International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

March 2014 | Issue 18 | Volume 6 | Number 1 | Pages 18-26

CLASSIFICATION OF DATABASES AND METHODS FOR SEISMIC DATA ANALYSIS AND EARTHQUAKE PREDICTION

N.S. Soleimani Zakeri S. Pashazadeh

Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran
n.zakeri90@ms.tabrizu.ac.ir, pashazadeh@tabrizu.ac.ir

Abstract- Earthquake is one the most important disasters in the world. In order to save lives and building substructures of countries, more research in this field should be carried out as a matter of severity. Computer modeling and different artificial intelligence algorithms are known as applicable tools for the earthquake hazards prediction and prevention. This article tries to review the recent studies that have been conducted in this field. For this purpose, the literature methods have been classified into three categories including machine learning, data mining, and seismic feature extraction methods. The machine learning methods are also divided into several subcategories such as Artificial Neural Networks (ANNs), fuzzy systems and Support Vector Machines (SVMs) methods. The similar condition goes with the data mining methods in categorization. Moreover, the seismic feature extraction methods explain the important features used by aforementioned methods. Most of the recent researches are related to the prediction issues (e.g., the ultimate goal of data mining is for predicting the location of earthquake). Furthermore, the clustering methods can help us to predict the high risk areas. Although the problem of earthquake prediction and the related issues have not been completely solved in the world, researchers have tried to reduce the prediction error to provide predictions that are more accurate.

Keywords: Earthquake, Seismic Features, Artificial Intelligence Systems, Hazard Prevention.

I. INTRODUCTION

The earthquake is a one of the most important natural hazards that affects the metropolises, the buildings, the facilities within it, and the life of the people in the regions of interest [1]. Therefore, studying and analysis in predicting the earthquake is important for our life security as it is important to save the world from the damages in irretrievable financial, human life and physical resources. Two crusts of earth moving in opposite directions form locking regions. The earth energy has been gathered in these regions for a long time and when it is greater than a threshold, this energy appears in the form of a large earthquake [1].

Due to the uncertainty of the occurrence of the earthquake, finding meaningful relations between the events is important to save the human life [1]. Nowadays, using of the computer computation and intelligent system methods has already been increased. Using these methods to extract the relationships between different earthquakes at different times and locations can help us to know this phenomenon more to make a prediction of earthquake's happening. The progresses in seismology have provided us within valuable knowledge of earthquakes.

Both the modern equipment and the improvements in the earthquake engineering have helped us to gain the valuable information. On the other hand, the artificial intelligence systems require large suitable amount of data for designing such systems [2]. Pattern recognition methods that are used to analyze the earthquake data are consisting of algorithms and tools based on statistical information or a pre knowledge based data [3]. There are different articles in the literature, which have been proposed on this issue using data collected from the earthquake prone areas all over the world.

In this paper, a classification of recent studies has been presented and the methods and the parameters the features have been compared with each other. It is possible to classify them in terms of methods and objectives of the corresponding method. Below, a classification of the methods is presented.

II. MATERIALS AND METHODS

A. Earthquake Prediction Process

In this part, the process of reviewing the earthquake analysis using different algorithms is presented, which are classified based on their features and methods. Panakkat and Adeli [4] have presented recent researches about the prediction of the time, location, and magnitude of earthquakes. The difference between the proposed methods is the type of methods used for prediction process including the theoretical geophysics, biology, statistics, mathematics, and computer modeling.

The issues that have been discussed are about the physical and chemical events in the earth's crust that can be perceived by animal [4]. Also, it is pointed out that according to several historical catalogs it has been

observed that when there isn't any normal seismic activity for a long time in a specific region, it is more likely to experience a major earthquake [4]. Another review paper in this issue is presented by Otari and Kulkarni [5] about the applications of data mining in earthquake prediction. This method divides the earthquake prediction system into forecasting the month or the year and the short-term prediction based on the hours or the days.

In the geological science, the historical earthquakes in the region and the fault characteristics usually determine forecasting the earthquake in a specific region. This article divides the recent methods into two categories, neural network and data mining methods [5]. The examples of data mining are including the statistic regression method, three point method, sequential classification rule method, non-hierarchical clustering and multi-dimensional scaling, fuzzy-based methods, deterministic and non-deterministic optimized algorithms [5].

B. The Seismic Features

This part presents classifying different articles by their features for analyzing the earthquake and explaining each feature. As many seismic features which have been used for clustering tasks with redundancy, it can produce a useless higher dimensional feature space. Therefore, eliminating the useless features is a pre-processing step in each classification problem. Although different tools are used to find unavoidable features, however these methods may lose some vital features.

Therefore, this paper presents a feature extraction method based on RDA criterion. For this purpose, it uses forward and backward algorithms to extract the rankings related to the features. When estimating the covariance matrix is hard, the covariance matrix is regularized. All the subsets should be examined to find the optimal solution, because it is not guaranteed that these algorithms will yield the optimal solution. Various experiments show that increasing the size of neural networks doesn't improve the quality of the response, so in order to decrease the complexity it is better to use small model networks [6].

There are two papers that use data spaces consisting of data vectors f_i (i = 1,2,...,N) and the feature space (i.e., abstract space that is obtained from a nonlinear transformation) F_j (j = 1,2,...,M) [7, 8]. A single feature in feature space is F_j = [NS_j, NL_j, CD_j, SR_j, AZ_j, TI_j, MR_j], where *NS* and *NL* are the degrees of non-randomness at short and long distances respectively. *CD* is the spatial correlation dimension, *SR* is the degree of spatial repetitiveness, *AZ* is the average depth, *TI* is the time interval for the occurrence of a constant number of events and *MR* is the ratio of values of two events falling into two different magnitude ranges [7, 8].

III. MACHINE LEARNING TECHNIQUES (ANN, SVM AND FUZZY SYSTEMS)

A. Artificial Neural Networks

Artificial neural networks as important prediction tools have been used in different branches of science. Sometimes this tool is combined by wavelet transformation and genetic algorithms to improve the quality of the prediction [34]. So far, Different types of the neural network models have been used for earthquake prediction such as feed-forward back propagation neural network, Recurrent Neural Network (RNN) [10], Radial Basis Function (RBF) [10] neural network, Probabilistic Neural Network (PNN) [11] and so on. In this part, these network models, their parameters and their obtained results are introduced.

The authors proposed a recurrent neural network to predict the time and location of the earthquakes working on the training and testing data belonging to the Southern California and the San Francisco Bay region [10]. The input vector of the neural network in this paper [10], is a vector of eight seismicity indicators. The precision of the prediction by this model were evaluated by four statistical measures named the Probability Of the Detection (POD), the False Alarm Ratio (FAR), the frequency bias and the true skill score or *R* score and two different methods are investigated [10].

Firstly, the seismic region is partitioned into small regions and the earthquake data are divided into several time intervals. Then, the indicators are calculated for each small region obtained from partitioning for each time interval. Then the relation of the obtained results and the largest event of that region in the following time interval have been investigated. In the second one, the time intervals are not equal and each time interval is between two large earthquakes. Then a recurrent neural network is applied on the results. After numerous examinations, the network with two hidden layers and 10 nodes was known to be the best one.

According to the location prediction by this network, the error of the result was about 15-39 miles, which is suitable for emergency handling. As a result, the recurrent neural network is known to be suitable for the large earthquake rather than small or moderate ones [9]. Adeli and Panakkat proposed another neural network for this application, namely Probabilistic Neural Network (PNN) [11] in which the same indicators are also used as the input vector of the neural network model and the same measurements were calculated to evaluate the accuracy of the network except the frequency bias.

The PNN predominantly was used for classification problems. Being five times faster than the back propagation neural network, classification robustness in noisy data and the ability of adding future data are the excellence of this network over the common back propagation neural network. The prediction problem is considered as a classification problem and the magnitude is predicted as one of the classes defined before using the Parzen windows [11].

Unlike other neural networks, PNN does not use any training rule to calculate the weights so escapes from complex calculations. However, in this neural network the output must be one of the defined classes. Two hidden layers namely the pattern and summation layer are considered for this network [11] which results in an approach with good predictions for the earthquakes prediction problem with moderate magnitudes that the

recurrent network couldn't predict them with good accuracy. To some extent, this method completes the recurrent neural network for predicting the earthquake parameters [11].

Other authors introduced a seismic alert system using artificial neural networks which used the CIRES (Centro de Instrumentacion y Registro Seismologico) database containing 2500 historical earthquakes in Mexico City [12] by using both the back propagation and the genetic algorithms. This method aims to learn a relation between two regions, the desired, and the most seismic region, by firstly considering the weights of the network as the individuals of the genetic algorithm and aiming to develop the weights to find the best solution as the weights of the artificial neural network.

In this article, 70% of the data is used to train the network and 30% of data is used for testing the neural network model. Finally, this method could successfully recognize pattern of desired region by using the Artificial Neural Networks (ANNs) and Genetic Algorithms (GAs) [12]. In another paper [13], the collected data from earthquakes in northern Red Sea area is used for prediction. The method used here is a feed forward artificial neural network with multi-hidden layers [13].

The statistical methods and other methods just extract the linear relation between data, but applying neural network helps to extract also the nonlinear relation between data [13]. The first step is collecting data from valid sources to produce valid outputs. The second step is applying preprocessing and omitting the noise and repetitive data. At the third step, the features, which are needed to be used in neural network are extracted. The last step is to construct the neural network model and eventually to train and test it in order to predict the magnitude of the future earthquake.

The features used in the neural network are the earthquake sequence number, the occurrence location, the magnitude value, and the depth parameter. The reason of using the sequence number instead of the date and time is due to the experiences that show that it is easier for neural networks to learn the sequence. Finally, the method was evaluated and the authors have reached to the conclusion that this method is at least 32% better than the other proposed methods [13].

Another paper have used neural network for earthquake prediction by using the data used provided from the seismic catalogue of China [14]. The preprocessing stage done in this study was to eliminate the influence of the aftershocks by deleting the samples with magnitude less than 0.5. The important parameters showing the seismicity are b, η, $A(b)$, Mf, C, D and so on [14]. Two sets of the above features, one related to the prior interval and the other related to the current interval are regarded as the inputs for the neural network.

W_0 by describing the seismicity increase is used to determine status of the region [14]. When this parameter is about unity, it shows that this region is an unusual region for future moderate or strong earthquakes. If W_0 is about zero, this region is regarded as a normal region. The threshold used in this paper is $W_0 \geq 0.7$. If this condition is true and being anomalous lasts for more than one year and this area is greater than four square degrees, it will be considered as an unusual region. In addition, if there is any moderate or strong earthquake in two years after considering a region as an unusual region, this is a true prediction. For this research, the success ratio in predicting earthquake was about ninety four percent [14].

B. Support Vector Machines

Another machine learning method based on the statistical learning theory is Support Vector Machines (SVM) in which the generalization is higher than artificial neural networks. This superiority and the ability to solve certain problems indicate that SVM is a robust method among machine learning methods. The SVM method was employed as a prediction method to study the collected data from the Historical Strong Earthquake Catalogue in China (Department of Seismic Hazard Prevention of State Seismological Bureau, 1995). In addition, Current day Earthquake Catalogue in China (Department of Seismic Hazard Prevention of State Seismological Bureau, 1999), and the Global Strong Earthquake Catalogue compiled by the Center for Analysis and Prediction, CEA [15].

The Back Propagation (BP) neural network was used over seventy-five strong earthquake samples occurred in the period of 1925-1999. From 75 samples, 65 samples were used for learning and 10 samples for testing the network. Then in order to make a comparison with SVM method, they used the same data plus the data in the period 2000 to 2003. Totally 14 samples were used for testing process and based on the reported results, the rate of correct prediction by BP neural network was 0.8 and the rate of correct prediction by SVM was 0.86 [15]

C. Fuzzy Systems

Fuzzy logic as an applicable tool is used by different case studies. As an example, its combination with clustering is used to recognize not allowed nodes in a network [35]. There is another paper that presents a fuzzy expert system for predicting earthquake by using earthquake data around Zagros range [2]. First, it used an expert's knowledge to construct a fuzzy rule base. By sending these rules to inference engine, a Fuzzy Inference System (FIS) is produced which is used to derive results.

The prediction process is based on the coupled earthquakes, two earthquakes that are close together both in time and space [2]. Based on the proposed idea, some of the experts in this field believe that a big or moderate earthquake may fool the coupled earthquakes. The first step is to covert the crisp variables to fuzzy variables. Then by using the Sugeno engine, it analyzes the fuzzy field inputs (converted fuzzy variables), and derives the output. In order to obtain a high performance in this system, the parameters provided by the expert must be error free [2].

IV. DATA MINING TECHNIQUES

A. General Definition of Data Mining and its Variants

Data mining is a technique for extracting the information, relevant data, and hidden facts in large

databases to find patterns and relationships. The various data mining methods are statics, clustering, visualization, association, classification, prediction, trend, evolution analysis and outlier analysis [16]. It also adds another categorization of knowledge discovery in large databases which is divided into event based mining and relation based mining [16].

Another paper indicates artificial intelligence and statistical analysis as a foundation of the data mining which lead to two models, descriptive and predictive models [1]. Clustering is the process of dividing the big data into smaller groups. Each obtained group is called a cluster that comprises of a hidden pattern used to extract knowledge from that pattern. Clustering algorithms are divided into two groups including the supervised one in which the clusters are predefined and the algorithm just impute data to clusters and the unsupervised algorithm in which the clusters are not defined before [17].

Also, a clustering algorithm uses particular error functions that minimize the distances within a cluster, however it is not always obvious that which clustering algorithm is the best one [18]. However, there is a general procedure for clustering, which firstly selects an analysis window or in other words the features on which the clustering is based, secondly, it selects the clustering algorithm, which seems to be the most appropriate one and then it assesses the used algorithm based on the appropriate criteria. The forth and last step is to analyze the extracted clusters to see if they are compatible with the correct cluster structures.

In order to extract important information by clustering we need to select the most suited method for feature production and clustering techniques [19]. There is a categorization of clustering methods that may overlap in some cases [20]:

- Hierarchical Methods: In the Hierarchical clustering [20], the clusters are created as the algorithm calculation proceeds. These methods are classified into two groups. Agglomerative methods, which use bottom-up approaches, starting with a single sample in a cluster and iteratively merge the clusters until the appropriate number of clusters are satisfied. Divisive ones which use top-down approaches starting with one cluster of all the samples and then divide this cluster to smaller ones until the criteria is satisfied [20].
- Partitioning Methods: These methods split data to subsets [20]. These methods are classified into five subgroups. The first one uses relocation and iteratively optimizes the clusters. Unlike the hierarchical methods in which the clusters are not examined after construction, this method iteratively improves the clusters. The second one is called probabilistic method that uses probabilistic parameters. The third and the forth ones are k-medoids and k-means which calculate the distances between clusters as the similarity of the clusters mutually in each iteration. The difference between these two methods is the point by which a cluster is presented.

In k-medoids, one of its points shows each cluster however in k-means each cluster is presented by the mean value of the points in that cluster. The last one is also known as the density based partitioning method. However, the first one is based on density and connectivity. In addition, both of these measures are calculated regarding the local distribution of the nearest neighbors. The second one uses a density function for clustering [20].

- Grid Based: Here, the difference with data partitioning is focusing on space instead of data [20].

Other categories of the clustering methods are listed as below [20]:

- Methods based on co-occurrence of categorical data
- Constraint-based (Gradient descent and Evolutionary)
- Scalable clustering
- Algorithms for high dimensional (Sub space, Projection, Co-clustering).

There are four main objective of doing clustering over different data [19]:

1- Data reduction which means, dividing the main data into smaller groups. Therefore, we can represent a cluster by just one representative data or feature.

2- Hypothesis generation uses cluster analysis to extract a theory from the nature of the data.

3- Hypothesis testing, which means validating the accuracy of the existing hypothesis.

4- Prediction which means using the obtained groups and their related features to do some predictions.

B. Different Approaches on Data Mining for Earthquake Data

There is a survey article about clustering time series data that refers to the earthquake data clustering [21], in which earthquake data clustering by a multivariable method is performed. This multivariable method uses agglomerative hierarchical clustering algorithm and *J* divergence and symmetric Chernoff information divergence as a distance measure. This article also stated that in order to identify the start times of the time series correctly, the time series clustering could be complete with a change-point detection algorithm [21].

A physical and stochastic model has been presented for earthquake clustering, which investigate the microscopic properties of the earth rather than macroscopic ones [22]. For this purpose, it used two universal and applicable laws, the Gutenberg-Richter (G-R) law, and the Omori law. According to epidemic model, each earthquake has an increasing effect on the future earthquakes. The proposed model was applied on Japanese seismicity data from 1970 to 2003. The authors concluded that the proposed model describes the observations with high likelihood [22].

A visual data mining methodology was proposed to compare different unsupervised clustering algorithms and to evaluate them [18]. While using unsupervised methods, the main goal is to perform clustering without using any prior information. Finally, the pattern in a cluster simulates its own points rather than points in other clusters. In order to compare the different algorithms it uses three types of data sets. In the first data set, clusters are well separated.

The clusters of the second one are intermediate separated and in the last one, clusters are approximately overlapped. The partition model as a hard competitive method, EM Strategy as a probabilistic model, the

agglomerative method, and the divisive method as a hierarchical model, neural gas algorithm, and SOM algorithm as a Soft Competitive Model were applied on these data sets. As result, it represents diagrams and offers self-organizing map algorithm as best of all methods [18].

Another paper in this concept introduced a web client server system named WEB-IS to analysis and visualize the seismic data around the world [19]. This system performs data mining as interactive system. In order to provide a prediction of the earthquakes and to facilitate observing of the similarities between them, we need easy and fast access. Therefore, the resources are located in a GRID framework using a distributed event brokering system named Narada Brokering.

Information exchange between the clients is accomplished by a java messaging system. Two main resources for data are used in this paper. Computer modeling which produces the synthetic data and seismic measurements that refers to the actual data. The authors believed that the failure in predicting the earthquake is because of the lack of communication of the investigators in this field and lack of quick and easy access to the data around the world. So they hope that this system would help a lot in this area in the future [19].

The Kohonen Self Organizing Maps (SOM) [23] were used to classify the seismic facies. The SOM is one of the most applicable methods and is assigned to the category of unsupervised pattern recognition methods. This method utilizes the wave forms as input data and groups seismic facies according to different seismic attributes such as amplitude, phase, frequency and etc. when using multi-attribute data as input, SOM yields better results [23].

In another study, two main methods including statistics and soft computing were used to make an effort for predicting the earthquake [24]. As we know, the earthquake data is following the time series methodology. This paper applies data mining methods on the collected data and then uses the fuzzy logic rules to predict the effects of the earthquake. The magnitude, the depth, and the impact are considered as the linguistic variables in this research. The feature extraction process extracts the linearly independent seismic parameters and the statistical properties of data [24].

In another paper, the authors used the waveforms and applied k-means and Gaussian Mixture on the waveforms [25]. There were also two other methods that were tested on the results, Linear Discriminant Function (LDF) and Quadratic Discriminant Function (QDF). The magnitude values of the waveforms were between 1.8 and 3.0. The results obtained by all of these four methods are acceptable however the accuracy of QDF is the most accurate one on the mentioned data [25]. The prediction in time series data mining refers to predicting an event along time series [26]. The algorithm proposed in this study uses fuzzy logic, which is applied on the synthetically produced earthquake time series.

This method is consisted of two steps. At first, it tries to embed the time delay and then nonlinearly converts the time series to the phase space. Then, the fuzzy logic by a Gaussian membership function is applied on the time series. So the optimal values of the important parameters are predicted. An advantage of this method is to omit the constraints of traditional time series analysis. The experimental results by true predictions indicate the success of this method [26].

The multi-resolution clustering methods in another paper are used to analyze numerous earthquake data which applies multivariate analysis by considering events with both low and large magnitude [27]. Four numerically simulated models achieved the synthetic data used in this research. These models are as below. Uniform properties (U), A Parkfield-type asperity (A), Fractal brittle properties (F), and Multi-size-heterogeneity fault zone (M). In order to extract abnormalities in two spaces including spatio-temporal and feature space, mutual nearest neighbor algorithm was used. The results show that the structures related to small events have a specific correlation with large earthquakes. according to this correlation there are clusters of small events before and after large earthquakes[27].

A pattern recognition method for the earthquake catalog was proposed using some models in which the used models have Uniform properties (U), A Parkfield-type asperity (A), Fractal brittle properties (F) and Multi-size scale heterogeneities (M) [28]. The analysis shows high complexity in the synthetic data however despite this, the local minimum or local maximums obtained by examining the parameters shows important association in large earthquakes. In some situations, distribution of the small events is used to statistically predict the large events. Some results are the same for all of the above models and some of them solely depend on the design of the model [28].

Another research proposes a new method for extracting association patterns which are frequently observed throughout multi-sequence data [29]. There are large amount of data for earthquake, which are an example of the multi-sequence data. This method can be used to discover the resemblance of patterns among the earthquake data [29]. The synthetic waveforms are used for performance assessment of the method. The proposed method includes two methods, the frequent pattern extraction from each sequence and the internal graph mining to extract association pattern.

In this method, the first step is to run preprocess for each sequence and normalize the data ranging from 0 to 1. The second step uses the maximum window width and a frequency threshold as input parameters to extract the frequent patterns (i.e., the intervals). Then, the interval graph is formed and finally the association patterns are extracted from the sets of intervals. In the experiment by using the precision and recall, the results that are calculated from below equations can be evaluated [29]:

$$precision = \frac{CDP}{DDP} \tag{1}$$

$$recall = \frac{CDP}{EDP} \tag{2}$$

where, *CDP* and *DDP* are precisely described in the related paper [29].

Another paper aimed to extract earthquake clustering patterns [30]. To achieve this purpose, transforming two dimensional superimposed Poisson process into one dimensional mixture density function is needed and *N*th order distance and a genetic algorithm to decompose mixture density function are employed. Finally, the background earthquakes and clustering earthquakes are separated from each other. The background earthquake is intended to the earthquakes which happen at initial stage of seismic activity period (i.e., the tranquility stage) [30].

The clustering earthquakes are those that are used to extract the clustering patterns. The research studies show that clustering earthquakes usually occur one to six month before the main quakes, so distinguishing them is beneficial. According to this paper, because of the noise effects utilizing the point density make it possible to extract the earthquakes clustering. The catalog data are provided from the Seismic Catalog of West China (1976-1979) $M \geq 1$ [30].

The other study applies the temporal clustering of the catalog of the Absheron-Prebalhkan region in the Caspian Sea [31]. The authors used both the historical and the instrumental catalogs in which the time was ranging from 1842 to 2012 and the magnitude is ranging from 2.5 to 6.8. The growing population and industry in the capital city of the Azerbaijan, Baku is depending on the status of the fault. Using the Gutenberg-Richter analysis obtains 4.0 as the completeness magnitude or threshold.

Therefore, the temporal clustering is applied on the events with magnitude $M \geq 4$. The methods, which are used for the temporal clustering, are Allan Factor and coefficient of variation. These events are divided into three sub-sequences. The analysis over the catalog exhibited the time-cauterization in the third sub-sequence (1995-2012) and a periodical behavior in second one (1949-1991) [31].

In another research, different methods and tools were applied to prevent seismic risks by predicting location, time or magnitude [1]. Designing a decision support system that is purpose of this paper exploits data mining techniques and uses a model to produce predictions. The system is composed of three parts, the Load Manager Process that uses SQL Server technology for its database server and Weka for data mining, the Multidimensional Schema that uses a bottom-up multidimensional modeling and finally that OLAP operation.

The descriptive and predictive models that are described in the above sections use k-Means for clustering and a priori algorithm for Association rules to complete the scheme of the descriptive model. The predictive model uses Classification Tress, Naïve Bayes, and Linear Regression. The assumption for this is that if an area has experienced several earthquakes in short intervals, it is also prone to large events. The main catalogs used are U.S. Geological Survey (USGS), National Seismic System of Mexico (SSN) and National Earthquake Information (NEIC) [1]. The average magnitude of the studied events is 6.16. The diagrams plotted in terms of hour and day which are just based on statistics demonstrate that most of the earthquakes have been occurred in about sixteenth day of a month and usually before 4 AM. [1].

The K-Nearest Neighbor Graph method was used as a supervised clustering method to recognize the seismically vulnerable regions to consider in the buildings infrastructures [17]. The advantage of this method to other classification algorithms is considering all the distributions of the training data points. The used data set in this research in taken from a digital map of Iasi city in the north eastern part of Romania [17].

Dzwinel and Yuen introduced their approach entitled 'nonlinear and multidimensional scaling and visualization of earthquakes clusters over data space and feature space', in 2003 and 2004 [7, 8]. According to their work, the Gutenberg-Richter power law earthquake size distribution shows that there are many small events around large earthquakes [7, 8]. The clustering method proposed here was also consisting of the feature extraction and the visualization methods. The features used in this research are described in the feature section (section 2-1) [7, 8].

The analysis is on both actual data set and synthetic data catalogs around the Japanese island in the period of 1997-2003 [7, 8]. The data catalogs are analyzed in both data space and feature space. In high resolution (data space), each event is a four dimensional vector $f = [m, z, x, t]$, where, m is the magnitude of the single event, z is the depth of the event, x is epicentral coordinate and t is the occurrence time of the event. An agglomerative algorithm applied on the data space produces clusters of similar events. Then, a nonlinear transformation is used to generate the feature space. In next step, a non-hierarchical clustering is done over the resulted feature space. Finally, it can be deduced that combining the result of the clustering in both spaces can give better results [7, 8].

C. Visualization of the Earthquake Data

One of the important issues in the earthquake data analysis is the visualization concept [32, 33]. Because of the multidimensional nature of the data, it is important to find the best way to visualize them. The better the visualization, the better the interpreting and extracting correct relation between data points [32, 33]. A paper proposed an estimation of the number of the earthquakes each year around the world [32]. These statics showed large numbers of earthquakes with magnitude 4.0 or more than that is considered as a risk in the major cities [32, 33].

Therefore, well understanding and visualization of the earthquake patterns is important to prevent risks in metropolises. In the above mentioned paper, visualization of different datasets are named Field-Measured seismic data sets, Numerical simulation seismic data sets, NEIC historical earthquake records [32, 33].

V. RESULTS AND DISCUSSION

The results of the mentioned literature methods have been concisely extracted from the published articles. These results are based on different machine learning methods used for prediction and their regional advantages and disadvantages.

In Table 1, various ANN and SVM models along with their outstanding dataset have been compared with each other. Although none of the datasets is the same, however

their obtained results are worthy regarding the region of interest and may be applicable only to those regions. Some of the methods have mentioned their point of failures but some of them have not mentioned any which seem they should have some disadvantages as well. The results shown in Table 1 indicate that to some extent these methods are complement and can be used together to achieve and overcome any existing short comes. However, a method that works well in every situation and in any area of interest is not accessible yet. Table 2 also classifies the data mining methods, shows the name of the used method, their advantages, and the data sets, which have been compared clearly.

Table 1. Classification of the intelligent system methods

Method	Data Set	Advantages	Disadvantages
Recurrent Neural Network [9]	Southern California & San-Fransisco Bay Region	Good results for large earthquakes	Not good results for small earthquakes
Probabilistic Neural Network [11]	Southern California & San-Fransisco Bay Region	Good results for moderate and small magnitude earthquakes (five time faster than BPNN) Classifier robustness in noisy data, ability of adding future data	Doesn't have good results for large magnitude earthquakes, the predicted magnitudes must be one of the defined classes
Alert System based on Back Propagation ANN and Genetic Algorithm [12]	CIRES (Centro de Instrumentacion y Registro seismologico) Mexico City	Recognizes a pattern in the desired region	None mentioned
Feed Forward Neural Network [13]	Northern Red Sea Area	Extract Nonlinear Relation between data, Using Sequence number instead of time and date	None mentioned
Neural Network to midterm prediction [14]	Seismic Catalog of China	Uses a Threshold to recognize usual and unusual regions	None mentioned
Support Vector Machine (SVM) [15]	Historical Strong Earthquake Catalogue in China and Current day Earthquake Catalogue in China	Generalization in this method is higher than artificial neural networks	None mentioned
A Fuzzy Expert System [2]	Earthquake Data Around Zagros Range	The used prediction is based on the coupled earthquakes	The Parameters provided by the expert must be error-free

Table 2. Classification of data mining methods

Method	Advantage of method	Data set
Clustering Time Series [21]	Uses Multi Variable Method and Agglomerative Hierarchical Clustering Algorithm	various places in Florida, Tennessee, and Cuba taken from the National Climatic Data Center
Physical And Stochastic Methods [22]	Describes the Model With Likelihood	Japanese Seismicity
Visual Data Mining [18]	By Comparing the Different Algorithms Concludes that the Self Organizing Map (SOM) is the best of all for this Purpose	None mentioned
Web Client-Server System Analysis of The Seismic Data [19]	Fast Access to Data All Over The World, Facilitating The Computing	Japan Meteorological Agency (JMA)
Kohonen Self Organizing Maps [23]	As a Suitable Algorithm When Using Multi Attribute Data	Dataset from Northern Gulf of Mexico, Synthetic Dataset
Data Mining Methods and Fuzzy Logic [24]	Using Linguistic Variables, Using Linearly Independent Properties	Indonesia earthquake data
K-means, Gaussian Mixture, LDF and QDF [25]	Conclude that The Accuracy of the QDF is Better Than The Others	Synthetically Generated Dataset
Multi Resolution Clustering [27]	The Small Events Structure Have Specific Correlation With Large Events	Synthetically Generated Dataset
Applying Pattern Recognition [28]	Extracts Important Association In Time, With Large Earthquakes	Synthetically Generated Dataset
Frequent Pattern Extraction [29]	Discovers The Similarity Among Earthquake Data	Synthetically Generated Dataset
Extract earthquake Clustering Pattern [30]	The Result Is not Effected By Noisy Data	Seismic Catalog of West China
Temporal Clustering of The Seismicity [31]	Gives a Special Property For Each Sub-Sequence	Catalog of The Absheron, Prebalkhan Region
Decision Support System [1]	Shows that The Concentration of The Earthquakes are Mostly in Sixteenth Day of The Month	USGS – SSN - NEIS
K-Nearest Neighbor Graph [17]	Considers All The Distributions of The Training Data	None mentioned
Nonlinear and Multi-Dimensional Scaling [7, 8]	Uses The Combination of Clustering in Two Spaces To Get Better Results	Japanese Island

VI. CONCLUSIONS

By considering the significance of studying the earthquake data as important natural phenomena, in this paper we review the literature sources about this important concept. Different methods in artificial intelligence and machine learning were used to analyze the earthquake data. Some of these methods include the artificial neural networks, data mining, clustering, fuzzy systems.

The spatial, temporal and the magnitude prediction are the important issues that are of interest to researchers. The spatial prediction is more concerned in data mining and clustering while the temporal and magnitude prediction in the most attractive issue in neural networks. Then, the advantages and the disadvantages of the methods, the used databases, and the used methods are specified and discussed in two tables.

The first table presents the machine learning methods and the second one presents the data mining methods. In conclusion, the presented tables and the above classifications in this paper are presented as a detailed comparison between the proposed literature methods and can be used by the interested researchers to study more in this field, which in result can be used in a hybrid form to achieve better and globally results.

REFERENCES

[1] M.J. Somodevilla, A.B. Priego, E. Castillo, I.H. Pineda, D. Vilariño, A. Nava, "Decision Support System for Seismic Risks", Journal of Computer Science & Technology, Vol. 12, No. 2, p. 71, 2012.

[2] A. Andalib, M. Zare, F. Atry, "A Fuzzy Expert System for Earthquake Prediction, Case Study - The Zagros Range", Third International Conference on Modeling, Simulation and Applied Optimization, Sharjah, U.A.E., 2009.

[3] D.A. Yuen, W. Dzwinel, Y. Ben-Zion, B.J. Kadlec, "Earthquake Clusters Over Multi-Dimensional Space, Visualization", Encyclopedia of Complexity and Systems Science, pp. 2347-2371, 2009.

[4] A. Panakkat, H. Adeli, "Recent Efforts in Earthquake Prediction (1990-2007)", Nat. Hazards Rev., Vol. 9, No. 2, pp. 70-80, 2008.

[5] G.V. Otari, R.V. Kulkarni, "A Review of Application of Data Mining in Earthquake Prediction", International Journal of Computer Science and Information Technologies, Vol. 3, No. 2, pp. 3570-3574, 2012.

[6] A. Javaherian, H. Hashemi Shahdani, "Seismic Attribute Redundancy Reduction Using Statistical Feature Extraction Technique", First International Petroleum Conference & Exhibition, Shiraz, Iran, 2009.

[7] W. Dzwinel, D.A. Yuen, K. Boryczko, Y. Ben-Zion, S. Yoshioka, T. Ito, "Cluster Analysis, Data-Mining, Multi-Dimensional Visualization of Earthquakes over Space, Time and Feature Space", Earth and Planetary Sci. Letters, pp. 1-11, 2003.

[8] W. Dzwinel, D.A. Yuen, K. Boryczko, Y. Ben-Zion, S. Yoshioka, T. Ito, "Nonlinear Multidimensional Scaling and Visualization of Earthquake Clusters Over Space, Time and Feature Space", Nonlinear Processes in Geophysics, Vol. 12, No. 1, pp. 117-128, 2005.

[9] A. Panakkat, H. Adeli, "Recurrent Neural Network for Approximate Earthquake Time and Location Prediction Using Multiple Seismicity Indicators", Computer-Aided Civil and Infrastructure Engineering, Vol. 24, No. 4, pp. 280-292, 2009.

[10] A. Panakkat, H. Adeli, "Neural Network Models for Earthquake Magnitude Prediction Using Multiple Seismicity Indicators", International Journal of Neural Systems, Vol. 17, No. 1, pp. 13-33, 2007.

[11] H. Adeli, A. Panakkat, "A Probabilistic Neural Network for Earthquake Magnitude Prediction", Neural Networks, Vol. 22, No. 7, pp. 1018-1024, 2009.

[12] C.M.A.G. Robles, R.A. Hernandez-Becerril, "Seismic Alert System Based on Artificial Neural Networks", World Academy of Science, Engineering and Technology, Vol. 66, pp. 813-818, 2012.

[13] A.S.N. Alarifi, N.S.N. Alarifi, S. Al-Humidan, "Earthquakes Magnitude Predication Using Artificial Neural Network in Northern Red Sea Area", Journal of Science, King Saud University, Vol. 24, pp. 301-313, 2012.

[14] W. Wang, G.F. WU, X.T. Song, "The Application of Neural Network to Comprehensive Prediction by Seismology Prediction Method", ACTA Seismologica Sinica, Vol. 13, No. 2, pp. 210-215, 2000.

[15] W. Wei, L. Yue, L. Guo-Zheng, W. Geng-Feng, M. Qin-Zhong, Z. Li-Fei, L. Ming-Zhou, "Support Vector Machine Method for Forecasting Future Strong Earthquakes in Chinese Mainland", ACTA Seismologica Sinica, Vol. 19, No. 1, pp. 30-38, 2006.

[16] K.R. Kumar, M. Sunferdeen, "Predicting Earthquakes through Data Mining", National Conference on Innovations in Communicative World, pp. 77-80, 2012.

[17] F. Leon, G.M. Atanasiu, "Data Mining Methods for GIS Analysis of Seismic Vulnerability", First International Conference on Software and Data Technologies. Vol. 2, pp. 153-156, 2006.

[18] I.D. Marroquín, J.J. Brault, B.S. Hart, "A Visual Data-Mining Methodology for Seismic-Facies Analysis, Part 1 - Testing and Comparison with other Unsupervised Clustering Methods", Geophysics, Vol. 74, No. 1, pp. 1-11, 2009.

[19] D.A. Yuen, B.J. Kadlec, E.F. Bollig, W. Dzwinel, Z.A. Garbow, C.R.S. Da-Silva, "Clustering and Visualization of Earthquake Data in a Grid Environment", Visual Geosciences, Vol. 10, No. 1, pp. 1-12, 2005.

[20] P. Berkhin, "Survey of Clustering Data Mining Techniques", Grouping Multidimensional Data, pp. 25-71, 2006.

[21] T.W. Liao, "Clustering of Time Series Data - A Survey", Pattern Recognition, Vol. 38, No. 11, pp. 1857-1874, 2005.

[22] R. Console, M. Murru, F. Catalli, "Physical and Stochastic Models of Earthquake Clustering", Tectonophysics, Vol. 417, No. 1-2, pp. 141-153, 2006.

[23] A. Roy, M. Matos, K.J. Marfurt, "Automatic Seismic Facies Classification with Kohonen Self Organizing Maps - A Tutorial", Geohorizons Journal of Society of Petroleum Geophysicists, pp. 6-14, 2010.

[24] G. Preethi, B. Santhi, "Study on Techniques of Earthquake Prediction", International Journal of Computer Applications, Vol. 29, No. 4, pp. 55-58, 2011.
[25] H.S. Kuyuk, E. Yildirim, E. Dogan, G.G. Horasan, "Application of k-Means and Gaussian Mixture Model for Classification of Seismic Activities in Istanbul", Nonlinear Processes in Geophysics, Vol. 19, No. 4, pp. 411-419, 2012.
[26] I. Aydin, M. Karakose, E. Akin, "The Prediction Algorithm Based on Fuzzy Logic Using Time Series Data Mining Method", World Academy of Science, Engineering & Technology, Vol. 51, p. 91, 2009.
[27] W. Dzwinel , D.Y.Y. Kaneko, K. Boryczko , Y. Ben-Zion, "Multi-Resolution Clustering Analysis and 3-D Visualization of Multitudinous Synthetic Earthquakes", Visual Geosciences, Vol. 8, No. 1, pp. 1-32, 2003.
[28] M. Encva, Y. Ben-Zion, "Application of Pattern Recognition Techniques to Earthquake Catalogs Generated by Model of Segmented Fault Systems in Three Dimensional Clastic Solids", Journal of Geophysical Research, Vol. 102, No. B11PP 24, pp. 513-528, 1997.
[29] T. Miura, Y. Okada, "Extraction of Frequent Association Patterns Co-Occurring Across Multi-Sequence Data", International Multi Conference of Engineers, IMECS, Vol. 1, pp. 452-455, Hong Kong, 14-16 March 2012.
[30] P. Tao, Z. Cheng-Hu, Y. Ming, L. Jian-Cheng, L. Quan-Lin, "The Algorithm of Decomposing Superimposed 2-D Poisson Processes and Its Application to The Extracting Earthquake Clustering Pattern", ACTA Seismologica Sinica, Vol. 7, No. 1, pp. 54-63, 2004.
[31] L. Telesca, G. Babayev, F. Kadirov, "Temporal Clustering of the Seismicity of the Absheron-Prebalkhan Region in the Caspian Sea Area", Nat. Hazards Earth Syst., Vol. 12, pp. 3279-3285, 2012.
[32] T.J. Hsieh, "Understanding Earthquakes with Advanced Visualization", ACM SIGGRAPH Computer Graphics - Visual Research, Evaluation and Assessment in the Age of Computer Graphics, Vol. 44, No. 1, pp. 1-13, 2010.
[33] G.H. Weber, M. Schneider, D.W. Wilson, H. Hagen, B. Hamann, B.L. Kutter, "Visualization of Experimental Earthquake Data", SPIE, Visualization and Data Analysis, Vol. 5009, pp. 268-278, 2003.
[34] H. Farsi, P. Etezadifar, "Video Quality Improvement Using Local Channel Encoder and Mixed Predictor by Wavelet, Neural Network and Genetic Algorithm", International Journal on Technical and Physical Problems of Engineering, Vol. 5, pp. 108-117, 2010.
[35] S.J. Dastgheib, H. Oulia, H. Gholamshiri, A. Ebrahimi, "A Proposed Fuzzy Method to Detect Faulty Nodes in Wireless Sensor Network Clustering Structure", International Journal on Technical and Physical Problems of Engineering. Vol. 4, pp. 71-75, 2012.

BIOGRAPHIES

Negar Sadat Soleimani Zakeri received her B.Sc. degree in Computer Software Engineering in 2009. Currently she is M.Sc. degree student of Computer Engineering, Artificial Intelligence in University of Tabriz, Tabriz, Iran Since 2011. Her research interests include artificial intelligence, artificial neural network, image processing, and seismic clustering analysis.

Saeid Pashazadeh received his B.Sc. degree in Computer Engineering from Sharif Technical University, Tehran, Iran in 1995. He obtained M.Sc. and Ph.D. degrees in Computer Engineering from Iran University of Science and Technology, Tehran, Iran in 1998 and 2010, respectively. Currently he is an Assistant Professor of Software Engineering and Chair of Information Technology Department at Faculty of Electrical and Computer Engineering in University of Tabriz, Tabriz, Iran. He is Member of IEEE and Senior Member of IACSIT and Member of Editorial Board of Journal of Electrical Engineering at University of Tabriz. He was lecturer in Faculty of Electrical Engineering in Sahand University of Technology, Tabriz, Iran from 1999 until 2004. His main interests are modeling and formal verification of distributed systems, computer security, wireless sensor/actor networks, and applications of artificial neural networks.