LARGE EARTHQUAKE MAGNITUDE PREDICTION IN TAIWAN BASED ON DEEP LEARNING NEURAL NETWORK

JP. Huang, *XA. Wang,* *Y. Zhao,* *C. Xin,* *H. Xiang*[†]

Abstract: In this paper, a deep learning-based method for earthquake prediction is proposed. Large-magnitude earthquakes and tsunamis triggered by earthquakes can kill thousands of people and cause millions of dollars worth of economic losses. The accurate prediction of large-magnitude earthquakes is a worldwide problem. In recent years, deep learning technology that can automatically extract features from mass data has been applied in image recognition, natural language processing, object recognition, etc., with great success. We explore to apply deep learning technology to earthquake prediction. We propose a deep learning method for continuous earthquake prediction using historical seismic events. First, we project the historical seismic events onto a topographic map. Taking Taiwan as an example, we generate the images of the dataset for deep learning and mark a label "1" or "0", depending on whether in the upcoming 30 days a greater than M6 earthquake will occur. Second, we train our deep leaning network model, using the images of the dataset. Finally, we make earthquake predictions, using the trained network model. The result shows that we can get the best result, when we predict earthquakes in the upcoming 30 days using data from the past 120 days. Here, we use R score as the performance metrics. The best R score is 0.303. Although the R score is not high enough, using the past 120 days' historic seismic event to predict the upcoming 30 days' biggest earthquake magnitude can be seen as the pattern of Taiwan earthquake because the R score is rather good compared to other datasets. The proposed method performs well without manually designing feature vectors, as in the traditional neural network method. This method can be applied to earthquake prediction in other seismic zones.

Key words: *large earthquake magnitude prediction, deep learning neural network, taiwan, pattern recognition*

---

*Jipan Huang; Xin'an Wang – Corresponding author; Yong Zhao; Chen Xin; The Research Center of Earthquake Monitoring and Forecasting Technology, Peking University Shenzhen Graduate School, Shenzhen, Guangdong, 518055, China, E-mail: huangjp@sz.pku.edu.cn, anxinwang@pku.edu.cn, zhaoyong@pkusz.edu.cn, xchiggs@sz.pku.edu.cn.

[†]Han Xiang; Guangdong Medical University, Dongguan, Guangdong, 523808, China, E-mail: hanxiangcn@outlook.com

1. Introduction

An earthquake is a highly destructive natural disaster. If earthquakes can be accurately predicted, many lives can be saved and economic loss can be avoided. China is a country with high seismic activity. A famous instrument for earthquake monitoring in ancient times was Di Dong Yi in the Han Dynasty. Unfortunately, it was not handed down and was lost.

In the last several decades, the first accurately predicted earthquake was the Haicheng earthquake in 1975 [29]. But, the Tangshan earthquake in 1976 was not predicted, and it led to 250,000 deaths and 160,000 injuries.

Taiwan is one of the districts with the highest seismic activity. The extension direction of the fault in Taiwan island is with the island contours in the same direction, mainly from north to south [6]. On September 21, 1999, at 01.47 h local time, a strong earthquake, with a 7.3 magnitude on the Richter scale and a depth of approximately 7.5 km, occurred near the town of Chi-Chi in central Taiwan N23.78, E120.84. The island suffered catastrophic damage, with 2,375 lives lost, over 10,000 injuries, and more than 30,000 collapsed buildings [31].

Methods used to predict earthquakes in the last decade can be divided into those based on probability and statistics and those based on artificial intelligence.

Among the first category, Petersen et.al assumes that the occurrence probability of earthquakes follows a Poisson distribution [25]. Shen et al. presented a probabilistic forecast method based on strain rate [26]. A model based on smoothed seismicity to predict earthquakes of magnitude equal to or larger than 5.0 in southern California was proposed by Kagan et al. [14].

Methods based on artificial intelligence can be divided into non-supervised learning frameworks and supervised learning frameworks. In a non-supervised learning framework, some methods based on clustering techniques [8, 21] or association rules [9, 12] have been applied to predict earthquake magnitudes in areas with high seismic activity in Chile, Portugal, and Spain. Mirrashid et al. used a neuro-fuzzy inference system based on a C-means algorithm to predicate earthquakes in Iran [20].

The supervised learning framework, namely regression or classification techniques, has been preferred by the majority of researchers to forecast earthquakes in recent years. H. Adeli et al. proposed a probabilistic neural network to predict earthquakes in a specific zone, using eight kinds of earthquake indicators to train the model, and obtained good results (two of four earthquakes were predicted correctly) for earthquakes with magnitudes ranging from 4.0 to 6.0 [1]. Then, they used a recurrent neural network to predict the time and location of earthquakes with magnitudes equal to or greater than 6.0, again with good results [23]. A multilayer perception neural network has been used to predict earthquakes using the total electron content (TEC) time series data [3]. An artificial neural network based on some earthquake predictors has been proposed to the prediction of earthquakes of medium-large magnitude for the city of Tokyo [5].

Recently, deep learning technology has been applied to image recognition, natural language processing, object detection, motion modeling, etc., with great success. The basic notion of a deep learning algorithm is extracting features from low- to high-level data with multiple structures. In the field of machine learning, the per-

formance of a model is often strongly influenced by how the data are represented. Thus, devising effective representations of the data is an important component of constructing a high-performing model [30]. Deep learning can automatically learn these underlying representations or features from the data itself. Deep learning technology is leading the third upsurge of artificial intelligence research.

At the same time, in the domain of traffic flow prediction [18], toxicity prediction [28], and event-driven stock prediction [7], people used deep learning technology and achieved good results.

Nonetheless, although artificial intelligence technology has been applied to earthquake prediction in many zones, to the authors' knowledge, a deep learning algorithm has not been used for prediction of earthquakes in Taiwan.

We will explore to apply deep learning technology to earthquake prediction in Taiwan. The remainder of the article is structured as follows: Section 2 shows the basic research methodology; Section 3 covers deep learning fundamentals; Section 4 addresses data set construction and optimization; Section 5 shows the deep learning model and optimization; Section 6 provides the result and discussion; and Section 7 is the summary.

2.   Methodology

This experimental setup meets the requirements established by Allen [4], whose work claimed that an earthquake prediction must provide

1. a specific location or area (Taiwan, N21-26,E119-123);

2. a specific span of time (within the next 30 days);

3. a specific magnitude range (magnitude greater than or equal to 6.0).

We selected the Taiwan seismic events from China Seismic Information net [2]. Then we constructed our dataset elaborately, to meet the Caffe deep learning framework [13]. We then trained our neural network model. Finally, we put the test images into the model and obtained the prediction result. The complete procedure is shown in Fig. 1.

3.   Deep learning fundamentals

3.1   Convolutional and subsampling layers

The convolutional layer and subsampling (also called pooling) layer are the key layers of a convolutional neural network (CNN). CNNs were proposed by LeCun [17], which are inspired by biological processes. The traditional multi-layer perception neural networks have too many weights and often overfitting. CNNs use three techniques to solve above problems : local receptive fields, shared weights and subsampling. Each unit of a layer receives input from a set of units located in a small neighborhood in the previous layer. The technique is called local receptive fields. The convolutional layer is composed of feature maps. All units in a feature map share the same weights. The shared weights technique will reduce many weight
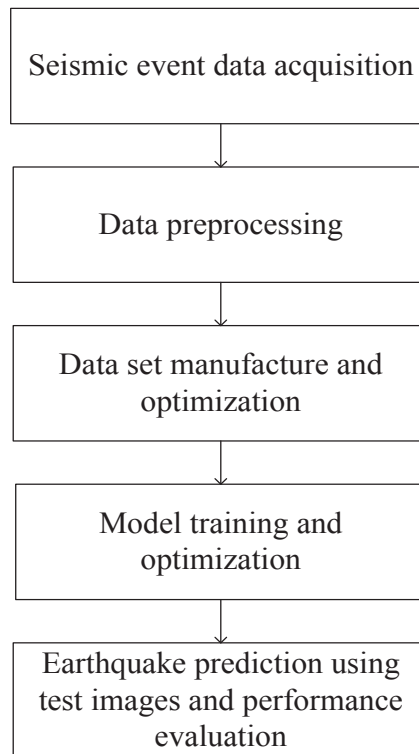
```
┌─────────────────────────────────────┐
│    Seismic event data acquisition    │
└─────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────┐
│          Data preprocessing          │
└─────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────┐
│       Data set manufacture and       │
│             optimization             │
└─────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────┐
│          Model training and          │
│             optimization             │
└─────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────┐
│      Earthquake prediction using     │
│     test images and performance      │
│              evaluation              │
└─────────────────────────────────────┘
```

Fig. 1 *The methodology procedure.*

parameters and form more useful feature maps. Subsampling layer is also called polling layer. The subsampling layer will reduce the resolution of the feature maps and reduce the sensitivity of the output to shifts and distortions.

3.2 Activation functions

In neural networks, the activation function of a node defines the output of that node given an input or set of inputs. The popular activations are Sigmoid and ReLU (rectified linear unit). In chapter 5.2, we will point out our activation function.

3.3 Reducing overfitting

A common problem in the training process of artificial neural network is over-fitting. There are many methods to reduce overfitting, such as label-preserving transformations [17], dropout [10] and so on. Here, we use the method "dropout". This method will set the output of each hidden neuron to zero with an adjustable probability (0.5 is very popular). Through the method, more useful robust feature can be learned [17].

4. Data set construction and optimization

In consideration that we must use more seismic events data whose magnitude is lower than 6. So, we used the data from the website of China Earthquake information net [2], on which the complete data started from January 1, 1970. They were in the CSV format, and were marked with earthquake time, magnitude, longitude, and latitude. The time range was between January 1, 1970, and May 25, 2016.

First, we preprocessed the data. We selected the seismic events in the Taiwan dangerous zone (from 119 to 123 degrees east longitude and from 21 to 26 degrees north latitude). The range of longitude and latitude was confirmed based on all the earthquakes occurring in Taiwan with magnitude greater than 7 (Fig. 2).

**Fig. 2** *All M7 and greater earthquakes in Taiwan.*

Then, we constructed the dataset. We set M6 as the seismic event threshold because M6 or greater earthquakes are often destructive. M6 or greater seismic events happening in the upcoming 30 days were positive samples, and less than M6 seismic events were negative samples. The seismic events superposition within N days were projected on the Taiwan map and marked with a "1" if the largest earthquake's magnitude in the upcoming 30 days was greater than or equal to 6. The background map is useless because all samples have the same background map but it make the sample looks not dull. In the same way, negative samples were marked with a "0". According to the Omori´s law [22], aftershock activity is predictable. So, the earthquake events within the radius of 50 Km within 7 days after the occurrence of the M6+ main-shock were deleted from the catalogue. From

the starting time of January 1, 1970, we got a sample every step forward 30 days. Finally, we got all samples in the dataset. The image size of each sample is 256 pixel length and 256 pixel width. Thus, we had a dataset in accordance with the Caffe deep learning framework [13]. Fig. 3 shows the whole procedure.

1.Seismic event data acquisition

2.From 1970/1/1, count N days, look if M6 earthquake occur in N~N+30 day. If yes, it is a positive sample with a label as 1. If no, it is a negative sample with a label as 0.
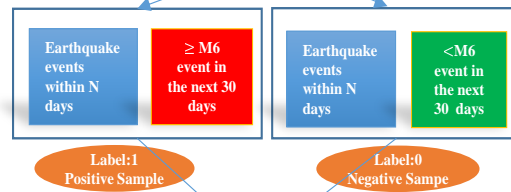
3.All earthquake events within current N days are projected on the map as points with size and brightness depending on their magnitudes. Then, one sample is completed.

4. Step forward 30 days, repeat step 2 and 3 ,and generate a new sample. Do this until 2016/4/23 and all the samples are completed.
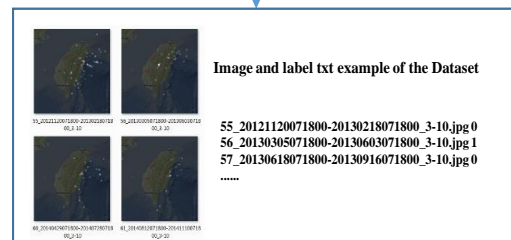
Fig. 3 *The procedure of data set construction and example.*

Finally, we carried out the dataset production optimization:

1. Earthquakes occurring in the same place within several seconds to several minutes are regarded as the same seismic event.

2. The sample ratio of the training set and testing set is 5:1.

Through changing the value of N, we obtained seven kinds of datasets: DS30-30-30, DS60-30-30, DS90-30-30, DS120-30-30, DS150-30-30, DS180-30-30, and DS210-30-30. For example, DS150-30-30, means every sample in the dataset uses the past

150 days to predict the largest earthquake's magnitude in the next 30 days. We listed details of all the data sets in Tab. I. We got 95 samples in the testing set in DS30-30-30, DS60-30-30. . . DS180-30-30. Due to the longer time window of the DS210-30-30, only 94 samples were got in the testing set.

Data set	Samples of Training set	The Last Day of the Training set	Samples of Testing set
DS30-30-30	468	Jul 10, 2008	95
DS60-30-30	467	Jul 10, 2008	95
DS90-30-30	466	Jul 10, 2008	95
DS120-30-30	465	Jul 10, 2008	95
DS150-30-30	464	Jul 10, 2008	95
DS180-30-30	463	Aug 9, 2008	95
DS210-30-30	463	Aug 9, 2008	94

Tab. I *Data set details.*

We obtained the above datasets automatically using software developed based on MATLAB R2014a.

5. Deep learning model and optimization

5.1 Model

We modified the Cifar10 model [16] to obtain our model because we only needed to classify two classes. At the same time, we added two dropout layers to reduce overfitting. The model structure can be seen in Fig. 4. The detailed configuration information are listed in Tab. II.
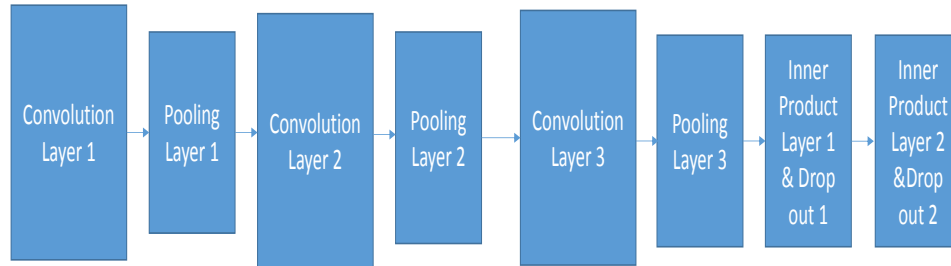

Fig. 4 *Our deep neural network model for earthquake prediction.*

5.2 Tune and optimization

We used momentum and weight delay technology in our model. The update rule for weight w was formula (1,2). In the formulas i is the iteration index, Δw is the weight variable, β is the momentum , λ is the weight decay, L is the cost function,

155

Layer	Kernel Size	Stride	Pad	Other
Convolution 1	5	1	2	–
Pooling 1	3	2	0	Max pooling
Convolution 2	5	1	2	–
Pooling 2	3	2	0	Average pooling
Convolution 3	5	1	2	–
Pooling 3	3	2	0	Average pooling

Tab. II *Model configurations.*

α is the learning rate and $\langle \frac{\partial L}{\partial w}\big|_{w_i} \rangle_{D_i}$ is the average over the ith batch D_i of the derivative of the objective with respect to w, evaluated at w_i

$$\Delta w_{i+1} = \beta * \Delta w_i - \lambda * \alpha * w_i - \alpha * \langle \frac{\partial L}{\partial w}\big|_{w_i} \rangle_{D_i}, \qquad (1)$$

$$w_{i+1} = w_i + \Delta w_{i+1}. \qquad (2)$$

In the selection of learning rate, we found that higher (0.1) and lower (0.0001) base learning rates would decrease the accuracy(only about 55%). We got the best result when the learning rate was 0.001(about 90%). The momentum were used to accelerated model convergence. In this paper, we used 0.9 as the momentum value. Weight decay is one kind of regularization and will decrease useless weight to zero [17]. We used a weight decay of 0.0005. The dropout rate of the two dropout layers was 0.5, meaning the relative hidden neuron weight would be set to zero with the probability 0.5 to avoid overfitting [17].

The loss function during the training of the models that achieved the best results reported in this paper was the standard Softmax loss function, that is, the log-likelihood error function described in [15]. For better normalization, we used the rectified linear unit activation function(ReLU), not the sigmoid activation function.

All experiments were run on a Quadro K1200 GPU. The Quadro K1200 server has 4 GB of GPU RAM and 32 GB of regular RAM.

6. Results and discussion

6.1 Performance evaluation

In the machine learning domain, there are some performance evaluation metrics for unbalanced samples, such as the confusion matrix, receiver operating characteristic (ROC), and area under curve (AUC). Among them, the confusion matrix is often employed when using artificial intelligence to predict earthquakes [1, 19]. Here, the true positive (TP) is the number of times that an upcoming earthquake has been correctly predicted. The true negative (TN) is the number of times that neither an earthquake prediction model has triggered an alarm nor an earthquake has occurred. A false positive (FP) is the number of times that an earthquake

prediction model has triggered an alarm but no earthquake has occurred. A false negative (FN) is the number of times that a classifier has not triggered an alarm but an earthquake did occur. The table below is the confusion matrix in our research.

Actual class	Predictied class	
	M6 earthquake or greater	less than M6 earthquake
M6 earthquake or greater	TP	FN
less than M6 earthquake	FP	TN

Tab. III *Confusion matrix for our reseach.*

Based on the confusion matrix, there is a popular performance evaluation metric in the earthquake prediction domain called the R score [1, 23, 24, 27]. It is defined by formula (3). The value domain of the R score is between -1 and 1. An R score is approximately 0 for completely random guesses, and approximately 1 for completely successful predictions [24]

$$R = \frac{TP}{TP + FN} - \frac{FP}{TN + FP}. \tag{3}$$

In our research, we got the best result on DS120-30-30. The R score is 0.303. The detailed results are in Table IV. The Result is better than random prediction. Next, We did the significance test to further validate the result. We randomized the catalogue and got different data set with the same parameters for the processing(length of windows, number of samples, numbers of neural network, convolution, activation function, and so on). These random catalogue were named as "RDS30-30-30", "RDS60-30-30" et al. Through the same process, we got another group R scores in Tab. V. At last, we did the one-way analysis of variance(ANOVA) [11] under the significant level α as 0.05 and got the p value 0.0185. The null-hypothesis was reject. There are obvious difference between the two R score groups.

Data Set	TP	FN	FP	TN	R Score
DS30-30-30	0	19	0	76	0
DS60-30-30	3	16	1	75	0.145
DS90-30-30	5	14	2	74	0.236
DS120-30-30	14	5	33	43	0.303
DS150-30-30	11	8	37	39	0.092
DS180-30-30	7	12	16	60	0.157
DS210-30-30	2	16	4	72	0.058

Tab. IV *Prediction results for different datasets.*

6.2 Discussion

Earthquake prediction is a complex, worldwide problem. In the domain of predicting earthquakes using neural networks, the number of test datasets is often

Data Set	TP	FN	FP	TN	R Score
RDS30-30-30	0	19	0	76	0
RDS60-30-30	1	18	2	74	0.026
RDS90-30-30	2	17	5	71	0.039
RDS120-30-30	10	9	35	41	0.065
RDS150-30-30	4	15	12	64	0.052
RDS180-30-30	6	13	22	54	0.026
RDS210-30-30	2	16	7	69	0.019

Tab. V *Prediction results for different randomized datasets.*

small. Our research uses 95 or 94 samples in the test dataset, and 19 or 18 positive samples. Using the past N days of seismic events to predict whether the largest earthquake magnitude in the upcoming 30 days is greater than M6, we found that we can get a relatively good result (0.303) when $N = 120$. We saw that we can not get the best result when N is too small or too large. We think it is mainly due to the destructive earthquakes need a certain time preparation process. If N is too small(much less than the preparation process time), the model can not learn enough useful information and can not do a good prediction. Similarly, if N is too big, the model will be interfered by useless information and also can not do a good prediction. At the same time, we did the significance test. The significance test shows that the result is better than random prediction.

7. Summary

Deep learning is a set of powerful machine learning algorithms and concepts that have seen groundbreaking success for the last 10 years. The main benefit of deep neural networks is their ability to learn complex, nonlinear hypotheses through data without the need to explicitly model features.

We applied deep learning technology to the large earthquake prediction domain and received a good result. We received the good result when using the past 120 days of seismic events to predict the largest earthquake magnitude in Taiwan in the upcoming 30 days. This temporal pattern of Taiwan earthquake was found. The pattern may be useful for further earthquake prediction and disaster reduction.

Next, we will try to further optimize the dataset and the model. At the same time, we will apply deep learning technology to other seismic danger zones (such as Japan and Chile).

Acknowledgment

References

[1] ADELI H., PANAKKAT A.: A probabilistic neural network for earthquake magnitude prediction. Neural Networks the Official Journal of the International Neural Network Society. 2009, 22(7), pp. 1018–1024.

[2] ADMINISTRATION C.E.: More than m2 earthquakes in the country (since 1965), 2016. http://www.csi.ac.cn/publish/main/813/5/index.html

[3] AKHOONDZADEH M.: A MLP neural network as an investigator of tec time series to detect seismo-ionospheric anomalies. Advances in Space Research, 2013, 51(11), pp. 2048–2057.

[4] ALLEN C.R.: Responsibilities in earthquake prediction: To the seismological society of America, delivered in Edmonton, Alberta, Bulletin of the Seismological Society of America. 1976, (6), pp. 2069–2074.

[5] ASENCIO-CORTÉS G., MARTÍNEZ-ÁLVAREZ F., TRONCOSO A., MORALES-ESTEBAN A.: Medium-large earthquake magnitude prediction in tokyo with artificial neural networks. Neural Computing & Applications. 2015, pp. 1–13.

[6] CAMANNI G., ALVAREZ-MARRON J., BROWN D., AYALA C., WU Y.M., HSIEH H.H.: The deep structure of south-central taiwan illuminated by seismic tomography and earthquake hypocenter data. Tectonophysics. 2016, 679, pp. 235–245.

[7] DING X., ZHANG Y., LIU T., DUAN J.: Deep learning for event-driven stock prediction. In: International Conference on Artificial Intelligence, 2015.

[8] FLORIDO E., MARTÍNEZ-ÁLVAREZ F., MORALES-ESTEBAN A., REYES J., AZNARTE-MELLADO J.L.: Detecting precursory patterns to enhance earthquake prediction in chile. Computers & Geosciences. 2015, 76, pp. 112–120.

[9] GUTENBERG B., RICHTER C.F.: Book reviews: Seismicity of the earth and associated phenomena. Science, 1950, p. 111.

[10] HINTON G.E., SRIVASTAVA N., KRIZHEVSKY A., SUTSKEVER I., SALAKHUTDINOV R.R.: Improving neural networks by preventing co-adaptation of feature detectors. Computer Science. 2012, 3(4), pp. 212–223.

[11] HOWELL D.C.: Statistical methods for psychology. Journal of the Royal Statistical Society. 2012, 43(43).

[12] IKRAM A., QAMAR U.: Developing an expert system based on association rules and predicate logic for earthquake prediction. Knowledge-Based Systems. 2015, 75(C), pp. 87–103.

[13] JIA Y., SHELHAMER E., DONAHUE J., KARAYEV S., LONG J., GIRSHICK R., GUADARRAMA S., DARRELL T.: Caffe: Convolutional architecture for fast feature embedding. Eprint Arxiv, 2014, pp. 675–678.

[14] KAGAN Y., JACKSON D.D., RONG Y.: A testable five-year forecast of moderate and large earthquakes in southern california based on smoothed seismicity. Seismological Research Letters, 2007, 78(1), pp. 94–98.

[15] KRÄHENBÜHL P., KOLTUN V.: Parameter learning and convergent inference for dense random fields. In: ICML. 2013, (3), pp. 513–521.

[16] KRIZHEVSKY A.: Convolutional deep belief networks on cifar-10, 2012.

[17] KRIZHEVSKY A., SUTSKEVER I., HINTON G.E.: Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 2012, 25(2).

[18] LV Y., DUAN Y., KANG W., LI Z., WANG F.Y.: Traffic flow prediction with big data: A deep learning approach. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2), pp. 1–9.

[19] MARTÍNEZ-ÁLVAREZ F., REYES J., MORALES-ESTEBAN A., RUBIO-ESCUDERO C.: Determining the best set of seismicity indicators to predict earthquakes. two case studies: Chile and the iberian peninsula. Knowledge-Based Systems, 2013, 50(3), pp. 198–210.

[20] MIRRASHID M.: Earthquake magnitude prediction by adaptive neuro-fuzzy inference system (anfis) based on fuzzy c-means algorithm. Natural Hazards, 2014, 74(3), pp. 1577–1593.

[21] MORALES-ESTEBAN A., MARTÍNEZ-ÁLVAREZ F., TRONCOSO A., JUSTO J.L., RUBIO-ESCUDERO C.: Pattern recognition to forecast seismic time series. Expert Systems with Applications, 2010, 37(12), pp. 8333–8342.

[22] OMORI F.: On aftershocks of earthquakes. J.coll.sci.imp.univ.tokyo, 1894, 7, pp. 111–200.

[23] PANAKKAT A., ADELI H.: Neural network models for earthquake magnitude prediction using multiple seismicity indicators. International Journal of Neural Systems, 2007, 17(1), pp. 13–33.

[24] PENG K., YIN X., ZHANG L.P.: A statistical investigation of the earthquake predictions using lurr. Pure & Applied Geophysics, 2006, 163(11-12), pp. 2353–2362.

[25] PETERSEN M.D., CAO T., CAMPBELL K.W., FRANKEL A.D.: Time-independent and time-dependent seismic hazard assessment for the state of california: Uniform california earthquake rupture forecast model 1.0. Seismological Research Letters, 2007, 78(1), pp. 99–109.

[26] SHEN Z.K., JACKSON D.D., YAN Y.K.: Implications of geodetic strain rate for future earthquakes, with a five-year forecast of m5 earthquakes in southern california. Seismological Research Letters, 2007, 78(1), pp. 116–120.

[27] SHI Y., ZHANG J.L.G.: An evaluation of chinese annual earthquake predictions, 1990–1998, 2001, 38A, pp. 222–231.

[28] UNTERTHINER T., MAYR A., KLAMBAUER G., HOCHREITER S.: Toxicity prediction using deep learning. Computer Science, 2015.

[29] WANG K.L., CHEN Q.F., SUN S.H., WANG A.D.: Predicting the 1975 haicheng earthquake. Bulletin of the Seismological Society of America, 2006, 96(3), pp. 757–795.

[30] WANG T., WU D.J., COATES A., NG A.Y.: End-to-end text recognition with convolutional neural networks. In: International Conference on Pattern Recognition, 2012, pp. 3304–3308.

[31] WANG W.L., WANG T.T., SU J.J., LIN C.H., SENG C.R., HUANG T.H.: Assessment of damage in mountain tunnels due to the taiwan chi-chi earthquake. Tunnelling & Underground Space Technology, 2001, 16(3), pp. 133–150.