Data Mining in Earthquake Prediction

by . Kumaran T

Submission date: 18-Mar-2019 03:41PM (UTC+0530)
Submission ID: 1095254986
File name: 162235_draft.docx (61.97K)
Word count: 1734
Character count: 9456

Data Mining in Earthquake Prediction

Kumaran T[b], Udhaya Kumar S[b], Vignesh J[b]

[b]*School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India*

Abstract

Earthquakes are the natural disasters that basis a lot of damage, both in human lives and property wise. There are two categories to forecast the earthquakes, they are forecasts and short-term predictions. Occurrence of earthquake can be identified and its severity can also be measured with the support of the parameters. Generally in forecasts we will get to know about the earthquakes much in earlier like months or years in advance, where as in short-term predictions we will know about the happening of earthquakes just before hours or days. The parameters we considered in predicting the earthquakes are latitude, longitude, magnitude (MAG), MMI, day, month and year. The data mining method such as ID3 (decision tree) will be implemented.

Keywords: earthquakes, data mining.

1. Introduction

Earthquake are the sudden movements that are created in the earth's crust by the unexpected release of stress that is accumulated in the geologic fault in the interior. The earthquake prediction has been going on for almost a century. Here a successful forecast, identifying the time, place and magnitude can save many lives and wealth too. The data that is obtained from these trainings is used to provide rough estimations of earth quake magnitude and reappearance intervals.

2. Problem Description

Two foremost problems are adversely disturbing the forecast of earthquakes; (1) effective contact to distributed unstructured earthquake data (both Web and offline sources); and (2) absence of infrastructure for precise and accurate earthquake info access and retrieval. If the earthquake information are not suitably integrated, info that is shared by lots of seismological observatories may lead to letdown of information building and prediction.

3. Motivation

Among the huge amounts of information related to earthquakes, we are implementing various algorithms to mine the data available and to check for the best and accurate way to forecast the earthquakes using the parameters like day, month, year, hours, minutes, seconds, place, latitude, longitude, magnitude, MMI and occurrence of the earth quake using Rapidminer tool.

4. Objective:

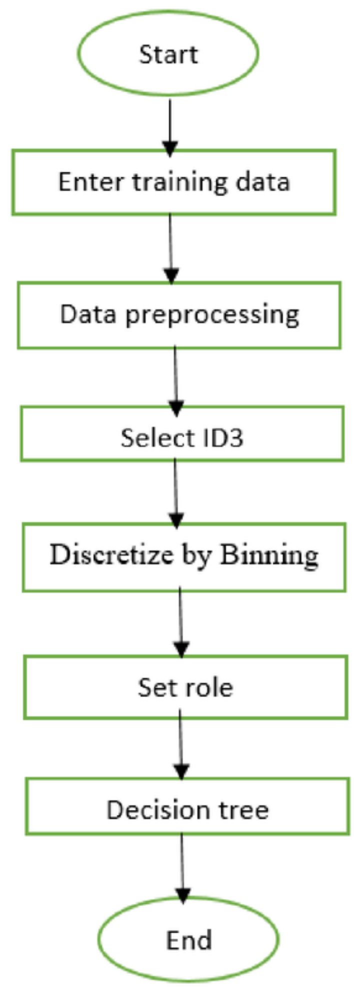
Here our main objective is to implement the algorithm like decision tree to the available pre-processed data set, which has been taken from an already existing huge data base. And after implementation the outcomes obtained can be compared to maintain the most accurate value in the forecast of earthquakes.

5. Literature Review

1) Earthquake data analyzing using apache Hadoop framework with sample earthquake dataset of year 2010 the analysis result shows next earthquake location and time.
2) Based on latitude lngtd and main depth the earth quake mag forecast model is built using svm and neural networks combination of svm and neural networks produces the more accurate than svm and neural network.
3) Improved LSTM network with two dimensional input based on spatial temporal perspective gives better results of prediction of earthquake. Simulation results shows better accuracy.
4) Simulation results shows BP neural network improved by IPSO algorithm is improved than the BP NN enhanced by PSO algorithm.
6) Seismographic records and analysis includes identification of measurements the usage of s parameter improved the efficiency of seismographic analysis.
7) Three types of earthquakes shallow subduction back arc and deep subduction are identified. Correlation and nonlinear regression were approved to improve GMPEs of Malaysia.
8) Using Bayesian network the outcome of 86% accuracy is obtained without simulated and annealing.
9) The efficient is not improved in BPGA NN for Himalayan area the analysis with BP NN shows the maximum magnitude and number of earthquakes.
10) The trained dataset is derived from RNM signal data. The software receives signal from different stations and records in a database.
11) ANFIS SVRNN and RBFN models predict the PGA values is 0.5 g in different areas.

6. Method:
ID3 Algorithm

The ID3 starts with original set as root node. In this algorithm it repeats over every vacant attribute of the set and computes the information gain of that element which has the minimum or main information gain value. The set is before split by the particular attribute to crop subset of data. For example Node can be divided into child nodes based upon the subgroup. The algorithm remains to persist on each subgroup, considering only elements never specific before. In this algorithm throughout, the decision tree is built with apiece internal node demonstrating the nominated element on which the data was divided, and leaf node representing class label of the last subset of this branch. By this algorithm calculates the info gain of every attribute of the data set s. Divided the set s into subsets using element for which the subsequent info gain is higher. Make a decision tree node having that attribute. Recur on subsets using lasting attribute.



7.Steps followed for analysis

8. Related Work

In this section we are explaining detail about the earth quake prediction. Earth quake forecast are classified in categories they are First, earthquake forecast can be done by the mathematical or statistical tool. Spatial connection theory help to predict earthquake epi centre. Poisson range identifier function are used in prediction. Boucouvalaset al. improve the Fibonacci, dual and lucasmethod and propose an predict earthquake s by using a trigger planter aspect date prior to a strong earthquake as a seed for the describing of fdl time 56+curved. And these methods used to test the restricted amount of data. Second, earthquake prediction also done by precursor signal studies. Precursor of significant earthquakes will take electromagnetic signal. Satellite pictures of clouds before earthquake have studied by Thomas et al. And Fan et al. earthquake predecessor based on lithosphere atmosphere-ionosphere combination and relations done by ko repanov. Conduct of animals are abnormal before 10 days of earthquake study was done by Hayakawa et al. With the precursor signal we cannot draw the conclusion and it will not prime to accurate prediction result. Third using machine learning we can also do the earthquake prediction. We can also use the data mining algorithm such as j48, Ada Boost, information network, k-nearest neighbors to forecast the earthquake using the seismic event based on the noted seismic event in the matching region. New seismic indicator which was developed by Gutenberg Richter Ratio is further useful than those conventionally used in earthquake prediction. Machine learning method uses seismic indicators for prediction and mostly time domain but not space domain correlations are studied.

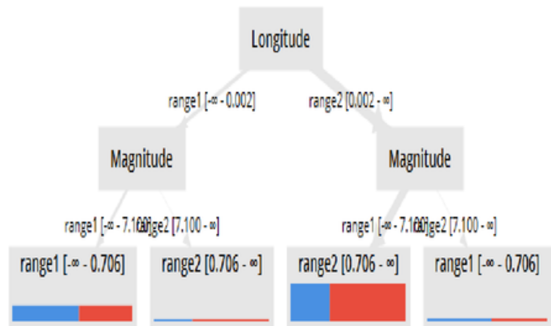
9. Dataset

The data were extracted from the earthquake dataset. We are using three variables to predict the earthquake. The variables we are using from the dataset are Latitude, Longitude, Magnitude.

10. Results:

Using ID3 Algorithm we implemented the experiment by using the three variables such as latitude, Longitude, Magnitude. As using ID3 Algo the gain value for attributes longitude is high, hence it is considered as the root node and then classification based on the magnitude gain value and further classification is done based on gain value of the latitude variable.

ID3 (Decision tree)



ID3 DESCRIPTION:

Longitude = range1 [-∞ - 0.002]
| Magnitude = range1 [-∞ - 7.100]: range1 [-∞ - 0.706] {range1 [-∞ - 0.706] =75, range2 [0.706 - ∞] =60}
| Magnitude = range2 [7.100 - ∞]: range2 [0.706 - ∞] {range1 [-∞ - 0.706] =1, range2 [0.706 - ∞] =2}
Longitude = range2 [0.002 - ∞]
| Magnitude = range1 [-∞ - 7.100]: range2 [0.706 - ∞] {range1 [-∞ - 0.706] =118, range2 [0.706 - ∞] =230}
| Magnitude = range2 [7.100 - ∞]: range1 [-∞ - 0.706] {range1 [-∞ - 0.706] =7, range2 [0.706 - ∞] =6}

11. Conclusion and Future Work

Prediction is done with ID3 as we could monitor the results of the built model's accuracy ranging 50-55%. In order to obtain more accuracy we planned to implement neural networks.

12. References

[1]. Shabariram, C. P., & Kannammal, K. E. (2017, January). Earthquake prediction using map reduce framework. In *2017 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-6). IEEE.

[2]. Zhou, W. Z., Kan, J. S., & Sun, S. (2017, July). Study on seismic magnitude prediction based on combination algorithm. In *2017 9th International Conference on Modelling, Identification and Control (ICMIC)* (pp. 539-544). IEEE.

[3]. Wang, Q., Guo, Y., Yu, L., & Li, P. (2017). Earthquake prediction based on spatio-temporal data mining: an LSTM network approach. *IEEE Transactions on Emerging Topics in Computing*.

[4]. Li, C., & Liu, X. (2016, May). An improved PSO-BP neural network and its application to earthquake prediction. In *2016 Chinese Control and Decision Conference (CCDC)* (pp. 3434-3438). IEEE.

[5]. Lin, J. W., Chao, C. T., & Chiou, J. S. (2018). Determining Neuronal Number in Each Hidden Layer Using Earthquake Catalogues as Training Data in Training an Embedded Back Propagation Neural Network for Predicting Earthquake Magnitude. *IEEE Access*, *6*, 52582-52597.

[6]. Mochizuki, M., Uehira, K., Kanazawa, T., Kunugi, T., Shiomi, K., Aoi, S., ... & Sekiguchi, S. (2018, May). S-Net project: Performance of a Large-Scale Seafloor Observation Network for Preventing and Reducing Seismic and Tsunami Disasters. In *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*(pp. 1-4). IEEE.

[7]. Klikushin, Y. N., Kobenko, V. Y., Koshekov, K. T., Belosludtsev, O. M., & Koshekov, A. K. (2016, November). Search of the operational earthquake precursors on the basis of the identification measurements of the seismographic records. In *2016 Dynamics of Systems, Mechanisms and Machines (Dynamics)* (pp. 1-6). IEEE.

[8]. Liew, M. S., Danyaro, K. U., Mohamad, M., Shawn, L. E., & Aulov, A. (2017). Ground Motion Prediction Equations for Malaysia Due to Subduction Zone Earthquakes in Sumatran Region. *IEEE Access*, *5*, 23920-23937.

[9]. Wicaksono, Y. A., Kurniawan, D., & Santoso, H. A. (2018, September). Bayesian Network Optimization for Earthquake Prediction Using Simulated Annealing. In *2018 International Seminar on Application for Technology of Information and Communication* (pp. 494-497). IEEE.

[10]. Tao, Z. (2015, August). Artificial Neural Network attempts for long-term evaluation of great earthquakes. In *2015 11th International Conference on Natural Computation (ICNC)* (pp. 1128-1132). IEEE.

[11]. Vahaplar, A., Tezel, B. T., Nasiboglu, R., & Nasibov, E. (2015, October). A monitoring system to prepare machine learning data sets for earthquake prediction based on seismic-acoustic signals. In *2015 9th International Conference on Application of Information and Communication Technologies (AICT)* (pp. 44-47). IEEE.

[12]. Ahmadi, M., Nasrollahnejad, A., & Faraji, A. (2017, March). Prediction of peak ground acceleration for earthquakes by using intelligent methods. In *2017 5th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)* (pp. 7-12). IEEE.

Data Mining in Earthquake Prediction

ORIGINALITY REPORT

5%	0%	5%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Shastri L Nimmagadda, Heinz Dreher. "Ontology based data warehouse modeling and mining of earthquake data: prediction analysis along Eurasian-Australian continental plates", 2007 5th IEEE International Conference on Industrial Informatics, 2007 Publication	4%
2	M. S. Liew, Kamaluddeen Usman Danyaro, Mazlina Mohamad, Lim Eu Shawn, Aziz Aulov. "Ground Motion Prediction Equations for Malaysia Due to Subduction Zone Earthquakes in Sumatran Region", IEEE Access, 2017 Publication	1%

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		