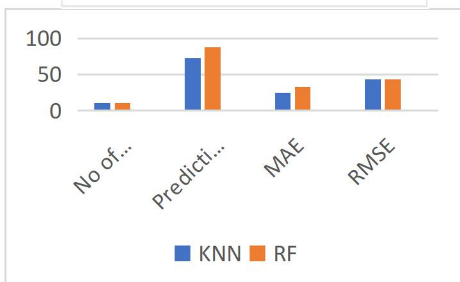
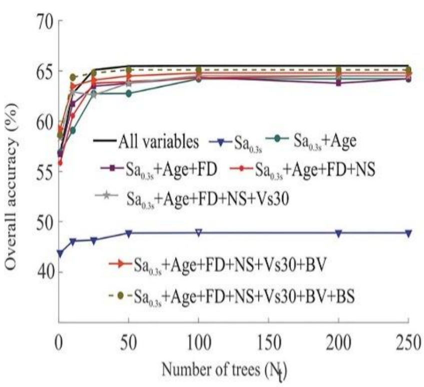


Fig 4: Comparison Prediction between KNN and RF Algorithms

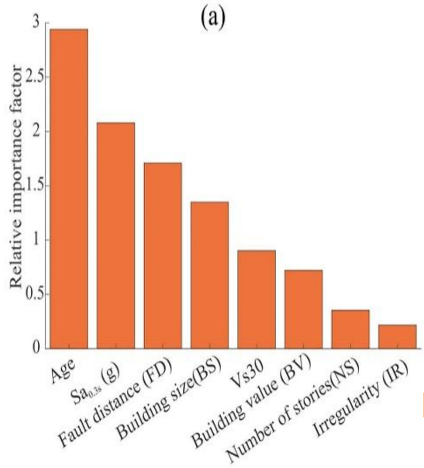
5.RESULTS

ISSN : 2057-5688

The random forest classification yielded the F1 score of 74.32 percent. This strategy is straightforward, practical, and simple to implement in computer programmes. Additionally, it aids in the quick assessment of earthquake damage and loss. It is also highlighted that even with Sa0.3s, fault distance, and age serving as the model's only predictor variables, a good prediction can still be made. With the help of several features, this study was able to construct various earthquake damage scenarios. In this study, we're looking at the various earthquake dam



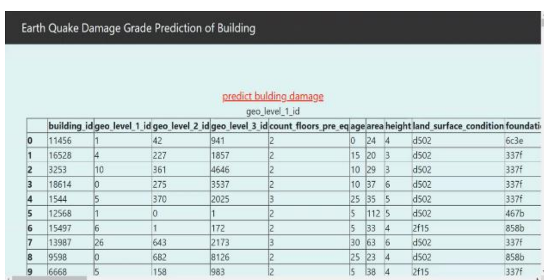
(a)



(b)

age grades in Nepal. This study uses the parameter-tuned Neural Network (NN) and Random Forest algorithms to evaluate the state of the building structures. The analysis reveals that Random Forest method has outperformed theneural network approach for building damage prediction.

Figure5.(a)Variationofoverallaccuracywiththenumberoftreesforvariouscombinationsofinputparameters.(b)Relativeimportanceoftheinputparametersusedinthe RFalgorithm



Earth Quake Damage Grade Prediction of Building

predict bulding damage

geo_level_1_id

	building_id	geo_level_1_id	geo_level_2_id	geo_level_3_id	count_floors_pre_eq	age	area	height	land_surface_condition	foundati
0	11456	1	42	941	2	0	24	4	d502	6c3e
1	16528	4	227	1857	2	15	20	3	d502	337f
2	3253	10	361	4646	2	10	29	3	d502	337f
3	18614	0	275	3537	2	10	37	6	d502	337f
4	1544	5	370	2025	3	25	35	5	d502	337f
5	12568	1	0	1	2	5	112	5	d502	467b
6	15497	6	1	172	2	5	33	4	2f15	858b
7	13987	26	643	2173	3	30	63	6	d502	337f
8	9598	0	682	8126	2	25	23	4	d502	858b
9	6668	5	158	983	2	5	38	4	2f15	337f

Fig 4: Geo-level-1-id

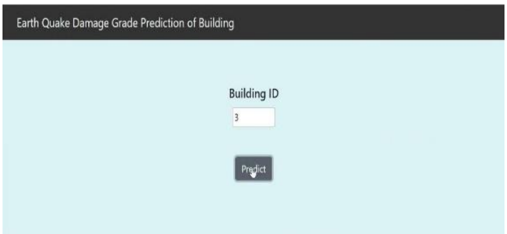


Fig 5: Building Id

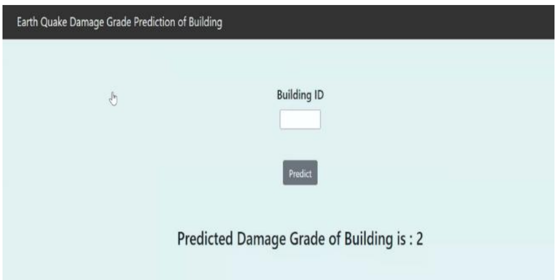


Fig 6: Damage Grade of Building

6.CONCLUSION

The dataset's spatial evaluation highlights the RF-based prediction model's capacity for precise forecasting. To enable broad application of machine learning methods for predicting earthquake-induced building damage, more work is required. This study examined data from a single earthquake that was deemed to have caused moderate physical infrastructure damage. Additionally, the data set only contained low-rise (one to three stories) residential buildings with particular categories of damage (such as chimney, cripple wall, and porch damage). The results might not be applicable to different building heights, construction types, or damage classes. Future studies should compare the effectiveness of machine learning techniques to the current state of the art.Despite all of these drawbacks, this study has demonstrated the enormous potential for employing artificial intelligence to guide emergency response and recovery planning following a disaster. Future research should concentrate on gathering more information that is diverse in terms of the scale, classes, and style of construction. The overall accuracy and generalizability of prediction models can be enhanced in future works by incorporating larger datasets (e.g., from multiple events) and additional site- (e.g., soil conditions) and building-specific

ISSN : 2057-5688

(e.g., presence of known vulnerabilities such as a soft-story) predictor variables. Last but not least, the use of machine learning algorithms can be expanded to include more danger types (like storms) and building effect metrics like financial losses and downtime.

REFERENCES

[1] Goda, K., Kiyota, T., Pokhrel, R.M., Chiaro, G., Katagiri, T., Sharma, K. and Wilkinson, S., 2015. The 2015 Gorkha Nepal earthquake: insights from earthquake damage survey. Frontiers in Built Environment, 1, p.8.

[2]https://www.drivendata.org/competitions/57/nepalearthquake/page/136/ [access date- 24 June, 2019]

[3] Bhargava, N., Katiyar, V. K., Sharma, M. L., & Pradhan, P. (2009). Earthquake prediction through animal behavior: A review. Indian J. Biomech, 78, 159-165.

[4] Alarifi, A. S., Alarifi, N. S., & Al-Humidan, S. (2012). Earthquakes magnitude predication using artificial neural network in northern Red Sea area. Journal of King Saud University-Science, 24(4), 301-313.

[5] Giþev, V., & Trifunac, M. D. (2012). A note on predetermined earthquake damage scenarios for structural health monitoring. Structural control and health monitoring, 19(8), 746-757.

[6] Xu, F., Chen, X., Ren, A., & Lu, X. (2008). Earthquake disaster simulation for an urban area, with GIS, CAD, FEA, and VR integration. Tsinghua Science and Technology, 13(S1), 311-316.

[7] Brunner, D., Bruzzone, L., & Lemoine, G. (2010, July). Change detection for earthquake damage assessment in built up areas using very high resolution optical and SAR imagery. In 2010 IEEE International

Geoscience and Remote Sensing Symposium (pp. 3210-3213). [8] Hosokawa, M., & Hoshi, T. (2001, July). Landform classification method using self-organizing map and its application to earthquake damage evaluation. In IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217) (Vol. 4, pp. 1684-1686).

[9] W. Jiang and B. X. Gu, "Forecasting Seismic Damage in Frame-Shear Wall Structure Based on Fuzzy Neural Network", Advanced Materials Research, Vols. 594-597, pp. 1692-1695, 201