International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 1 | ISSN : 2456-3307

Big Data Analytics using Back Propagation Algorithm for Foretelling Quake

K. Kanmani

Department of Computer Science, SRM University, Chennai, India

ABSTRACT

Quakes are without a doubt most appalling, unavoidable, common cataclysm on Earth. Seismic tremor is a sudden development of the worlds outside layer brought on by the arrival of the anxiety collected along geologic deficiencies by volcanic activity. Earthquake occurrence is one of the significant events in nature that causes both irretrievable financial and physical harm. The Precarious condition emerging because of tremors could be kept away from just by making a solid indicative to anticipate the area, greatness and time of eminent earth quakes. The mechanism of the quake stays to be researched, however a few inconsistencies connected with earthquakes have been found by DEMETER satellite perceptions. It is a helpful and practical approach to utilize the self-versatile counterfeit neural system to develop relations between different manifestation variables and seismic earthquake occurrences. Big data is much more than storage of and access to data. Analytics plays an important role in making sense of that data and exploiting its value. But learning from big data has become a significant challenge and requires development of new types of algorithms. The back-propagation neural network is quite suitable to express the nonlinear relation between earthquake and various anomalies. This paper presents a new approach which can work efficiently with the neural networks on large data sets. Since an Artificial Neural Network is a powerful modeling method, it has been widely used in the earthquake forecasting code. Artificial neural systems are present day machines that have great potential to improve the quality of our life. Advances have been made in applying such systems for problems found difficult for computation.

Keywords : Earthquake, Big Data, Artificial Neural Networks, Seismic systems and Earth electric field

I. INTRODUCTION

Earth Quakes are common marvels that happen with impact of a considerable measure of parameters, for example, seismic movement, changing in the ground waters' movement, changing in the water's temperature, and so forth. The impact of the event is traumatic because it affects large area to occur suddenly and unpredictably. Earthquake can cause large-scale loss of life, property and violates the basic services such as water, sanitation, energy, communication, transportation and so on. Earthquakes not only destroy cities and villages, but the effects lead to the destabilization of the economic and social fabric of the nation. The impact of the event is traumatic because it affects a large area to occur suddenly and unpredictably. Over the years, earthquake prediction has been a controversial subject that has challenged even the brightest researchers.

Artificial neural network system is a versatile nonlinear, dynamic framework. It can remove the causal relationship which is certain in the specimens by taking in a substantial number of tests and breaking down the more convoluted and nonlinear framework . So, it is very effective to apply the neural network to the earthquake prediction. Use of ANN to predict the earthquake is order of the day and a few literatures shows successful implementation of predicting earthquakes.[1] ANN is considered to be a form of AI which has high ability to model non-linear relationships.

II. CAUSES OF EARTH QUAKE

The earth has four noteworthy layers: the inward center (inner core), external center (outer core), and mantle and outside (crust). Reasons for Earthquake shown in fig.1

The crust and the top of the mantle make up a thin skin on the surface of our planet. But this skin is not all in one piece – it is made up of many pieces like a puzzle covering the surface of the earth. Not only that, but these puzzle pieces keep slowly moving around, sliding past one another and bumping into each other. We call these puzzle pieces tectonic plates, and the edges of the plates are called the plate boundaries. The plate boundaries are made up of many faults, and most of the earthquakes around the world occur on these faults. Since the edges of the plates are rough, they get stuck while the rest of the plate keeps moving. Finally, when the plate has moved far enough, the edges unstuck on one of the faults and there is an earthquake. While the edges of faults are stuck together, and the rest of the block is Ease of Use moving, the energy that would normally cause the blocks to slide past one another is being stored up. When the force of the moving blocks finally overcomes the friction of the jagged edges of the fault and it unstuck, all that stored up energy is released. The energy radiates outward from the fault in all directions in the form of seismic waves like ripples on a pond. The seismic waves shake the earth as they move through it, and when the waves reach the earth's surface, they shake the ground and anything on it. The size of the earthquake is called its magnitude. There is one magnitude for each earthquake. Scientists also talk about the intensity of shaking from an earthquake, and this varies depending on where you are during the earthquake.

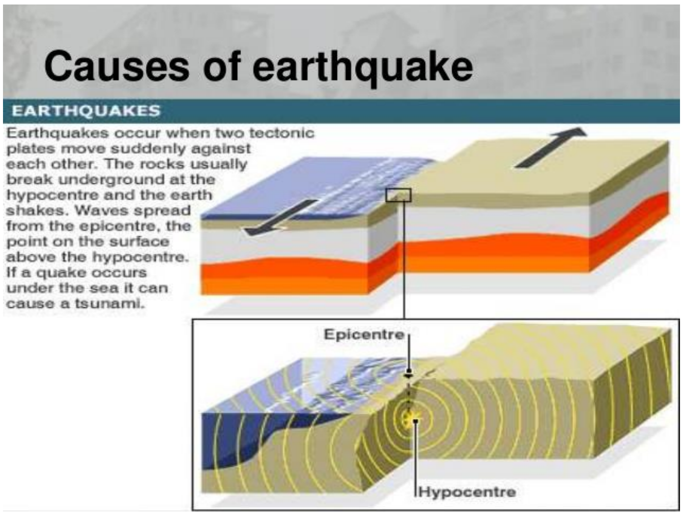

Figure 1 Reasons for Earthquake

III. REVIEW OF LITERATURE

For the Past 20 Years many Researchers have been working in this field. They were investigated various Methods for Prediction of Earthquake. For Example

- Mutual Information based Feature selection for Bagging
- Wrappers for Feature Subset Selection to find earth quake
- Toward Integrating Feature Selection Algorithms for classification and clustering
- Parallel, Distributed Processing, Explorations in the Microstructure of Cognition
- Recurrent Model and Robust Time – series prediction

Big Data

Big Data refers to massive amount of structured, unstructured and semi-structured data generated by applications Big data is a broad term for data sets so large or complex that traditional data processing applications are inadequate. Challenges include analysis, capture, data curation, search, sharing, storage, transfer, visualization, and information privacy Big data analytics over the repositories in the applications has recently received a great deal of attention from the research communities. Big Data' is used for collections of large datasets that include huge volume, high velocity, and a variety of data that is increasing day by day. Using traditional data management systems, it is difficult to process Big Data. Volume, Velocity, Variety is the key elements in big data. Volume", refers to the quantity of data. Velocity: timeliness or rate at which the data arrives and time in which it must be acted upon. Variety: heterogeneity of data types, data formats, semantic interpretation

Artificial Neural Network

An artificial neural network (ANN) is made out of interconnected counterfeit neurons that copy a few properties of natural neurons. Fig. 2 Shows model of an ANN. The neural network system is a very self adaptive non – linear dynamics. The network can extract the portion of hide samples through learning sample mass and can analyze a complex non linear system.

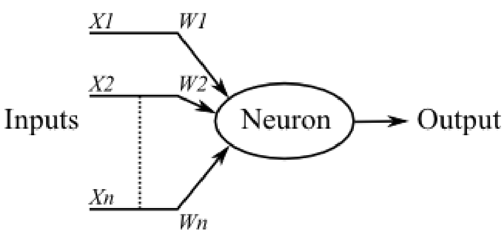

Figure 2 Model of an Artificial Neural Network

Seismic Networks

Seismograms prove to be useful for finding quakes as well and having the capacity to see the P wave and the S wave is critical. You figured out how P and S wave every shake the ground in various courses as they go through it. P waves are also faster than S waves, and this fact is what allows us to tell where an earthquake was. To understand how this works, let's compare P and S waves to lightning and thunder. Light travels faster than sound, so during a thunderstorm you will first see the lightning and then you will hear the thunder. If you are close to the lightning, thunder will boom right after the lightning, but if you are far away from the lightning, you can count several seconds before you hear the thunder. The further you are from the storm, the longer it will take between the lightning and the thunder. P waves are like the lightning, and S waves are like the thunder. The P waves travel faster and shake the ground where you are first. Then the S waves follow and shake the ground also. If you are close to the earthquake, the P and S wave will come one right after the other, but if you are far away, there will be more time between the two. By looking at the amount of time between the P and S wave on a seismogram recorded on a seismograph, scientists can tell how far away the earthquake was from that location.

DEMETER satellite

DEMETER satellite information have been utilized to concentrate on the plasma parameters variety amid the Haiti earthquake. One day (11 January 2010) preceding the earthquake there is a huge improvement of electron thickness and electron temperature close to the epicenter. Decrease of electron temperature is observed few days after temperature near the epicenter. Decrease of electron temperature is observed few days after the earthquake. Anomalous plasma parameter variations are detected both in day and night times before the quake. Launch of the DEMETER (Detection of Electromagnetic Emissions Transmitted from Earthquake Regions) satellite on 29 June 2004 has provided a good opportunity to study the ionosphere perturbations associated with earthquakes.

Related Work

Ashif pannakkat and Hojjat Adeli presents a method using ANN to predict magnitude of the earthquake. They have taken seismicity indicators inputs for the ANN model. They did not take previous magnitude for the model.[1]

C.R.Arjun and Ashok kumar propose a method for the estimation of peak ground acceleration using ANN by taking inputs like magnitude, hypo-central distance and average shear wave velocity.[2]

Alarifi.A.S. and Alarifi N.S. proposed a method using ANN to forecast earthquake in northern red sea area. They presented different statistical methods and data fitting such as linear ,quadratic and cubic regression.[3]

By the study of the Earth's electric field, Electric pre-seismic signals had observed that were generated some time during large Earthquakes [EQs]. The observed preseismic electric signals had attributed to various physical generating mechanisms while the same signals were represented by different type of "electric signature".

After doing analysis in Athens (ATH), Pyrgos (PYR) and Hios Island (HIO) in Greece. The specific data set spans for more than eight (8) years for the ATH and PYR monitoring sites and for almost four (4) years for the HIO case.

Fangzhou Xu, in 2010 used Data Mining technology overthe world for years from 2007 to 2008 by taking the way ofThree-Layer Back-Propagation Neural Network onDEMETER data, a series of physical quantities measured by the DEMETER satellite including Electron density, Electron temperature, ions temperature and oxygen ion density,together with seismic belt information in the range of a 30kmx 30km region around the epicenter for analysis. The timespan is about 30 days before the earthquake and he gets a totalaccuracy is about 69.96% [4]

In 2013 North Taiwan searched by Jui-Pin Wang using First-Order In 2013, the Qeshm Island, south of Iran searched by Adel Moatti using the clustering method has been performed to pattern recognition from the years of 1995 to2012 for magnitude greater than 6.0.The b-value of Gutenberg Richter law has been considered as precursor to earthquake prediction decrease as large earthquakes precursor [5].

Morales-Esteban used Feed Forward and Recurrent neural networks in the Iberian Peninsula area by The database of the Spanish Geographical Institute concluding that The networks predict the occurrence of large earthquakes for a seven-day [6].

In 2014, southwest areas of China searched by Feiyan Zhou using Back-Propagation (BP) Neural Network based on the Leven berg–Marquardt (LM) Algorithm, a nonlinear optimization method between Newton's method and gradient descent method For the overly parameterized problems has been performed for magnitude greater than or equal to three within half a year and The result is convergence speed of LM algorithm is fast and it has a good predictive effect and high accuracy [7].

IV. METHODOLOGY

One of the most popular NN algorithms is back propagation algorithm BP algorithm could be broken down to four main steps. After choosing the weights of the network randomly, the back propagation algorithm is used to compute the necessary corrections. The algorithm can be decomposed in the following four steps:

i) Feed-forward computation
ii) Back propagation to the output layer
iii) Back propagation to the hidden layer
iv) Weight updates

I'll just look at the connection between the neuron in the output layer and in the hidden layer. Fig3. Connection learning in a Back Propagation Network.

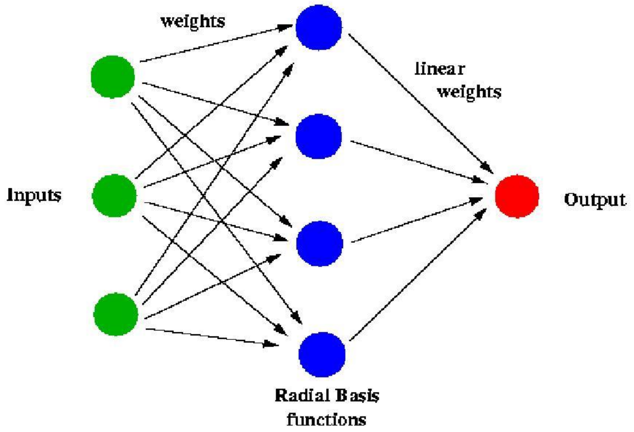

Figure 3 Basic Structure of BP neural network

Accept that we isolate the first preparing set (N) into n sub information sets, in particular N_1 to Nn. Each similarly isolated set will contain X/n occasions and every sub set will be prepared by the same BP system structure with a solitary concealed layer.

V. PROPOSED METHOD

This research will focus on the analysis of huge data by applying more neural networks to learn several sub dataset Our investigation is gone for enhancing the first preparing calculation to have the capacity to prepare an expansive dataset by isolated similarly sub datasets. The thought behind the proposed technique is that the littler sub information set will expend less preparing time than a major one. However the mistake from every sub information set is temperamental what not weights are isolated. Thus, we will gather the weight from every hub in the least mistake system and use these weights to supplant the prepared weight of different systems with the same structure. In this paper we describes a method which has two stages, as a Initial stage we have to take seismic network signal ,Electric pre-seismic signals and average magnitude of previous earthquakes, if any recorded in the past datas and in the final stage, those datas collected in the Preliminary stage are given to the ANN as inputs. ANN has to be trained with earthquake knowledge representation and employs non linear and back propagation algorithm to produce the precise prediction.

Initial Stage

Seismometers work using the principle of inertia. The seismometer body rests securely on the sea floor. Inside, a heavy mass hangs on a spring between two magnets. When the earth moves, so does the seismometer and its magnets, but the mass briefly stays where it is. As the mass oscillates through the magnetic field it produces an electrical current which the instrument measures.

Statistical prediction technique is nothing but using the historical data of the particular area, on the basis of data, we have to place seismometers at a particular distance apart from one another and approximately 10 km from the ground level, and we have to place the ocean bottom seismometer. By making a Seismometer network we can receive data from land and ocean. If we synchronize the signal generated by the land and ocean seismometer we can predict the larger earthquake Electric preseismic signals generated by earth's electric field is also an important data in the prediction method.[8]

Final stage - Knowledge Representation to ANN

By elements of encounters of the seismic tremor expectation specialists and the quake cases [7], we took the accompanying organization to speak to the learning of earthquake forecasting:

IF R1N1¨T11¨T12
R2N2¨T21¨T22B/P
… …
RKNK¨TK1¨TK2B/P
THEN RT1T2CiMtM1M2Mm P
RKNK¨TK1¨TK2B/P
THEN RT1T2CiMtM1M2Mm P

Every rule is composed of several premises (condition) and one conclusion, and the premises must meet certain relationship in time order. There are 5 attributes in a premise, which are defined as follows:
RI..= Area or station where anomaly appears.
NI .= A certain precursory anomaly

¨Ti1.= Delay period. This anomaly must appear delaying Ti1 after the preceding anomaly or the first anomaly appears.
Ti2=. Effective period. After this anomaly ends it is still possible that earthquakes occur during ¨Ti2 period.

P/F= P represents that this anomaly must delay after the preceding anomaly, F represents that this anomaly must delay after the first anomaly. The conclusion contains the following 9 attributes:
R = Area of coming earthquake
T1= Parameter 1 for calculating origin time.
T2= Parameter 2 for calculating origin time.
Ci =. Parameter 3 for calculating origin time.
Mt = Method for calculating origin time. For instance, when Mt=4, the Ci represents the order number of the premise, T1 represents the onset time of the anomaly related to the premise, T2 represents the end time of this anomaly. During the T1— T2 period, the earth-quake will occur.
Mg1= Parameter 1 for calculating magnitude
Mg.= Parameter 2 for calculating magnitude
Mm=Method for calculating magnitude. For instance, When Mm=1 M1 is the minimum magnitude of the coming earthquake, M2 is the maximum magnitude of the coming earthquake.
P = Probability of Earthquake Occurrence (PEO) .
Ri may be defined as any area or any station in a rule. It means this rule can fit any-where. It is very easy to store and manipulate the knowledge represented in such format.
Figure 4 shows the . Architecture of proposed work.

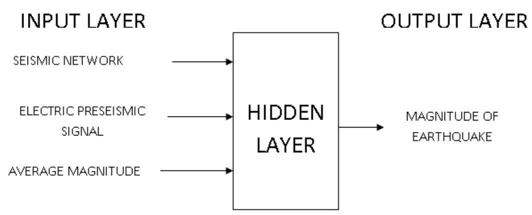

Figure 4. The Architecture of proposed work.

VI. CONCLUSION

This paper has proposed another technique which can apply Back propagation Neural Network to huge datasets. We can see from our analysis that weights of little dataset focalized speedier than the first dataset.

In any case, the weights prepared from sub dataset improved result when they were tried on the first data set.

Foretelling earthquake magnitude class by a neural network can be very effective.. The seismic network is faster and the accuracy also much higher. Each of the two networks can be realized by software, which is very simple and easy to perform. But some of the measures should be taken to further improve the accuracy of the neural network, such as learning of the neural network with specific data much more efficient data processing. Results showed that BP neural network is simple but effective too and the method of applying a certain certain value in the prediction of earthquakes in the future. This model can help in early evacuation according to future earthquake predications, as well as prevent causalities, loss of lives and economic losses that occur due to earthquakes happening worldwide every year

VII. REFERENCES

[1]. Ashif Pannakkat and Hojjat Adeli," Neural network models for earthquake magnitude prediction using multiple seismicity indicators", International journal of neural systems, Vol.17,Iss: 1,PP: 13 - 33.

[2]. C.R.Arjun and Ashok kumar, " Artificial neural network - based estimation of peak ground acceleration", ISET journal of Earthquake technology, Paper no.501,Vol.46,No.1 March 2009,PP:19 - 28.

[3]. Alarifi A.S. and Alarifi N.S.,"Earthquake magnitude prediction using Artificial neural network in northern red sea area", American Geophysical union, fall meeting 2009.

[4]. Fangzhou Xu, Xianfeng Song" Neural Network Model for EarthquakePrediction using DMETER Data and Seismic Belt Information" SecondWRI Global Congress on Intelligent Systems, 2010

[5]. Adel Moatti, Mohammad Reza Amin-Nasseri" Pattern Recognition onSeismic Data for Earthquake Prediction Purpose" InternationalConference on Environment, Energy, Ecosystems and Development,2013

[6]. A. Morales-Esteban, F. Martínez Álvarez"Earthquakeprediction inseismogenic areas of the Iberian Peninsula based on computationalintelligence" A. Morales-Esteban et al. / Tectonophysics 593

[7]. Feiyan Zhou, Xiaofeng Zhu "Earthquake Prediction Based on LM-BPNeural Network" Proceedings of the 9th International Symposium on Linear Drives for Industry Applications, Volume 1, 2009

[8]. Guang-yu Geng, Chuang-hui Li" Research on Seismo-IonosphericAnomalies Using Artificial Neural Network" IEEE,2010.

[9]. HUANG Sheng-Zhong" The prediction of the earthquake based onneutral networks", International Conference on Computer Design andApplications (ICCDA), 2010.

[10]. Habib Shah, Rozaida Ghazali, and Nazri Mohd Nawi "Using ArtificialBee Colony Algorithm for MLP Training on Earthquake Time SeriesData Prediction", Journal of Computing, 2011.

[11]. K. Tomiyasu "lunar, solar and earthquake projected positions of 138mag. 8.25-5.2 events in california from 1769 to 2004" IEEE,2012.

[12]. J. Reyes, A. Morales-Esteban" Neural networks to predict earthquakes iChile" Reyes et al. / Applied Soft Computing 13, 1314-1328, 2013.

[13]. Zhuowei Hu, Lai Wei "Spatial Prediction of Earthquake-InduceSecondary Landslide Disaster in Beichuan County Based on

GIS"Research Journal of Applied Sciences, Engineering and Technology 6(20): 3828-3837, 2013.

[14]. S. Niksarlioglu, F. Kulahci "An Artificial Neural Network Model for Earthquake Prediction and Relations between Environmental Parametersand Earthquakes" World Academy of Science, Engineering andTechnology, 2013.

[15]. Thanassoulas et.al., "The Earth's oscillating electric field (T = 1 day) in relation to the occurrence time of large EQs (Ms ≥ 5.0R). A postulated theoretical physical working model and its statistical validation."Cornell university library, pages:10

[16]. Z.Shengkai, Wang Chenjmin and Ma li,"Application of Artificial Intelligence in earthquake forecasting", http://ftp.esscc.uq.edu.au, PP:477 - 481.

[17]. Mark A.Kramer, " Non linear principal component analysis using auto associative neural networks by Laboratory for intelligent systems in process engineering", AIChe Journal, Feb 1991, Vol.37,No.2.PP:233-243.

[18]. The Earth's oscillating electric field (T = 1 day) in relation to the occurrence time of large EQs (Ms ≥ 5.0R). A postulated theoretical physical working model and its statistical validation.by Mr.Thanassoulas, Mr.C., Klentos, Mr.V., Verveniotis, Mr. G., Zymaris, N.

[19]. K. Shibata and Y. Ikeda, Effect of number of hidden neurons on learning in large-scale layered neural networks,ICROS-SICE International Joint Conference, 2009, p50085013.

[20]. K. Kaikhah and S. Doddameti, Discovering Trends in Large Datasets Using Neural Networks, Applied Intelligence, Springer Science + Business Media, Inc.,

[21]. Netherlands, vol. 24, 2006, pp. 51-60. [4] B. Liang and J. Austin, A neural network for mining large volumes of time series data, IEEE Transactions on Neural Networks, 2005, pp.688-693.

[22]. D. Xia, F. Wu, X. Zhang, and Y. Zhuang, Local and global approaches of affinity propagation clustering for large scale data, Journal of Zhejiang University SCIENCE, 2008, p1373-1381.

[23]. J. Heaton, Introduction to Neural Networks for C#, Second Edition, Heaton Research, Inc., Chesterfield, St. Louis, United States, Second Edition, 2008, pp. 137-164.