Earthquake Prediction Using Seismic Information

Madhushree N * and Thiruukumaran

Abstract

Earthquake is one of the most hazardous, devastating natural calamities and yet a very least predictable natural disaster that occur. Prediction of earthquake has been a challenging research for many researchers. With the increasing amount of earthquake dataset collected, many researchers try to solve the task of predicting the earthquake in future time. Even though many data mining techniques are been used, the prediction rate is not still accurate due to lack of feature extraction technique. The proposed methodology enhances the performance of earthquake prediction. As obtained precursory pattern features CART algorithm is used to get accurate prediction of both the magnitude range and an effective time range of future earthquakes.

Keywords

Earthquake; Precursory pattern; Magnitude; Time range; Cart

Introduction

Earthquake is one of the most complex, hazardous natural calamities. Usually an earthquake is the shaking of the surface of the earth which produce or release energy over the earth’s lithosphere that form the seismic waves. The seismicity or seismic activity of an area is termed as frequency, type and size of earthquake, which is taken over a period. The scale of measuring the earthquake magnitudes was developed by Richter in the year 1935[1]. Even though the media reports of earthquake magnitude as Richter magnitude or Richter scale, standard practice by most of the seismological authorities express as an earthquake strength on the moment magnitude scale, which is usually based on actual energy released by the earthquake.

An overall estimation is about 500,000 earthquake that occur every year, about 100,000 of these can be felt minor earthquakes occur in places like California and Alaska in USA as well as Iran, Chile, Turkey, Japan, Greece, Italy, Mexico, India, Nepal, Pakistan etc. As the data set of earthquake collected is increasing, many researchers are trying to solve the prediction of earthquake that is used to count time, magnitude, places of the future earthquake. With data mining techniques a numerous research scholars discovered patterns of earthquake from seismic time series, but the prediction is not accurate, due to lack of feature extraction.

Feature Extraction is a method of enhancing machine learning by finding out the characteristics in the data which tries to help the particular problem. While features can be obtained or taken by set of sequence comparison techniques followed which are like dynamic time wrapping and by subsequent discovery technique. Features usually contain the properties of the data set. Extracting data mining techniques cannot be used on time series data. Thus, a dimension reduction is required through feature extraction technique, which map each and every time series. While, the project here proposes a precursory pattern based feature extraction technique for predicting the earthquake. With data mining techniques a numerous research scholars discovered patterns of earthquake which is capable of predicting the magnitude of the occurrence of earthquake and also the effective time range of earthquake.

A precursory pattern of earthquake represent the part involved in seismic records before activating the main shock this is considered as precursory pattern of earthquake. In order to obtain, at first a seismic data is divided into a fixed day time period from the magnitude occurred by the largest earthquake of each time period it involved is known as main shock. The set of sequence involved of the last events during last time that has occurred before the present time is here considered as precursory pattern. Seismic indicators are considered from the precursory patterns by CART algorithm [2].

Objectives

1. Feature extraction (Precursory pattern based) to better capture the characteristics of earthquake, thus it can be used to enhance the earthquake prediction.

2. Applying machine learning algorithm Classification and Regression Tree (CART) algorithm on the extracted features to predict the magnitude and effective time range of earthquake on Richter scale.

Literature Survey

Gutenberg, had made use of statistical feature method approach and formed a seismic indicators. Likewise, magnitude, energy, acceleration of the earthquake and so on [1]. Nuannin, made use of sliding time and space window which is eventually based on seismic events to gain the indicators of earthquake [3]. Florido, from the above obtained seismic events eventually considered a few number of fixed seismic events which formed before the main earthquake happened, like a precursory pattern in order to extract the necessary features but the proposed method was in efficient to determine the magnitude of the earthquake of the main shock.

Narayana kumar, had extracted the seismic features indicating the fixed number of events that occurred in prior to main shock, using BP neural network. The results demonstrate that accuracy is better for medium and large earthquakes [4]. Adeli, proposed a new feature extraction where it solved the problem of magnitude and an effective time range of prediction. The author used the probabilistic neural netwok (PNN) which provided with good prediction results of earthquake whose magnitude lies in between 4.5 to 6.0 Richter [5]. Mirrashid, had established the prediction of earthquake with the magnitude 5.5 or even more by making use of ANFIS (adaptive Neuro Fuzzy Interface System). The obtained results claimed best results. Asencio, proposed a clustering method mainly for the seismogenic zone and then perform machine learning algorithm to make the
prediction of earthquake, which included ANN, KNN, NB, decision tree, SVM.

Mart, in order to obtain the relevant feature selection redundant features were to be eliminated, the author used information gain of each one of the seismic indicators for the selection. Asim proposed a hybrid embedded feature selection method where precise prediction was established. Hamze-Ziabari, recently build on bagging ensemble model and CART (classification and regression tree) algorithm were used to predict the ground acceleration.

**Proposed methodology**

This project proposed a precursory pattern based feature extraction method along with the CART (classification and regression tree) algorithm, to predict the earthquake.

**Methodology**

In order to form the learning sample, initially the raw seismic data time period while the magnitude formed of the longer earthquake including with each time period termed as the main shock. Sequence which contain of prior events in the due of prior time period before the present time and events usually are prior the main shock of present time period is considered as the earthquake precursory pattern. From this the seismic indicators obtained based on precursory patterns with CART [6] (classification and regression tree) algorithm which in turn leads to the prediction of earthquake. While to predict the magnitude as well as obtain the good time range of the upcoming earthquake, we are using the precursory pattern feature extraction method. Initially the seismic records obtained is divided into a N day time period, while the magnitude of the higher earthquake equivalent to the time period will be considered as a systematic seismic indicators.

In general, a decision tree which usually contains the root node, internal node, leaf node. Thus, the key of this algorithm is to pick the partitioning attribute based on the data purity. Information gain is made to compute the purity of data, By applying CART algorithm we will be able to predict the earthquake magnitude. The following figure represent the pictorial representation on how the methodology works as shown in the below Figure 1.

Steps involved:

- Loading the data from seismic wave data servers.
- Extracting seismic information and preprocessing.
- Extraction of feature which is formation of precursory pattern
- Use precursory pattern to form statistical feature called as seismic indicator.
- Once this is formed input these feature with labels will be a input to decision tree algorithm.
- Pruning will be performed for generalization of model heuristically.
- Feature importance scores will be formed to understand effect of the seismic indicators.

**CART (classification and regression tree)**

CART (classification and regression tree) algorithm was introduced in 1984, It is a binary decision tree where the node is divided into two sub node started with root node. It is tree which output either a classification or a regression tree. CART algorithm is structured as a sequence of question, the answer to which determine what the next question.

Classification tree: The target variable is unqualified and the tree is used to identify the class where a target variable would be. Below figure represent the classification tree involved. As shown in Figure 2. Regression tree: The target variable is continuous and tree which is obtained is used to predict the values. Below figure represent the regression tree involved. As shown in Figure 3.

**Elements of cart**

An act for dividing the data at a node based on the value of a single variable. Stopping rules for deciding when a branch is terminal and cannot be further divided. Eventually, prediction for the target variable in each extreme node.

**Pruning**

Pruning is an approach in machine learning where mainly deal with the over fitting for the decision tree. It mainly focus by reducing the size of the decision tree by removing parts of the tree that is not able to provide to classify the instance. Therefore it becomes an essential aspects to pruning the tree in order to avoid over fitting.
Limitations

This is a single-center retrospective study, and the results need to be validated with a large multicenter prospective study. There were fewer patients in the PG group during the study period since it is a newer technique compared to the older and more established RG modality. In addition, there were a few differences between the two groups in terms of baseline characteristics. Patients in the PG group were younger since more imagers opted for this technique in this population in order to minimize radiation exposure. Furthermore, RG was more commonly chosen in obese patients in order to obtain better image quality. Despite these differences, the multivariate analysis showed only the presence and extent of coronary artery disease remaining as significant predictors. Furthermore, our patient population and event rates were similar to previously published studies. We did not include ventricular function data, which was available for the RG group, since it could only add incremental information to the coronary data.

Conclusion

In this project, proposed a precursory pattern based feature extraction method along with CART algorithm for the accurate prediction. By analysing the results, CART along with precursory pattern will provide the best performance, to solve the task of earthquake prediction for predicting magnitude and an effective time range.

References

1. Gutenberg B, Richter CF (1956) Earthquake magnitude, intensity, energy, and acceleration (second paper): Bulletin of the Seismological Society of America 46: 105-145.
2. Lei Zhang, Langchun Si, Haipeng Yang, Yuanzhi Hu, Jianfeng Qiu (2019) Precursory pattern based feature extraction technique for earthquake prediction: IEEE Access 7:30991-31001.
3. Nuannin P, Kulhanek O, Persson L (2005) Spatial and temporal b value anomalies preceding the devastating off coast of nw sumatra earthquake of december 26, 2004: Geophysical Research Letters, 32:339–357.
4. Narayanakumar S, Raja K (2016) A BP artificial neural network model for earthquake magnitude prediction in himalayas, india: Circuits and Systems, 7:3456–3468.
5. Adeli H, Panakkat A (2009) A probabilistic neural network for earthquake magnitude prediction: Neural Networks, 22:1018-1024.
6. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Cart: Classification and regression trees: Encyclopedia of Ecology, 40:582–588.

Author Affiliation

Department of Engineering, New Horizon College of Engineering, Bangalore, Karnataka, India.