International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

Time Series Earthquake Prediction Model

Dr. P. Srinivasa Rao*[1], Aditi Mantha[2], Vaishnavi Loya[2], Tummala Kranthi Priya[2], Thundla Aruna[2]
[1]Professor, CE, Department of Computer Science and Engineering, JB Institute of Engineering & Technology, Hyderabad, Telangana, India.
[2]Students, Department of Computer Science and Engineering, JB Institute of Engineering & Technology, Hyderabad, Telangana, India

ARTICLEINFO

Article History:

Accepted: 01 April 2023
Published: 10 April 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
335-339

ABSTRACT

Natural calamities like earthquake cause damage to life and property. Assessment of harm grade to structures is fundamental for post-disaster reaction and recuperation end of the monotonous course of manual approval and confirmation of property harm prior to giving alleviation assets to individuals. By taking into account essential perspectives like structure area, age of the building, development subtleties and it's auxiliary purposes, taken from the Gorkha seismic tremor dataset, this paper investigates different multi-class grouping AI models and procedures for anticipating the harm grade of designs. The proposed engineering of the model includes three significant stages, Component Choice, genetic algorithm adaboost Randomforest Classifier, and adaboost decision tree Classifier. and adaboost decision tree Classifier. The paper gives the after effects of the analyses highlight designing, preparing varieties and gathering learning. The paper dives into the examination of each model, to figure out the explanation for their presentation. This paper likewise gathers the specialists that play a significant job in choosing the earthquake damage of the structures. The proposed classifier in the paper gives critical contribution to understanding earthquake damage and also provides a paradigm to model other natural disaster damage.

Keywords: Earthquake, Flutter, Adaboost, Random Forest, Classifier.

I. INTRODUCTION

There is a significant need to survey the damage done by earthquakes in disaster hit locales. Anticipating the harm done to structures helps in recognizing the recipients qualified for government help for lodging recreation. It is a basic piece of the post-disaster recuperation stage. Manual recognizable proof and characterization demand huge investment and assets and prompts trouble among the populace. There have been endeavors made to anticipate the seismic damage utilizing satellite symbolism , of earthquakes hit locales, yet all at once these are register escalated. The April 2015 Nepal tremor, the Gorkha earthquake, was quite possibly of the most obliterating debacle that has at any point hit Nepal. It had a second greatness of 7.8 on the

Richter Scale. It killed around 9000 individuals, additionally causing wounds to 22000 others. In the result of the Gorkha earthquake, Nepal played out a broad family overview by utilizing portable innovation to assess the degrees of building harms in those areas impacted by the quake. The study, which is currently one of the biggest post-calamity datasets ever incorporated, contains helpful data on the effect of the quake, the family conditions, and the financial socioeconomics of the fiasco. This paper expects to assess and estimate the harm caused to each structure by the 2015 Gorkha earthquake in view of the qualities of the structure area and its development.

II. LITERATURE SURVEY

Earthquake magniture prediction for Hindukush region has been carried out in this research using temporal sequence of historic seismic activities in combination with the machine learning classifiers. Prediction has been made on the basis of mathematically calculated eight seismic indicators using the earthquake catalog of the region. These parameters are based on the well-known geophysical facts of Gutenberg–Richter's inverse law, distribution of characteristic earthquake magnitudes and seismic quiescence. In this research, four machine learning techniques including pattern recognition neural network, recurrent neural network, random forest and linear programming boost ensemble classifier are separately applied to model relationships between calculated seismic parameters and future earthquake occurrences

The problem is formulated as a binary classification task and predictions are made for earthquakes of magnitude greater than or equal to 5.5 ($$M \ge$$M⩾ 5.5), for the duration of 1 month. Furthermore, the analysis of earthquake prediction results is carried out for every machine learning classifier in terms of sensitivity, specificity, true and false predictive values. Accuracy is another performance measure considered for analyzing the results. Earthquake magnitude prediction for the Hindukush using these aforementioned techniques show significant and encouraging results, thus constituting a step forward toward the final robust prediction mechanism which is not available so far.

III. PROPOSED METHODOLOGIES

In proposed system, we used some machine learning algorithms Gradient Boosting,Gradient Boosting is a famous boosting algorithm. In gradient boosting, every indicator revises its predecessor's error. As opposed to Adaboost, the loads of the preparation occasions are not changed, all things considered, every indicator is prepared involving the remaining errors of predecessor as labels There is a strategy called the Gradient Boosted Trees whose base learner is CART (Classification and Regression Trees). we used adaboost random forest ,adaboost decision tree Classifier models also. The spiral model is an SDLC model that combines elements of an iterative software development model with a waterfall model. It is advisable to use this model for expensive, large and complex projects. In its diagrammatic representation, we have a coil having many cycles or loops. The number of cycles varies for each project and is usually specified by the project manager. Each spiral cycle is a stage in the software development process.

Plan objectives and find alternate solutions

This phase includes requirement gathering and analysis. Based on the requirements, objectives are defined, and different alternative solutions are proposed.

Risk analysis and resolving

In this quadrant, all the proposed solutions are analyzed, and any potential risk is identified, analyzed, and resolved.

Develop the next version of the product

This phase includes the actual implementation of the different features. All the implemented features are then verified with thorough testing.

Plan the next phase

In this phase, the software is evaluated by the customer. It also includes risk identification and monitoring like cost overrun or schedule slippage and after that planning of the next phase is started.

IV. EXPERIMENT

This section describes the experiment and how it was planned, as well as its objectives and the procedure itself. It is done to help the reader understand how the experiment was carried out and planned.

Here for a machine learning projects .we generally collect the input from online websites and filter the input data and visualize them in graphical format and then the data is divided for training and testing . That training is testing data is given to the algorithms to predict the data .

1. First, we take dataset.
2. Filter dataset according to requirements and create a new dataset which has attribute according to analysis to be done
3. Perform Pre-Processing on the dataset
4. Split the data into training and testing
5. Train the model with training data then analyze testing dataset over classification algorithm
6. Finally you will get results as accuracy metrics.

Data collection is a process in which information is gathered from many sources which is later used to develop the machine learning models. The data should be stored in a way that makes sense for problem. In this step the data set is converted into the understandable format which can be fed into machine learning models.Data used in this paper is a set of cervical cancer data with 15 features . This step is concerned with selecting the subset of all available data that you will be working with. ML problems start with data preferably, lots of data (examples or observations) for which you already know the target answer. Data for which you already know the target answer is called labelled data.

Organize your selected data by formatting, cleaning and sampling from it. Three common data pre-processing steps are:

Formatting: The data you have selected may not be in a format that is suitable for you to work with. The data may be in a relational database and you would like it in a flat file, or the data may be in a proprietary file format and you would like it in a relational database or a text file.

Cleaning: Cleaning data is the removal or fixing of missing data. There may be data instances that are incomplete and do not carry the data you believe you need to address the problem. These instances may need to be removed. Additionally, there may be sensitive information in some of the attributes and these attributes may need to be anonymized or removed from the data entirely.

Sampling: There may be far more selected data available than you need to work with. More data can result in much longer running times for algorithms and larger computational and memory requirements. You can take a smaller representative sample of the selected data that may be much faster for exploring and prototyping solutions before considering the whole dataset.

Next thing is to do Feature extraction is an attribute reduction process. Unlike feature selection, which ranks the existing attributes according to their predictive significance, feature extraction actually transforms the attributes. The transformed attributes, or features, are linear combinations of the original attributes. Finally, our models are trained using Classifier algorithm. We use classify module on Natural Language Toolkit library on Python. We use the labelled dataset gathered. The rest of our labelled data will be used to evaluate the models. Some machine learning algorithms were used to classify pre-processed data. The chosen classifiers were Random forest. These algorithms are very popular in text classification tasks. Model Evaluation is an integral part of the model development process. It helps to find the best model that represents our data and how well the chosen model will work in the future. Evaluating model performance with the data used for training is not acceptable in data science because it can easily generate overoptimistic and over fitted models. There are two methods of evaluating models in data science, Hold-Out and Cross-Validation. To avoid over fitting, both methods use a test set (not seen by the model) to evaluate model performance.

V. CONCLUSION

The time intervals are not uniform, as expected from many sensor data with limited sensitivity to events; some earthquake data have 8 milliseconds intervals and some nearly 1 day, hence time series is irregular. This complicates the time-series modeling which I will cover in the future blog post. 5. EDA — 50 years of earthquake dat

VI. FUTURE WORK

In the future post, we will use the developed features, specifically, the clustering features to perform more advanced time-series analysis. Based on early evaluation, we found LSTM as an appropriate method for my purposes.

Two aspects specifically brought our interest towards LSTM sequential analysis: 1 — the irregular nature of my time-series. 2 — the shock-like nature of the earthquake without much apparent precursors.

VII. REFERENCES

[1]. Andrews, D. F. A robust method for multiple linear regression. Technometrics 16, 4 (1974), 523–531.

[2]. Asim, K., Mart´ınez-´Alvarez, F., Basit, A., and Iqbal, T. Earthquake magnitude prediction in hindukush region using machine learning techniques. Natural Hazards 85 (01 2017), 471–486.

[3]. Bhandarkar, T., K, V., Satish, N., Sridhar, S., Sivakumar, R., and Ghosh, S. Earthquake trend prediction using long short-term memory rnn. International Journal of Electrical and Computer Engineering (IJECE) 9 (04 2019), 1304.

[4]. Breiman, L. Random forests. Mach. Learn. 45, 1 (Oct. 2001), 5–32.

[5]. Geller, R., Jackson, D., Kagan, Y., and Mulargia, F. Earthquakes cannot be predicted. Science 275 (1997), 1616 – 1616.

[6]. Graves, A., Mohamed, A.-r., and Hinton, G. Speech recognition with deep recurrent neural networks. In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (2013), pp. 6645–6649.

[7]. Hochreiter, S., and Schmidhuber, J. Long short-term memory. Neural computation 9 (12 1997), 1735–80.

[8]. Kuyuk, H. S., and Susumu, O. Real-time classifification of earthquake using deep learning. Procedia Computer Science 140 (2018), 298–305.

[9]. Li, A., and Kang, L. Knn-based modeling and its application in aftershock prediction. In Proceedings of the 2009 International Asia Symposium on Intelligent Interaction and Affffective Computing (USA, 2009), ASIA '09, IEEE Computer Society, p. 83–86.

[10]. Mallouhy, R., Abou Jaoude, C., Guyeux, C., and Makhoul, A. Major earthquake event prediction using various machine learning algorithms. In International Conference on Information and Communication Technologies for Disaster Management (Paris, France, Dec. 2019).

[11]. Shearer, P. M. Introduction to Seismology, 2 ed. Cambridge University Press, 2009.

[12]. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfifitting. Journal of Machine Learning Research 15 (06 2014), 1929–1958.

[13]. USGS. Signifificant earthquakes - 2021. https://earthquake.usgs.gov/earthquakes/browse/signifificant.php. ”Accessed: 2021-17-06”.

[14]. Zhang, A., Lipton, Z. C., Li, M., and Smola, A. J. Dive into Deep Learning. 2019.

Cite this article as :

Dr. P. Srinivasa Rao, Aditi Mantha, Vaishnavi Loya, Tummala Kranthi Priya, Thundla Aruna, "Time Series Earthquake Prediction Model", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 2, pp. 335-339, March-April 2023.

Journal URL : https://ijsrset.com/IJSRSET2310249