American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-5, Issue-4, pp-106-111
www.ajer.org
Research Paper Open Access

Seismic Hazard Prediction Using Seismic Bumps: A Data Mining Approach

Musa Peker
Department of Information Systems Engineering, MuglaSitkiKocman University, Mugla, Turkey

Abstract: *Due to the large number of influencing factors, it is difficult to predict the earthquake which is a natural disaster. Researchers are working intensively on earthquake prediction. Loss of life and property can be minimized with earthquake prediction. In this study, a system is proposed for earthquake prediction with data mining techniques. In the study in which Cross Industry Standard Process for Data Mining (CRISP-DM) approach has been used as data mining methodology, seismic bumps data obtained from mines has been analyzed. Extreme learning machine (ELM) which is an effective and rapid classification algorithm has been used in the modeling phase. In the evaluation stage, different performance evaluation criteria such as classification accuracy, sensitivity, specificity and kappa value have been used. The results are promising for earthquake prediction.*

Keywords -*Earthquake prediction, extreme learning machine, data mining, CRISP-DM*

I. INTRODUCTION

Earthquake occurs as a result of unexpectedly emerged energy in the earth's crust. Due to this energy, seismic waves occur and these waves create a shock effect on earth. There are many factors that affect the earthquake. Increase in factors increases the complexity. Complexity makes it difficult to predict location, intensity and time of the earthquake.

Researchers use different information for prediction of possible earthquake. Earth's crust form change, slope change, radon gas changes in well and springs, elastic variable wave velocities, changes in groundwater level, seismic pulses, are examples ground magnetic field change can be given as example.

Some studies carried out in the literature using seismic bumps are as follows. Bilen et. al. [1] proposed a system for earthquake prediction by analyzing seismic bump data. 94.11% classification accuracy was achieved in the study in which k nearest neighbor algorithm was used as classification algorithm. Celik et al. [2] proposed an intelligent system for earthquake prediction by analyzing seismic bump data. 91% classification accuracy was achieved in the study in which support vector machine was used as classification algorithm. Dehbozorgi and Farokhi [3] used neuro-fuzzy system algorithm in their study conducted by using seismometer data. In the study, 82% accuracy rate was obtained. Zhang et. al. [4] proposed multi-scale wavelet analysis for single-component recordings. Colaket. al. [5] used the wavelet method and average energy value for the detection of seismic wave arrival time in three-component stations. Xu et. al. [6] carried out analysis on data obtained from DEMETER satellite. From the satellite they obtained information such as seismic band information electron density, electron temperature, ion temperature and oxygen ion intensity. In this study, seismic bump data obtained from the coal mines of Poland has been used for earthquake prediction. Knowledge discovery has been performed from obtained data by using data mining methods. ELM which is an effective and fast algorithm was used at the classification stage.

II. MATERIAL AND METHODS

2.1. Data

In this study, a dataset which is used to predict seismic bumps bigger than 10^4 J has been preferred [7]. These data were taken from the mines in Poland with 8 hour shift intervals. There are 18 features in this data set consisted of total 2584 samples. There are two different classes in the data set, 0 and 1. 0 means that there is no earthquake hazard in the next interval, 1 means the presence of the earthquake hazard. Features and descriptions in the data set are presented in Table 1.

Table1. The features of the dataset

Feature	Description
seismic	result of shift seismic hazard assessment in the mine working obtained by the seismic method (a - lack of hazard, b - low hazard, c - high hazard, d - danger state);
seismoacoustic	result of shift seismic hazard assessment in the mine working obtained by the seismoacoustic method;
shift	information about type of a shift (W - coal-getting, N -preparation);
genergy	seismic energy recorded within previous shift by the most active geophone (GMax) out of geophones monitoring the longwall;
gpuls	a number of pulses recorded within previous shift by GMax;
gdenergy	a deviation of energy recorded within previous shift by GMax from average energy recorded during eight previous shifts;
gdpuls	a deviation of a number of pulses recorded within previous shift by GMax from average number of pulses recorded during eight previous shifts;
ghazard	result of shift seismic hazard assessment in the mine working obtained by the seismoacoustic method based on registration coming form GMax only;
nbumps	the number of seismic bumps recorded within previous shift;
nbumps2	the number of seismic bumps (in energy range $[10^2,10^3)$) registered within previous shift;
nbumps3	the number of seismic bumps (in energy range $[10^3,10^4)$) registered within previous shift;
nbumps4	the number of seismic bumps (in energy range $[10^4,10^5)$) registered within previous shift;
nbumps5	the number of seismic bumps (in energy range $[10^5,10^6)$) registered within the last shift;
nbumps6	the number of seismic bumps (in energy range $[10^6,10^7)$) registered within previous shift;
nbumps7	the number of seismic bumps (in energy range $[10^7,10^8)$) registered within previous shift;
nbumps89	the number of seismic bumps (in energy range $[10^8,10^{10})$) registered within previous shift;
energy	total energy of seismic bumps registered within previous shift;
maxenergy	the maximum energy of the seismic bumps registered within previous shift;
class	the decision attribute - '1' means that high energy seismic bump occurred in the next shift ('hazardous state'), '0' means that no high energy seismic bumps occurred in the next shift ('non-hazardous state').

2.2. Data Mining

Data mining is the process of determining of qualified knowledge previously unknown in large databases [8]. For this purpose, such methods statistic, machine learning, artificial intelligence, database management and data visualization are utilized. In the literature, there are numerous methods available presented as data mining methodology. CRISP-DM model has been used in this study.

According to CRISP-DM process, data mining process is an interactive and iterative process consisting of six phases. CRISP-DM has been developed by a consortium with the original members of Daimler-Benz, SPSS and NCR. Fig. 1 shows the process steps of this methodology [9]. According to this methodology, data mining process consists of the following steps:

1. Business Understanding: The aim, objectives and pre-strategy are determined
2. Data Understanding: Data is collected, the data quality is evaluated
3. Data Preparation: Last data is prepared, events and variables are selected for analysis. It is the longest process. Data mart will be created at the end of this process.
4. Modelling: Suitable modeling techniques are selected and applied.
5. Evaluation: One or more than one models are evaluated, it is checked that whether or not the aim has been reached.
6. Deployment: Reporting is performed and it is deployed.

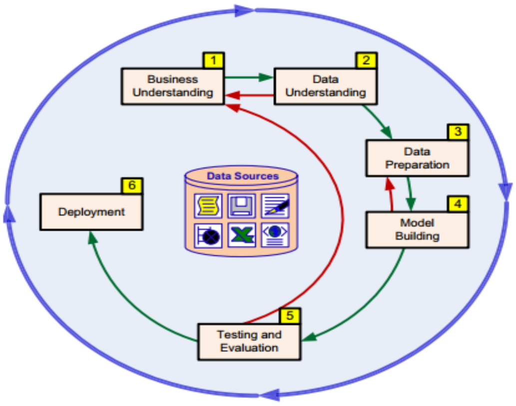


Fig1. CRISP-DM process model for data mining applications.

2.3. Extreme Learning Machine

In this study, ELM method is used in the modeling stage of the data mining. ELM which is proposed by Huang et al., is a learning algorithm developed for a feedforward neural network with one hidden layer [10, 11]. Unlike the gradient-based feed-forward networks, the input weights and threshold values are generated randomly in ELM. In addition, analytical methods are used in the calculation of output weights. In this way, learning process is accelerated. As well as the ability to learn faster, ELM has better generalization performance than feed-forward networks learning by back-propagation algorithm. Fig.2 shows the structure of ELM algorithm.

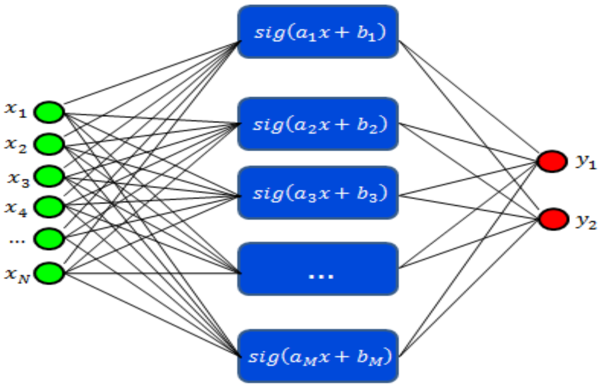


Fig.2. Structure of the ELM

In Fig.2, $x = (x_1, x_2, x_3, \dots, x_N)$ and $y = (y_1, y_2, y_3, \dots, y_N)$ denote input and output features, respectively. Mathematical expression of a network which has M hidden layer neuron is calculated by using equation (1).

$$\sum_{i=1}^{M} \beta_i g(w_i x_j + b_i) = o_j, \quad j = 1, \dots, N \tag{1}$$

where $w_i = w_{i1}, w_{i2}, \dots, w_{in}$ are weight values between input and hidden layer, $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{im})$ are weights between output and hidden layer, $b = b_1, b_2, \dots, b_M$ threshold values of the hidden layer neurons and o_j denotes output values of the hidden layer neurons. $g(.)$is the activation function.

In an ELM which has *M* hidden neuron and $g(x)$ activation function, the error approaching to zero is a desired situation. As a result, $\sum_{j=1}^{N} \|o_j - y_j\| = 0$ is a targeted situation. Equation (1) can be written by reinterpreting equation (2).

$$\sum_{i=1}^{M} \beta_i g(w_i x_j + b_i) = y_j, \quad j = 1, \dots, N \tag{2}$$

Nequations given in equation (2) can be canceled as in equation (3):

$$Y = H\beta \tag{3}$$

where, H denotes hidden layer output matrix. βdenotes output layer weights and Y denotes output vector. These values are calculated using equation (4)-(6).

$$H = \begin{bmatrix} g(w_1.x_1 + b_1) & \dots & g(w_N.x_1 + b_1) \\ \vdots & \dots & \vdots \\ g(w_1.x_M + b_1) & \dots & g(w_N.x_M + b_N) \end{bmatrix}_{NxM} \tag{4}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_M^T \end{bmatrix}_{Mx1} \tag{5}$$

$$Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{Nx1} \tag{6}$$

In ELM, input weights and threshold values of hidden layer neurons are generated randomly and H hidden layer output matrix is obtained analytically. Equation (3) represents the structure of linear equation solved by ELM structure to generate output. In the conventional structures, processes occur iteratively during the training phase of the network. ELM finalizes the training by solving a linear equation at once without iteration. Equation (7) is used to obtain β values from Equation (7):

$$\beta = H^{+}Y \tag{7}$$

Here, H^{+} has been defined as MoorePenrose matrix, generalized inverse of H output matrix.
Data mining approach proposed in this study is given in Fig. 3.

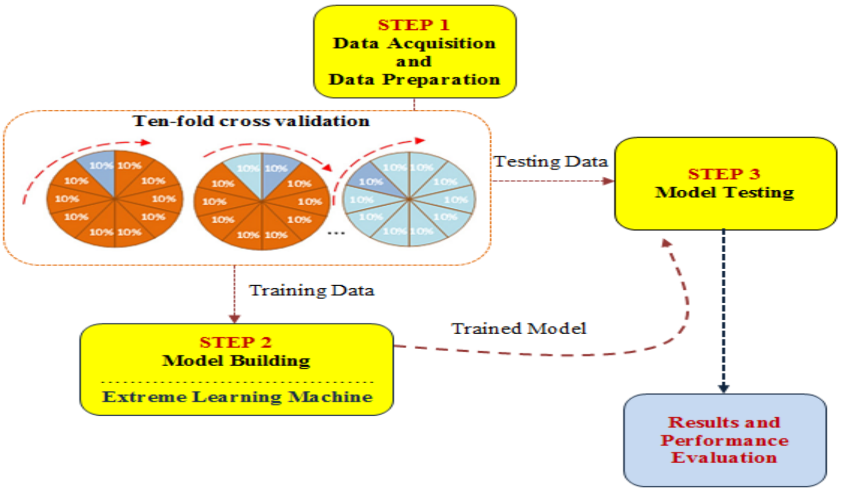


Fig.3.The proposed data mining approach for earthquake prediction

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

Data have been normalized in the range of 0-1 during the data preparation. Equation (7) has been used for normalization.

$$x' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{7}$$

In this equation, x' denotes normalized data; x_i indicates input value; x_{min} is the smallest number within input set; x_{max} is the maximum number within input set.
10-fold cross-validation method has been used to increase the validity of experimental study. In modelling stage, ELM has been used. Activation function is an important parameter of this algorithm. In this study, experiments have performed by using three different activation functions (Sigmoid, Sin, Hardlim). As a result of the experiments, the best result has been obtained by sigmoidal activation function. One of the important parameters in ELM is the number of hidden layer neurons. Search algorithm has been used to determine the number of neurons which give high accuracy rate. The number of neurons has been started from 5 in the search algorithm. And the optimum number of neurons has been investigated with the increment of 5 till 500 neurons. The obtained results are presented in Fig. 4. Accordingly, the highest accuracy rate has been obtained with 90 hidden layer neurons.

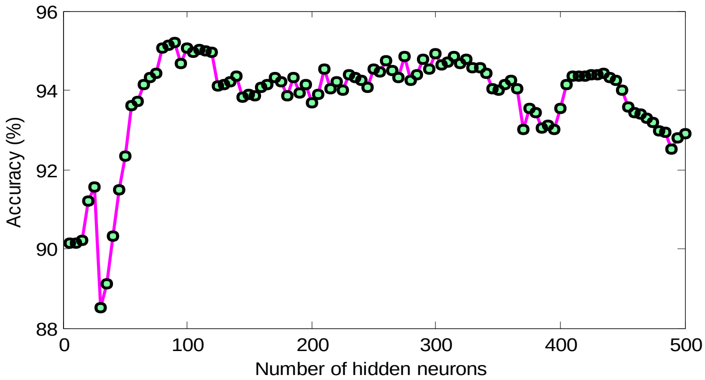


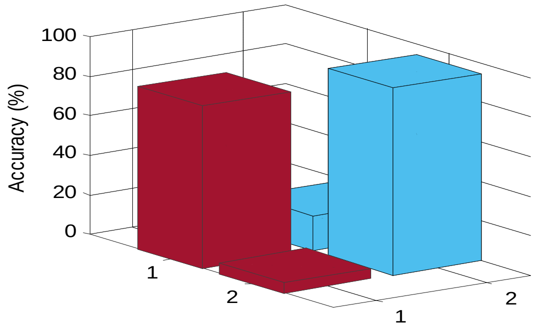
Fig4. Classification accuracies by number of hidden layers.

Obtained statistical results are given in Table 2. The results obtained by using 3 different activation functions are also presented in the table.

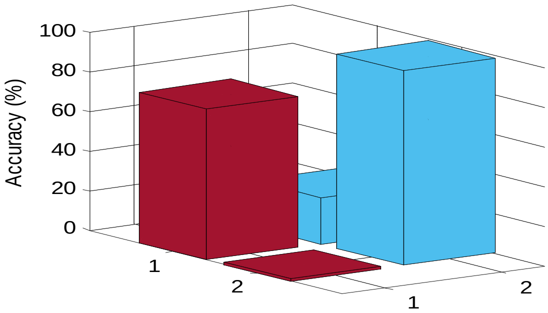
Table2. The features of the dataset

Activation Function	Performance metrics	Value
Sigmoid	Accuracy	95.22%
	Sensitivity	65.55%
	Specificity	97.42%
	Kappa statistic value	0.6298
Sin	Accuracy	94.77%
	Sensitivity	65%
	Specificity	97.1%
	Kappa statistic value	0.606
Hardlim	Accuracy	94.73%
	Sensitivity	65%
	Specificity	96.96%
	Kappa statistic value	0.6041

In addition, the confusion matrixes obtained in different folds are presented in Fig.5.



(a)



(b)

Fig 5. Confusion Matrix a) The result obtained in Fold 4 b) The result obtained in Fold 7
1: Hazardous, 2: Not Hazardous

The comparative analysis made with the studies in the literature carried out on the same data set is as follows: Bilen et. al. [1] proposed a method based on Principal Component analysis and k-nearest neighbors algorithm. They achieved 94.11% classification accuracy. Celik et al. [2] proposed a method based SVM. They obtained 91% classification accuracy. As a result, developed method gave better results than the methods proposed in previous studies. The data used in this study is unbalanced data. Class distribution is not balanced. In spite of this, good results have been obtained with the proposed method.

IV. CONCLUSION

In this study, a data mining based decision support system has been developed in order to predict the earthquake using seismic bump data. After passing the data mining stages, seismic data which have been obtained from coal mines have been classified with ELM in the modelling phase. 95.22% accuracy rate has been achieved with the proposed method. It is aimed to strengthen the data set in future studies. This aim will be pursued by the data obtained from different mines and fields. Performing of comparative analysis with different algorithms is aimed in modelling stage. In addition, feature selection algorithms will be used for the detection of features which affect the earthquake prediction more.

REFERENCES

[1] M. Bilen, A.H. Işık, T. Yiğit, Seismic hazard prediction with classification of seismic pulses, International Burdur Earthquake & Environment Symposium (IBEES2015), Burdur, Turkey, 2015, 41-48.

[2] E. Celik, M. Atalay, and H. Bayer, Earthquake prediction using seismic bumps with Artificial Neural Networks and Support Vector Machines, Signal Processing and Communications Applications Conference, Trabzon, Turkey, 2014, 730-733.

[3] L. Dehbozorgi, F. Farokhi, Effective feature selection for short-term earthquake prediction using neuro-fuzzy classifier", II. International Conference on Geoscience and Remote Sensing, Qingdao, 2010, 165-169

[4] H. Zhang, C. Thurber, and C. Rowe, Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings, Bulletin of the Seismological Society of America, 93(5), 2003, 1904-1912.

[5] O.H. Colak, T.C. Destici, H. Arman, and O. Cerezci, Detection of P-and S-waves arrival times using the discrete wavelet transform in real seismograms, The Arabian Journal for Science and Engineering, 34, 2009, 79-89.

[6] F. Xu, X. Song, X. Wang, J. Su, Neural network model for earthquake prediction using DEMETER data and seismic belt information, II. WRI Global Congress on Intelligent Systems, Wuhan, 2010, 180-183.

[7] J. Kabiesz, B. Sikora, M. Sikora, and L. Wrobel, Application of rule-based models for seismic hazard prediction in coal mines, ActaMontanisticaSlovaca, 4, 2013, 262-272

[8] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, From data mining to knowledge discovery in databases. AI magazine, 17(3), 1996, 37.

[9] C. Shearer, The CRISP-DM model: the new blueprint for data mining, J Data Warehousing, 5, 2000, 13-22.

[10] G.B. Huang, Q.Y. Zhu, and C.K. Siew, Extreme learning machine: theory and applications, Neurocomputing, 70(1), 2006, 489-501.

[11] O.F. Alcin, A. Şengür, and M.C. İnce, Forward-backward pursuit based sparse extreme learning machine, Journal of the Faculty of Engineering and Architecture of Gazi University, 30(1), 2015, 111-117.