

Aftershock predict based on convolution neural networks

Zhijun Li¹, Gege Jin², Hongmei Yin³, Chen Chen⁴, Jiyong Hua^{1,*}

¹Yangzhou University, Yangzhou, China

²Yangzhou Technician Branch, Jiangsu United Vocational and Technical College, Yangzhou, China

³Yangzhou Customs, Yangzhou, China

⁴Jiangsu College of Tourism, Yangzhou, China

*Corresponding author: 576461158@qq.com

Abstract: Earthquake prediction is a difficult task. Constrained within a certain spatiotemporal range, earthquakes are only a probability event. In a large area, predicting earthquakes based on geographical events that have already occurred is reliable. Predicting the duration of aftershocks under the condition that a major earthquake has already occurred is the research content of this article. Extract 6 features from seismic phase data to predict the aftershock period. We constructed a convolutional neural network model, sorted out 855 data from 1351 data, and trained the network. The accuracy of training verification reaches 90%, and the accuracy of testing reaches 100%. After further refinement, this model can be used to predict the duration of aftershocks in earthquakes. Provide data guidance for earthquake rescue.

Keywords: Convolution Neural Network; Aftershock Predict; Earthquake Predict

1. Introduction

An earthquake is a random event. A large number of earthquake events have left behind rich observational data. With our understanding of natural laws, we may be able to identify the patterns of earthquakes from big data. Cattania et al. [Cattania,2019] believe that earthquakes cannot be considered as an isolated event for research. To study the possibility of earthquakes occurring within a larger regional space. Chang Qing Li [Chang-Qing Li,2018] used the LSTM model to predict the location and direction of fractures in granite fracture experiments conducted in the laboratory. Sehrish et al. believe that neural networks can express the mapping relationship between earthquake occurrence signs and probabilities. They use BAT-ANN networks to avoid the algorithm falling into local optima and missing out on global optima. Asmae Berhich et al. predicted the likelihood of earthquakes based on their time, location, and magnitude.

We believe that the aftershock period can be predicted on the premise that the earthquake has already occurred. Obviously, neural networks are currently the best tool available. In order to prevent overfitting of the model, we chose the convolutional model. In order to make the data more comprehensive, we selected 856 data from 2351 data of an earthquake. In order to make the data features more comprehensive, we selected 7 feature data from the seismic phase data block to form the input vector.

2. Relate works

Helene et al. [1] conducted research on earthquake prediction. In the early days, seismologists believed that prediction was the logical goal of earthquake research. For most of the 20th century, optimism towards predicting earthquakes persisted. As bonuses flow into seismology, it drives predictive research towards conclusions. China seems to have successfully predicted earthquakes, which makes the development of earthquake prediction methods imminent. The goal of seismological research is to predict without any problems, but it should be carried out under the premise of rational and correct use of information and understanding of inherent difficulties. The public's response to earthquake prediction shows that 60-85% of people believe that earthquakes can be predicted.

Asmae Berhich et al. [2] divided the Chilean earthquake dataset into two types: large earthquakes

and small earthquakes. They believe that there are four methods for predicting earthquakes, namely precursor signals, statistical algorithms, machine learning, and deep learning. They take latitude, longitude, depth, year, month, day, hour, minute, second, and magnitude as 10 characteristic parameters from the seismic dataset. Large earthquakes with magnitudes greater than 5.0 are considered major earthquakes, while earthquakes with magnitudes 0.2 to 5.0 are considered minor earthquakes. By constructing an LSTM network with 10 neurons, four prediction results are output: magnitude, latitude, longitude, and year. Normalize the input data to [0,1]. Take 80% of the dataset for training and 20% for testing. The experimental results were evaluated using MAE and MSE.

Saba Sehrish et al.[3] divided earthquake prediction into three categories. Mathematical statistics, artificial intelligence, and hybrid methods. Input time and seismic signals to the ANN network to predict earthquake magnitudes. The evaluation algorithms are MAE and MSE. They believe that different technologies can predict earthquakes. Continuous improvement can lead to more accurate results. Among all technologies, neural networks exhibit better possibilities. It can achieve complex mapping relationships between input and output. Adding the BAT algorithm to ANN can effectively avoid overfitting of the network.

Chang Qing Li et al. [4] conducted fracture tests on granite in the laboratory. Analyze using six parameters: dry density, statistical friction coefficient, Young coefficient, general force, loading rate, and shear force. Obtain digital images through a two-dimensional mapping scanner. Study the transmission characteristics of sound waves during granite fracture process. Explain the mechanism of earthquake occurrence.

Molchan et al. [5] believe that there is no standard method for earthquake prediction and evaluation. It is necessary to carefully examine the theoretical analysis. One important point to emphasize is that algorithms based on early warning mechanisms are not trustworthy.

Cattania et al. [6] proposed that the prediction of large earthquakes should be studied in a large spatiotemporal space. Relatively speaking, small earthquakes are caused by the slow rupture of isolated convex bodies while large earthquakes have already occurred. These fractures are periodically repeated and can be predicted. They conducted research on earthquake prediction from a temporal and spatial perspective.

Qianlong W et al. [7] constructed a two-dimensional input LSTM to reveal the spatiotemporal relationship of historical earthquakes. Divide LSTM into small parts to reduce algorithm complexity. They noticed that most neural network algorithms use different feature inputs. Not fully considering the spatiotemporal relationship of earthquakes. In the time domain, there seems to be a reasonable pattern of seismic activity. In the spatial domain, adjacent geographical activities can trigger each other. RNN is not suitable for handling long-term time dependence. LSTM uses functions to store information, replacing memory units. The unit state is transmitted along the entire path, only undergoing some linear interaction in the middle, and the information can be well maintained to the output end. Compared with one-dimensional input, the algorithm verification accuracy has improved from 79.6% to 87.8%.

Gitis et al. [8] believe that a dense network of GPS receiving stations can monitor the movement of the Earth's surface. Can these measurement data be effectively used for system earthquake prediction? The paper studied data from Japan and California. They propose the minimum alarm area method to analyze the daily time series of horizontal displacement on the Earth's surface, clearly distinguish the spatial and temporal regions of the location before the epicenter of a strong earthquake. Reflecting abnormal changes in seismic structures and geodynamic processes can be predicted.

Rui L et al. [9] divided earthquake sequences into multiple learning samples and precursor patterns. Based on these patterns and samples, eight dominant features are extracted, while implicit features are also extracted. Based on the attention mechanism, combine explicit and implicit features. A dynamic loss function was designed in the model optimization using a small batch gradient descent optimization method. Adapting to different training data and balancing different categories of algorithms by combining explicit and implicit features is an effective earthquake prediction method.

William et al. [10] wrote a collection of 20 papers. It is divided into seven parts, including historical earthquake phenomena, physical models, precursor earthquakes, surface geochemistry, seismic related atmosphere signals, ionospheric processes, and interdisciplinary earthquake prediction methods. They Believe that earthquake warning can promote building standards. Buildings and facilities that can withstand earthquakes. It can reduce the cost of future earthquakes and reduce the number of injuries and deaths.

Danijel et al. [11] pointed out that CSEP is a global network infrastructure used for prospective

evaluation of earthquake prediction models and algorithms. The global CSEP collaboration has been conducting predictive experiments in various tectonic environments worldwide. The experiment provides a large number of results, providing information for operable earthquake prediction systems and earthquake disaster models. New and surprising insights have been provided on the predictability of earthquakes.

Gualberto et al. [12] explored seismic indicators on the Chilean National Earthquake Service dataset. After fully adjusting these indicators, the accuracy of prediction can be improved. The results indicate that by adjusting the input appropriately, the predictive ability of the classifier is significantly exceeded. Optimize and develop adaptive systems that utilize all available information, discover new metrics to provide more information to the system.

Elshin Oleg et al. [13] introduced Terra Seismic, which can predict most major earthquakes 2-5 months in advance. The geological pattern and pressure accumulation of earthquake development are usually the same. Terra Seismic currently provides earthquake prediction for 25 key earthquake prone areas. Successfully detected approximately 90% of major earthquakes in the past 50 years.

The above literature shows that aftershocks can be predicted. Neural networks are the most suitable method to establish corresponding prediction models.

3. Proposed Method

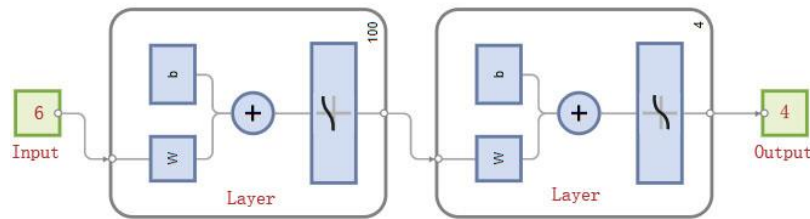


Figure 1: Network Structure

Using a forward propagation network, input the $R \times Q1$ matrix. $Q1$ input vector containing R elements. The output is the $SN \times Q2$ matrix. $Q2$ target vectors containing SN elements. Adopting a three-layer structure as figure1. The hidden layer has 100 neurons, with an input layer width of 7 and an output width of 4. The transfer function takes the S-shaped transformation function Logig and the hyperbolic tangent S-shaped function transig, respectively. Try not to use linear functions. The gradient descent function adopts trainingdx, which is an adaptive learning rate training function. The weight of the input layer is determined by the input, and the hidden layer may come from the input layer. The weights of each layer are updated by the learning function. Training is performed by the training function. Performance is measured by a performance function.

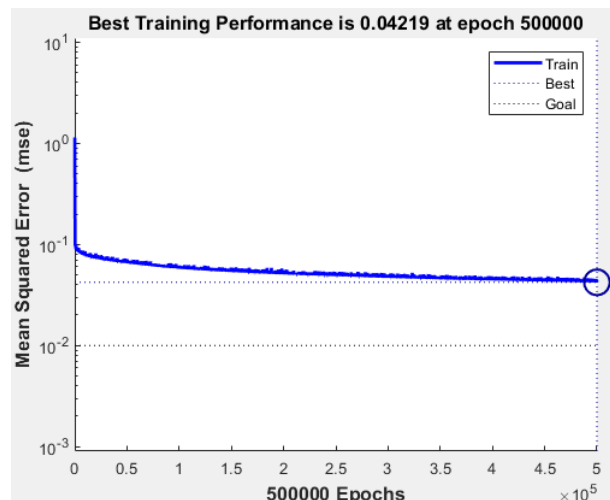


Figure 2: Training State

The learning rate is taken as 0.04219. Display intermediate results every 50 rounds. The target error of neural network training is taken as 0.01. The maximum number of iterations is taken as 50000.

Training State is as figure2.

4. Experiment and Results

The data is from China Earthquake Networks Center and National Seismological Science Data Center (<http://data.earthquake.cn>). We selected the seismic phase data block DPB from the QinghaiMaduo 7.9 magnitude earthquake phase dataset on May 22, 2021 at 02:04. Original data shows in figure4. There are 1351 recorders. Training performance shows as figure 3.

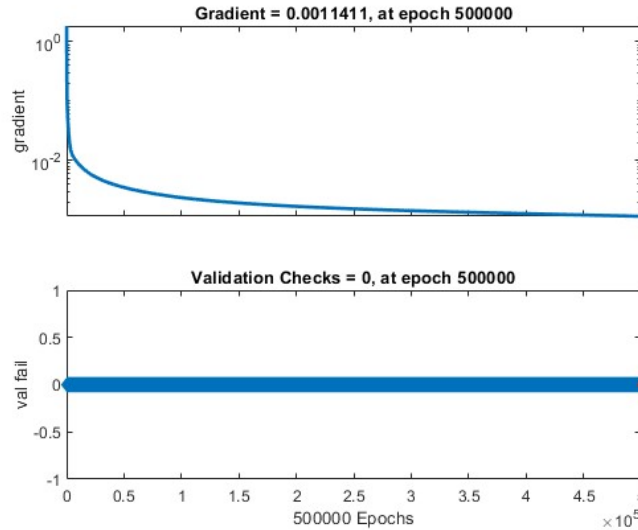


Figure 3: Training Performance

Original data format:

	A	B	C	D	E	F	G
1	02:05:46.10	2.4	6.1	341	3837	11.2	7.4
2	02:05:46.10	2.4	6.1	341	3837	11.2	7.4
3	02:05:46.10	2.4	6.1	341	2971	13.5	
4	02:05:46.10	2.4	6.1	341	2504	7.2	7.4
5	02:05:46.10	2.4	6.1	341	25	1.4	
6	02:05:46.10	2.4	6.1	341	22.3	1.6	6.4
7	02:07:30.00	4.3	6.1	341	22.3	1.6	
8	02:06:06.00	4.3	6.1	341	22.3	1.6	
9	02:06:06.00	4.3	15.1	56	8641	20.9	8
10	02:06:06.00	4.3	15.1	56	4891	19.6	
11	02:06:06.00	4.3	15.1	56	3048	19.1	7.9
12	02:10:57.80	4.7	15.1	56	3048	19.1	
13	02:10:46.40	0.6	15.1	56	3048	19.1	
14	02:10:39.30	3.5	15.1	56	3048	19.1	
15	02:07:53.60	-0.5	15.1	56	3048	19.1	
16	02:07:46.60	-1.9	15.1	56	3048	19.1	
17	02:07:46.60	-1.9	16.4	294	4294	17.1	7.8
18	02:07:46.60	-1.9	16.4	294	3972	15.8	
19	02:07:46.60	-1.9	16.4	294	1781	17.1	7.9
20	02:11:11.50	5.2	16.4	294	1781	17.1	

Figure 4: Original Data

Seismic phase refers to seismic wave groups with different properties or propagation paths displayed on seismic maps. Various seismic phases have different characteristics. Specifically, in terms of arrival time, waveform, amplitude, period, and particle motion mode. The seismic phase characteristics are related to the source, propagation medium, and receiving instrument. These wave groups all have a certain duration. The waveforms of different seismic phases overlap with each other, causing interference, resulting in a complex pattern in the seismic map. One of the tasks of seismology is to analyze and explain the causes and physical meanings of various seismic phases. We use various seismic phase characteristics to determine the basic parameters of earthquakes. Then, we study the mechanical properties of seismic sources, and explore the internal structure of the Earth.

Filter the raw data. Select 7 features. They are: Phase when the seismic phase arrives_ Time, travel time residual Resi, epicenter distance, station azimuth Azi, amplitude Amp, magnitude Mag_ Val and Period. Due to the fact that the dates are on the same day, only hours, minutes, and seconds are taken. For the ease of operation, we subtract the initial time from the time and take the offset as the time

characteristic value. Figure 5 shows the organized data. Data without magnitude removed from the data. We get 856 valid data. We divided data into four categories based on duration. Data of category 4 is the duration more than 20 seconds. Category 3 is in the range from 10 seconds to 20 seconds. Category 2 ranges from 5 seconds to 10 seconds. Category 1 is defined as below 5 seconds.

Organized data format:

1	Phase_time	Resi	Distance	Azi	Amp	Mag_val	Period_val	Period
2	186.16	2.4	6.1	341	3837	7.4	3	11.2
3	186.16	2.4	6.1	341	3837	7.4	3	11.2
4	186.16	2.4	6.1	341	2504	7.4	2	7.2
5	186.16	2.4	6.1	341	22.3	6.4	1	1.6
6	246.06	4.3	15.1	56	8641	8	4	20.9
7	246.06	4.3	15.1	56	3048	7.9	3	19.1
8	386.66	-1.9	16.4	294	4294	7.8	3	17.1
9	386.66	-1.9	16.4	294	1781	7.9	3	17.1
10	751.56	5.2	16.4	294	70	7.3	2	5.4
11	751.56	5.2	16.4	294	6.3	6.9	1	1.2
12	442.56	-2.8	15.2	64	82	6	4	20.1
13	442.56	-2.8	15.2	64	728	7.2	3	14
14	393.46	-4.7	15.2	64	26.9	6.6	2	9
15	393.46	-4.7	15.2	64	0.2	5.3	1	1.7
16	386.36	-2.3	10.1	338	4333	7.6	3	13.5
17	386.36	-2.3	10.1	338	2823	7.9	3	14.7
18	386.36	-2.3	10.1	338	53.6	7	3	11.5
19	386.36	-2.3	10.1	338	0.6	6.2	1	0.9
20	278.46	-2.4	24.7	54	2558	7.9	3	14.2

Figure 5: Computed Data

We divided the dataset into two subsets. 80% of the dataset is a training subset, and 20% is a testing subset.

Input data shows in figure 6.

1	2	3	4	5	6	7	8	9	10	11	12
-0.8429	-0.8429	-0.8429	-0.8429	-0.7923	-0.7923	-0.6736	-0.6736	-0.3656	-0.3656	-0.6264	-0.6264
0.2043	0.2043	0.2043	0.2043	0.4086	0.4086	-0.2581	-0.2581	0.5054	0.5054	-0.3548	-0.3548
-0.9971	-0.9971	-0.9971	-0.9971	-0.9912	-0.9912	-0.9903	-0.9903	-0.9903	-0.9903	-0.9911	-0.9911
0.9044	0.9044	0.9044	0.9044	-0.6985	-0.6985	0.6400	0.6400	0.6400	0.6400	-0.6535	-0.6535
-0.7923	-0.7923	-0.8645	-0.9988	-0.5323	-0.8350	-0.7676	-0.9036	-0.9962	-0.9997	-0.9956	-0.9606
0.2222	0.2222	0.2222	-0.3333	0.5556	0.5000	0.4444	0.5000	0.1667	-0.0556	-0.5556	0.1111

Figure 6: Input Data

Train the data using BP network and CNN network respectively. Predict the duration of aftershocks from input. Figure 7 shows the input vectors.

1	2	3	4	5	6	7
186.1600	2.4000	6.1000	341	3837	7.4000	3
186.1600	2.4000	6.1000	341	3837	7.4000	3
186.1600	2.4000	6.1000	341	2504	7.4000	2
186.1600	2.4000	6.1000	341	22.3000	6.4000	1
246.0600	4.3000	15.1000	56	8641	8	4
246.0600	4.3000	15.1000	56	3048	7.9000	3
386.6600	-1.9000	16.4000	294	4294	7.8000	3
386.6600	-1.9000	16.4000	294	1781	7.9000	3
751.5600	5.2000	16.4000	294	70	7.3000	2
751.5600	5.2000	16.4000	294	6.3000	6.9000	1
442.5600	-2.8000	15.2000	64	82	6	4
442.5600	-2.8000	15.2000	64	728	7.2000	3
393.4600	-4.7000	15.2000	64	26.9000	6.6000	2
393.4600	-4.7000	15.2000	64	0.2000	5.3000	1
386.3600	-2.3000	10.1000	338	4333	7.6000	3
386.3600	-2.3000	10.1000	338	2823	7.9000	3
386.3600	-2.3000	10.1000	338	53.6000	7	3
386.3600	-2.3000	10.1000	338	0.6000	6.2000	1
278.4600	-2.4000	24.7000	54	2558	7.9000	3
278.4600	-2.4000	24.7000	54	1446	8	3

Figure 7: Input Vector

Output data shows in figure 8.

	1	2	3	4
1	0	0	1	0
2	0	0	1	0
3	0	1	0	0
4	1	0	0	0
5	0	0	0	1
6	0	0	1	0
7	0	0	1	0
8	0	0	1	0
9	0	1	0	0
10	1	0	0	0
11	0	0	0	1
12	0	0	1	0
13	0	1	0	0
14	1	0	0	0
15	0	0	1	0
16	0	0	1	0
17	0	0	1	0
18	1	0	0	0
19	0	0	1	0
20	0	0	1	0

Figure 8: Output of the Neural Network Model

After training, we test the model with new dataset. Figure 9 shows the test results.

	1	2	3	4	5	6	7	8
1	657.3600	-4.3000	25.9000	58	1033	7.8000	3	3
2	1.0559e+03	2.6000	25.9000	58	34.5000	7.1000	2	2
3	1.0559e+03	2.6000	25.9000	58	0.2000	5.5000	1	1
4	588.1600	1.4000	27.6000	56	3125	7.9000	3	3
5	588.1600	1.4000	27.6000	56	1669	8.1000	3	3
6	645.0600	-4.5000	27.6000	56	9.2000	6.8000	1	1
7	645.0600	-4.5000	27.6000	56	0.2000	5.6000	1	1
8	641.8600	-0.9000	17.3000	93	5285	7.9000	3	3
9	641.8600	-0.9000	17.3000	93	5267	8.1000	3	3
10	765.8600	-0.8000	17.3000	93	45.9000	7.1000	1	2
11	765.8600	-0.8000	17.3000	93	0.7000	5.6000	1	1
12	454.6600	-1.7000	15.9000	107	5563	7.9000	3	3
13	454.6600	-1.7000	15.9000	107	4185	8	3	3
14	697.9600	3.3000	15.9000	107	36.2000	7.1000	1	2
15	697.9600	3.3000	15.9000	107	1.1000	6	1	1
16	398.1600	-0.6000	22.4000	47	5769	8.1000	3	3
17	398.1600	-0.6000	22.4000	47	2591	8.3000	3	3
18	955.5600	2.1000	22.4000	47	34.9000	7.2000	1	1
19	955.5600	2.1000	22.4000	47	1.8000	6.4000	1	1
20	553.0600	-0.2000	8.6000	159	8067	7.6000	4	4

Figure 9: Test Result

5. Evaluation

From the rendering, it can be seen that the built-in trainingdx training function has a large output value of the entire model, resulting in significant errors that make the model unusable. In addition to output constraints and setting an output upper limit, the training is good and the approximation effect is good. This indicates that the improved constrained trainingdx training function can handle similar situations where the model output value is too large or too small, resulting in better results.

Mean square error (MSE) is a measure that reflects the degree of difference between the estimator and the estimated quantity. Let t be the overall parameter determined based on the sample θ . An estimator of $(\theta - t)$ is mathematical expectation of 2. It is called the mean square error of the estimator t . It is equal to $\sigma^2 + b^2$, where σ^2 and b are the variance and bias of t .

Consistent estimation (or consensus estimation) is the standard for evaluating estimators in large samples. When the sample size is not large, people tend to use small sample based evaluation criteria. In this case, variance is used for unbiased estimation and mean square error is used for biased estimation.

Generally, when the sample size is fixed, the criterion used to evaluate the quality of a point estimation is always a function of the distance between the point estimation and the true value of the parameter. The most commonly used function is the square of the distance. Due to the randomness of the estimation, the expectation of this function can be obtained, which is the mean square error given by the following equation:

6. Conclusions

Predicting the duration of aftershocks is feasible on the premise that an earthquake has already occurred. In different regions, aftershock warning mechanisms can be established based on the geological conditions of the region. Expanding to larger regions and for a longer period of time, based on existing earthquakes, predicting future earthquakes should also be feasible. This is the direction that this article strives to explore.

Earthquake prediction is not the goal. The goal of this study is to provide data support for earthquake relief. In the event of an earthquake, minimize personnel and property damage as much as possible.

In future research, we will delve deeper into the use of deep learning algorithms and construct new aftershock prediction models using typical residual models. We use all parameters in seismic phase data to make detailed predictions of aftershocks.

References

- [1] Helene J, Tiziana R, Caroline B, Cliodhna O, et al. *Stigma in Science: The Case of Earthquake Prediction* [J]. *Disasters*, 2018, 42(1): 81-100.
- [2] Asmae Berhich, Fatima-Zahra Belouadha, Mohammed Issam Kabbaj. *Lstm-Based Models for Earthquake Prediction*[C]. *International Conference on Networking*, 2020: 46:1-46:7.
- [3] Saba Sehrish, Faraz Ahsan, Sajjad Mohsin. *BAT-ANN based earthquake prediction for Pakistan region*. [J]. *Soft computing*, 2017, 21(19): 5805-5813.
- [4] Chang-Qing Li, Xiao-Ping Zhou. *Laboratory Earthquake Prediction of Granite* [J]. *Tribology International*, 2021, 160.
- [5] G. Molchan, L. Romashkova, A. Peresan. *On some methods for assessing earthquake predictions* [J]. *Geophysical journal international*, 2017, 210(3.0): 1474-1480.
- [6] C. Cattania, P. Segall. *Crack Models of Repeating Earthquakes Predict Observed Moment-Recurrence Scaling* [J]. *Journal of geophysical research. Solid earth*, 2019, 124(1): 476-503.
- [7] Qianlong W, Yifan G, Lixing Y, Pan L, et al. *Earthquake Prediction based on Spatio-Temporal Data Mining: An LSTM Network Approach*[J], *IEEE Transactions on Emerging Topics in Computing*, 2020, 8(1): 148-158.
- [8] Valeri Gitis, Alexander Derendyaev, Konstantin N. Petrov. *Analyzing the Performance of Gps Data for Earthquake Prediction* [J]. *Remote sensing*, 2021, 13(9): 1842.
- [9] Rui L, Xiaobo L, Shuowei L, Haipeng Y, Jianfeng Q, Lei Z, et al. *Dlep: A Deep Learning Model For Earthquake Prediction*[C]. *IEEE International Joint Conference on Neural Network*, 2020: 1-8.
- [10] William H. Prescott. *Pre-earthquake processes: a multidisciplinary approach to earthquake prediction studies* [J]. *International geology review*, 2019, 61.0(16): 2080.0-2082.0.
- [11] Danijel S, M. W, Warner M, Thomas H J, Yosihiko O, David D J, Sum M, David A R, Matthew C G, Naoshi H, Maria L, Philip J M, Anne S, Matteo T, Stefan W, Jeremy D Z, Jiancang Z, et al. *The Collaboratory for the Study of Earthquake Predictability: Achievements and Priorities* [J]. *Seismological Research Letters*, 2018, 89(4): 1305-1313.
- [12] Gualberto A C, Francisco M, A. M, Jorge R, et al. *A sensitivity study of seismicity indicators in supervised learning to improve earthquake prediction*. [J]. *Knowledge-based systems*, 2016, 101(C): 15-30.
- [13] Elshin Oleg, Tronin Andrew A. *Global Earthquake Prediction Systems*, *arXiv preprint arXiv*, 2020, 2003(07593).