International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 03 | March - 2022 ISSN: 2582-3930

ANALYSIS AND PREDICTION OF EARTHQUAKE IMPACT-A MACHINE LEARNING APPROACH

Mr.PARUCHURU.SRIKANTH, Mr.NETHI HEMANTH SIVA SAI , Mr.S.GOKULKRISHNAN

B.E Graduate(IV year) Computer Science and Engineering,SCSVMV,Kanchipuram,

B.E Graduate(IV year) Computer Science and Engineering,SCSVMV,Kanchipuram,

Assistant Professor,Computer Science and Engineering,SCSVMV,Kanchipuram

ABSTRACT:-

An earthquake is a natural calamity that is well-known for the devastation it causes to both natural and man-made structures such as buildings. Bungalows and residential locations to name a few. Earthquakes are measured using seismometers, that detect the vibrations due to seismic waves travelling through the earth's crust. The damage caused by an earthquake was divided into damage grades in this study, with values ranging from one to five. A previously collected data set was utilised to forecast the damage grade of a given structure using a variety of criteria , which is associated with a Unique Identification String. A survey of available machine learning classifier methods was used to make the forecast. The machine learning algorithms used in this work were Logistic Regression, Naive Bayes Classifier, Random Forest Classifier and K-Nearest Neighbors. Based on an evaluation of a set of attributes, the most appropriate algorithm was considered. A detailed analysis was done on the predicted attribute by the given algorithm, followed by data analysis that provided details that could help mitigate the impact of an earthquake in future.

Keywords: Logistic regression, Naive Bayes Classifiers, Random Forest Algorithm, K-Nearest Neighbors, XGBoost .

I. INTRODUCTION

Leveraging the power of machine learning is a viable

option to predict the degree of damage that is done to buildings post an earthquake. It can help identify safe and unsafe buildings which helps to predict damage prone areas and thus avoiding death and injuries resulting from the aftershock of an earthquake, while simultaneously making rescue efforts efficient. This is done by classifying these structures on a damage grade scale based on various factors like its age, foundation, number of floors, material used and several other parameters. Then the number of families and the probable casualties ward-by-ward in a district are taken into account. This enables distribution of relief forces proportionately ward-wise and its prioritization based on the extent of damage. Models of this kind can help save as many lives as quickly as possible and turn out to be an efficient and costeffective solution. It can be further improved by the inclusion of distribution of resources like food, clothes, medical, monetary supplies based on the extent of human casualties and the damage incurred by the various structures

1.1 Scope of the Project:

Time and quantity of the organization's resources are limiting factors, and organization managers face several difficulties when it comes to the distribution of the resources.

Leveraging the power of machine learning is a viable option to predict the degree of damage that is done to buildings post an earthquake.

It can help identify safe and unsafe buildings which helps to predict damage prone areas and thus avoiding death and injuries resulting from the aftershock of an earthquake, while simultaneously making rescue efforts efficient.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 03 | March - 2022 ISSN: 2582-3930

1.2 Literature Survey:

This paper introduces a rapid assessment tool for earthquake emergencies. The method has two options for extracting damage information from remote sensing images: one uses a damage index and the other uses picture classification. Traditional visual interpretation is used in the damage index mode. After experts have provided the damage index, ground intensity data can be obtained, and loss estimation parameters can then be obtained from the experiential vulnerability matrix. The image categorization mode is a technique used in digital image processing. These loss estimate parameters can be determined using the classification results, which are sorted by building type and ranged by damage severity. The action of multi-resourced estimate data is explained to show how to discover parameters in varied data while the assessment models are introduced.

II. PROJECT DESCRIPTION

2.1 Problem Statement:

An earthquake is a calamitous occurrence that is detrimental to human interest and has an undesirable impact on the environment. Earthquakes have always caused incalculable damage to structures and properties and caused the deaths of millions of people throughout the world. In order to minimize the impact of such an event, several national, international and transnational organizations take various disaster detection and prevention measures. Time and quantity of the organization's resources are limiting factors, and organization managers face several difficulties when it comes to the distribution of the resources.

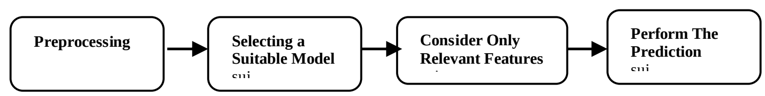
Leveraging the power of machine learning is a viable option to predict the degree of damage that is done to buildings post an earthquake. It can help identify safe and unsafe buildings which helps to predict damage prone areas and thus avoiding death and injuries resulting from the aftershock of an earthquake, while simultaneously making rescue efforts efficient.

III. METHODOLOGIES

1. **Modelling :** Logitic, Navie bayes, Random Forest, KNN, and xgboost . Combine the training using machine learning algorithms and establish a classification model.

This covers the technique and flow of events that were used to perform the prediction process. The prediction methodology itself is composed of three integral steps: data preprocessing, model selection and the final prediction process.

Flow diagram of the methodology step



2. Preprocessing : In the dataset, a building was uniquely identified by 4 attributes: Building Identification, District Identification, Municipality Identification, Ward Identification.

These attributes were added to the training data for identifying the building damage grade. Many attributes in the files related to the building structure and ownership details of the buildings had string data which were converted to their vectorized representation using Label Encoding technique.

3. Model Selection : For any classification problem, a plenitude of machine learning algorithms are available and thus to choose the best among them, it is necessary to evaluate them on the same data based on a suitable parameter.

Here we chose F1 score with 'weighted' average, as calculated in accordance with the content mentioned in the Hackerearth website as our evaluation metric.

Here, TP refers to True Positive, i.e. When the model rightly predicts a positive result.

FP refers to False Positive i.e. When the model wrongly predicts a positive result FN refers to False Negative i.e. When the model wrongly predicts a negative result

4. Prediction The models developed by the individual algorithms were trained on the training dataset and then test data was used for final prediction of damage grades

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 03 | March - 2022 ISSN: 2582-3930

and for evaluation. Since on evaluation, Random Forest Classifier algorithm was found to possess the highest F1 score, the model was considered for the prediction process.
The feature importance method was used to obtain the importance of the features in our model [15]. Out of all the features that were considered in the initial model, the least important features were dropped, thus producing a more accurate prediction of the damage grade.
The threshold value for a feature to be relevant to our model was set to be 0.000200, i.e. any parameter with feature importance less than this value would be dropped from our dataset.

IV. MODULES :

1. **Data Collection :** Collect sufficient data samples and legitimate software samples.

2. **Data Preporcessing :** Perform effective data processing on the sample and extract the features.

3. **Train and Test Modelling :** Split the data into train and test data Train will be used for trainging the model and Test data to check the performace

4. **Modelling :** Logitic, Navie bayes, Random Forest, KNN, and xgboost . Combine the training using machine learning algorithms and establish a classification model.

V. RESULTS :

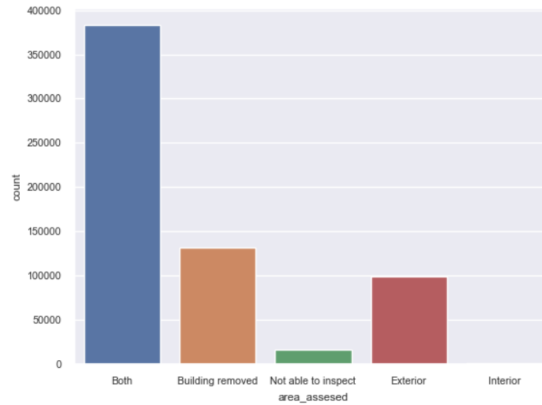


Fig.1. Result Diagram

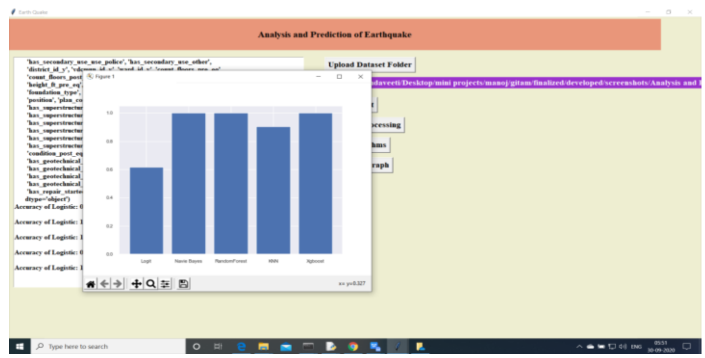


Fig.2. Result Diagram

This work presents that the Random Forest Classifier algorithm has the highest accuracy in predicting the damage due to earthquakes, based on the F1 score calculated for each of the four algorithms previously mentioned in this work.
Earthquakes are well known to excite electromagnetic pulse, that cause tremors under the Earth's crust. These electromagnetic pulses are shielded effectively by Reinforced Concrete. Reinforced concrete has a low tensile strength, and hence Steel bars are used, which are embedded in the concrete sets.

VI . CONCLUSION

This work presents that the Random Forest Classifier algorithm has the highest accuracy in predicting the damage due to earthquakes, based on the F1 score calculated for each of the four algorithms previously mentioned in this work.

K-Nearest Neighbors has been observed to be the second most preferred algorithm for earthquake damage prediction. On analysis of the materials that help curb damage to buildings during an earthquake, the work concludes that Reinforced Concrete is the material most suited to the cause.

Earthquakes are well known to excite electromagnetic pulse, that cause tremors under the Earth's crust. These electromagnetic pulses are shielded effectively by Reinforced Concrete. Reinforced concrete has a low tensile strength, and hence Steel bars are used, which are embedded in the concrete sets.

This provides Reinforced Concrete with immense ability to withstand natural calamities such as Earthquakes. This fact justifies the reason for the widespread presence of Reinforced Concrete among the buildings with Earthquake Damage grade 1, and its minimal presence in buildings with Earthquake Damage grade 5. The applications of this work can be further extended to predict damage caused by Earthquakes in areas for which a similar and relevant dataset can be obtained.

VII. REFERENCES

1. C.P Shabariram and K.E Kannammal "Earthquake prediction using map reduce framework" 2017 International Conference on Computer Communication and Informatics (ICCCI), pp. 1–6, 2017.

2. You-Po Su and Qing-Jie Zhu "Application of ANN to Prediction of Earthquake Influence" 2009 Second International Conference on Information and Computing Science, vol. 2 pp. 234-237, 2009.

3. Cao Li and Xiaoyu Liu, "An improved PSO-BP neural network and its application to earthquake prediction," in 2016 Chinese Control and Decision Conference 2016, pp. 3434–3438.

4. Xueli Wei, Xiaofeng Cui, Chun Jiang, Xinbo Zhou "The Earthquake Probability prediction based on weighted factor coefficients of principal components "2009 Fifth International Conference on Natural Computation, vol. 2, p. 608-612, 2009

5. Dezhang Sun, Baitao Sun "Rapid prediction of earthquake damage to buildings based on fuzzy analysis"2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, vol. 3, p. 1332-1335, 2010

6. Jianwen Zhang, Changyong Wang "Sheer Capacity computation of steel reinforced lightweight concrete beams" International Conference on Mechanic Automation and Control Engineering, 2010. p 1502-1506

7. Python Seaborn Library pydata official website

8. A Competitive Machine Learning series 'Machine Learning Challenge #6' problem 'Predicting the damage to a building'