Scholar Publishing Group
Machine Learning Theory and Practice
https://doi.org/10.38007/ML.2021.020306
ISSN 2790-0983 Vol. 2, Issue 3: 44-51

Landslide Risk Assessment Based on Artificial Neural Network Model

Khadijah Ragab*

Chandigarh University, India

**corresponding author*

Keywords: Landslide Risk, BP Neural Network, Influencing Factors, Neural Model

Abstract: According to the public data, the overall trend of economic loss caused by landslide disaster is growing year by year, and improving the prediction level of landslide disaster can effectively reduce the loss caused by landslide disaster. Therefore, the study of landslide prediction and prediction has important practical significance, and landslide displacement prediction is an important content of landslide prediction and prediction. This paper mainly studies landslide risk assessment based on artificial neural network model. In this paper, based on the results of field investigation, literature summary and expert experience, collected samples of 100 landslides, analyzes the distribution and the main influence factors of landslide, and establishes the surface elevation, and vegetation index, cutting slope, annual average rainfall, surface density, overlying soil types for the impact factor of regional landslide hazard assessment system. According to the requirement of BP neural network (BPNN) for input data, the landslide sample data processing method which is suitable for Sichuan landslide risk assessment system is established.

1. Introduction

Landslide refers to the phenomenon that the rock and soil mass or various deposits on the slope slide along a certain sliding surface under the action of its own gravity or external force. Landslides are mainly affected by various internal and external conditions such as local stratigraphic lithology, topography and landform, meteorology and hydrology, and human engineering activities [1-2]. Landslides mainly occur in the dangerous terrain of mountainous areas and hilly areas, which has a large amount of land, especially in the southwest region is relatively high, in a small area with obvious elevation changes, complex topography, steep mountain, history experienced many large earthquakes, leading to frequent geological disasters in the region [3-4]. With the development of economy and the growth of population, the country's landslide monitoring is no longer only satisfied with the front and back of the house, and it needs to monitor the landslide situation in a larger area. By monitoring and analyzing the risk of slope, we can effectively understand which regions and periods have a high probability of landslide, as well as the possible energy levels of landslide, so as to make corresponding countermeasures. Therefore, it is particularly important to carry out risk analysis of landslide [5-6].

Landslide risk performance reflects an important attribute of geological disasters, is an important step of risk assessment of geological disasters, is to judge the liability of landslides after further research, is to determine the main research areas in the future the possibility of the probability of landslide in space and time, and the future of the scale of landslide, the movement mode and intensity of landslide, For landslide risk assessment, the main contents should include landslide instability mechanism, the probability of dormant landslide or unstable slope transforming into landslide, and the acceleration of displacement of the landslide with deformation [7-8]. The landslide risk assessment of the study area is mainly to extract and analyze the landslide risk probability index, which can be obtained by the product of time, space and strength. Usually old said liability evaluation is equivalent to spatial probability, probability and time need to conjoint analysis with external factors, through these factors will change time compared with the time of the landslide, time probability can be obtained, probabilities can be based on the strength of landslide volume, area, frequency analysis, such as speed, activity, and then estimate the strength of the landslide [9-10]. Some scholars combined GIS with statistical analysis model to analyze the relationship between regional geological conditions and landslide occurrence under different conditions, and obtained risk assessment results on the basis of this analysis [11]. International scholars have adopted various methods of landslide hazards are studied, but there are still some shortage has not been solved, such as the current most of the research is based on historical data, but the outside conditions is a change in the study area, the cause is due to changes in the process of risk, how to combining the existing situation and the historical data is not too much study. However, with the development of technology and the deepening of research, these unsolved problems will be solved one by one [12].

The establishment of a set of scientific and reasonable landslide evaluation index system and the landslide susceptibility and risk evaluation scheme suitable for urban characteristics will help to identify the prone areas of landslides, reduce the potential loss caused by landslides, and provide scientific basis for national economic construction and national personal safety guarantee.

2. Landslide Risk Assessment Model Based on BPNN

2.1. Artificial Neural Network

Artificial neural network (Ann) is a computational structure that simulates biological process and reflects some characteristics of human brain on the basis of modern neurobiological research. At first, due to the lack of network structure and learning ability, neural networks are only used to solve simple linear problems. With years of optimization and improvement, the characteristics and functions of neural networks are more and more close to the human brain, such as the implementation of parallel operation in information processing, the establishment of nonlinear mapping, the ability of associative memory and classification recognition, and so on. The three major factors that determine the performance of neural networks are the information processing characteristics of neural networks, network topology structure and learning rules, which are also the classification standards of neural networks [13-14].

At the beginning of modeling, scholars established mathematical models through mathematical methods and mechanical theories, and achieved the ability to solve problems through the process of model, data operation and final solution. However, in the face of many procedures and complex problems, it is difficult to establish mathematical models in this way [15]. Based on this phenomenon, artificial neural network provides a good solution, which omits a large number of mathematical formulas and theories and directly establishes a model through sample data [16].

Neurons are the basis of neural networks, which mimic the way the brain is organized and are made up of three parts called "dendrites", "axons" and "cell bodies". Among them, dendrites are the input layer of cells in the brain, which senses the signals transmitted through a large number of terminal factors; cell bodies, as the processing central system, have the ability to process data; axons, as the output layer of brain cells, transmit the data processed by cell bodies to the next neuron. Neurons relay way into a state of excitement and inhibition, when entering the urge to make potentials in the membrane potential reaches a certain value, the neurons in the excited state, in accordance with the "dendritic cell body, axons", this value is about 40 mv, when did not reach the numerical input impulse, inhibitory neurons, Does not transmit these nerve impulses [17-18]. The neuron model is shown in Figure 1.

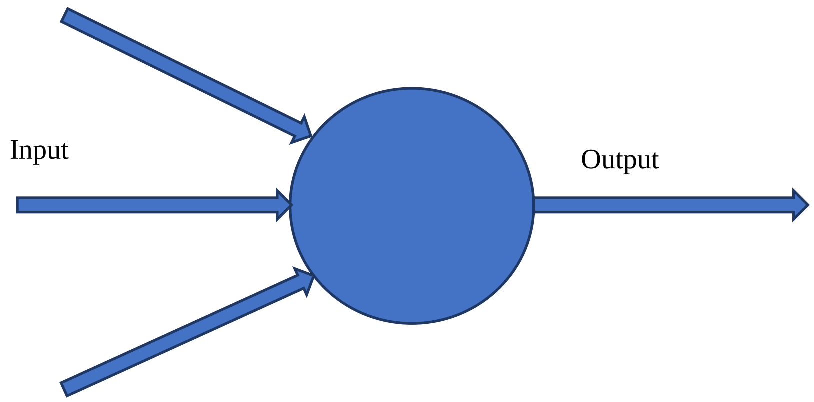

Figure 1. Model of neuron

Artificial neuron is to simulate the transmission mode of cell neuron, which simulates the function of neuron and simplifies the complex system of neuron. The figure shows the structure of the neuron system. The relationship between input and output of neurons can be expressed as follows:

$$I_i = \sum_{j=1}^{n} w_{ij} x_j - \theta_i \tag{1}$$

$$y_i = f\left(I_i\right) \tag{2}$$

In the formula, xj (j=1,2,3... ,n) is the information transmitted from each neuron organization, wij represents the weight transmitted between neuron i and neuron j.

A hierarchical network divides neurons into input layers, hidden layers and output layers according to their functions. There is usually no information flow between neurons in the set layer, but with further exploration of the structure, researchers found that the performance of neural networks for specific purposes can be effectively improved if the input layer and output layer are connected or the hidden layer is interconnected.

2.2. BPNN Evaluation Model

In general, the landslide risk assessment system determines the structure of BPNN, that is, the geometric topological relationship of BPNN and the number of nodes. The reasonable number of hidden layers and nodes can effectively improve the efficiency of landslide risk assessment. In general, more complex problems require more nodes. However, the research shows that with the increase of the number of hidden layers and nodes, the computation speed of BPNN will decrease, and the initial parameters of the network (initial weight and initial threshold) will have an increased impact on the results.

The logical topology structure of BPNN is the important foundation of all its functions. The logical topology structure should be established according to specific problems. For the study of landslide risk assessment, landslide risk assessment system directly determines the logical topology structure of BPNN.

The logical topology structure of BPNN mainly refers to the node Settings of input layer. For the input layer, this paper takes the influence factor as the input factor, so that the six influence factors in the landslide risk assessment system, such as surface elevation, vegetation index, slope, average annual rainfall, surface cutting density and overlying soil type, correspond one to one with the six nodes in the input layer of BPNN. For the hidden layer, the number of nodes in the hidden layer is set according to the trial-and-error formula 3, and the number of nodes in the hidden layer is 6.

$$m = \sqrt{N + L} + a \tag{3}$$

The training samples are used for adaptive learning and data mining by BPNN. The weights and thresholds in BPNN are changed, and the global optimal weights and thresholds are obtained. Test samples are used to test whether the prediction results of trained BPNN for general data (non-training data) meet the expectation. Training samples and test samples are usually determined in accordance with the following principles: training samples and test samples can fully reflect the characteristics of the data to avoid essential differences between training samples and test samples; The data scale, quality and acquisition method of training and test samples should remain the same.

In this paper, 100 landslide samples were determined through field investigation and literature summary of a landslide in a province. Now these 100 landslide samples are divided into training samples (75 samples) and test samples (25 samples) according to the ratio of 3:1.

In the modeling process, some controlling parameters need to be set in advance. This paper refers to the parameter selection in the existing research results to set the parameters of the above model. For BPNN, the parameters to be set include momentum factor, maximum training times, learning rate and error limit. The values of specific parameters are shown in Table 1.

Table 1. BPNN parameter setting

	Momentum factor	Maximum training times	Learning rate	Error limits
Value	0.8	10000	0.03	0.01

3. Landslide Risk Assessment Experiment

Based on the model evaluation, the landslide risk index of the study area needs to be transformed into a zoning map according to the size of the index to provide services for regional landslide disaster management intuitively. The commonly used partitioning methods are: equal distance method, expert experience method, natural discontinuity point method, etc. Scholars have made a thorough comparison of the regional geological hazard evaluation methods, and found that under the same geological conditions, the natural discontinuity point method makes the most ideal evaluation results. There are significant differences among the regions divided by the natural discontinuity point method, and the differences within each category are very small, which is similar to cluster analysis from the results. This method is now the most mature and extensive method in geological hazard assessment.

One of the criteria for evaluating the performance of machine learning models is the receiver operating characteristic (ROC) curve. ROC curves can be generated by the definition of true positive rate (TPR) and false positive rate (FPR) as axes. The performance of different models is measured by the area under the ROC curve, which is defined as the AUC value. The value of AUC ranges from 0.5 to 1.0. The better the model performance, the higher the AUC value. Other statistical indicators commonly used for auxiliary verification include Recall, Precision, and ACC. These measures include true positive (TP, the number of disaster points correctly predicted), false positive (FP, the number of disaster points predicted as non-disaster points), true negative (TN, the number of disaster points predicted as non-disaster points), Correct prediction of the number of non-disaster points), false negative (FN, the prediction of non-disaster points as the number of disaster points).

4. Analysis Of Evaluation Results

After the training of each model is completed, 25% of the divided test set, namely 25 data samples, is used to test the performance of each model combined with auxiliary verification indicators.

Table 2 lists the values of common metrics used to assist in evaluating model performance.

Table 2. Auxiliary validation of statistical standards

	Recall	Precision	ACC
SVM	0.81	0.83	0.822
ANN	0.92	0.93	0.917
BPNN	1	0.96	0.987

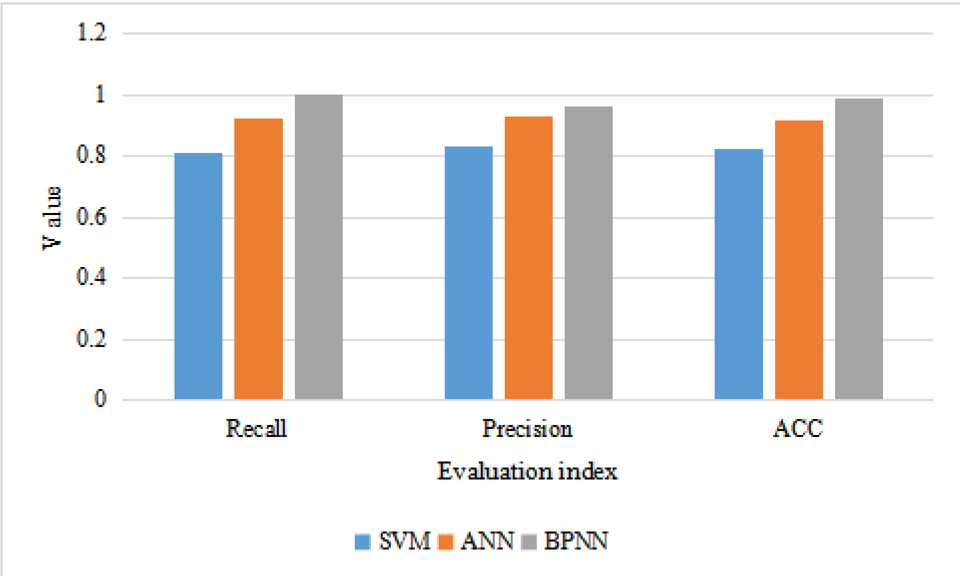

Figure 2. Statistics of evaluation indexes of different models

As can be seen from the table, in terms of recall, BPNN model achieves the maximum value (100%), followed by ANN model (92%) and SVM model (81%).

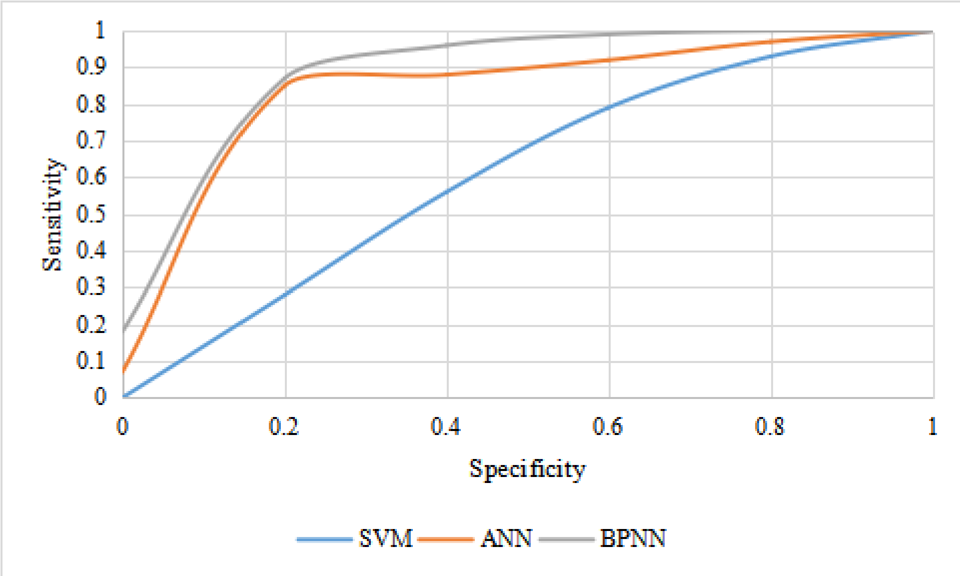

Figure 3. ROC curve of landslide risk assessment based on machine learning model

The ROC curve analysis of the machine learning model is shown in Figure 2, and the AUC value of each model is also used as the standard to evaluate various models. Comparing the AUC values of each model, it is found that the prediction result of SVM model is not ideal (AUC<0.8). The ANN model performs well in the spatial prediction of geological disasters (AUC>0.8). Among all machine learning models, BPNN model has the best generalization ability, and its AUC value is 0.957.

5. Conclusion

Landslide hazard mapping is often used to guide landslide early warning and post-disaster risk analysis, which is also the first step of landslide disaster management. This paper takes a province as the research area, collects 100 landslide samples in a province through field investigation and literature summary, forms a landslide sample database, establishes a landslide risk assessment system suitable for a province, and constructs a landslide risk assessment method based on BPNN and its optimization algorithm. The applicability of the three models is compared and analyzed. In this study, there are the following deficiencies and problems, which are worthy of in-depth analysis and discussion in future research: The BPNN training usually require a large amount of data samples to obtain ideal accuracy, due to the particularity of the landslide, sample collection requires a lot of manpower and material resources, this paper within the scope of the ability to use the 100 landslide training samples, the number of samples, there may be the result of the model, in the future should continue to expand the landslide database, Conduct in-depth analysis and research.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Asghari S, Balvasi I. A Comparison of Artificial Neural Network Model with Fuzzy logic model In Landslide Hazard Assessment. Quantitative Geomorphological Research, 2018, 7(2): 158-182.

[2] Rajabi A M, Khodaparast M, Mohammadi M. Earthquake-induced landslide prediction using back-propagation type artificial neural network: case study in northern Iran. Natural Hazards, 2021, 110(1): 679-694. https://doi.org/10.1007/s11069-021-04963-8

[3] Hammad M, Van Leeuwen B, Mucsi L. Integration of gis and advanced remote sensing techniques for landslide hazard assessment: A case study of northwest syria. Isprs Annals of the Photogrammetry Remote Sensing and Spatial Information Sciences, 2020, 6(6): 27-34. https://doi.org/10.5194/isprs-annals-VI-3-W1-2020-27-2020

[4] Lee J H, Kim H, Park H J, et al. Temporal prediction modeling for rainfall-induced shallow landslide hazards using extreme value distribution. Landslides, 2021, 18(1): 321-338. https://doi.org/10.1007/s10346-020-01502-7

[5] Aslam B, Zafar A, Khalil U. Development of integrated deep learning and machine learning algorithm for the assessment of landslide hazard potential. Soft Computing, 2021, 25(21): 13493-13512. https://doi.org/10.1007/s00500-021-06105-5

[6] Ngo P T T, Panahi M, Khosravi K, et al. Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. Geoscience Frontiers, 2021, 12(2): 505-519. https://doi.org/10.1016/j.gsf.2020.06.013

[7] Huqqani I A, Tien T L, Mohamad-Saleh J. Landslide hazard analysis using a multilayered approach based on various input data configurations. Geosfera Indonesia, 2021, 6(1): 20-39. https://doi.org/10.19184/geosi.v6i1.23347

[8] Panchal S, Shrivastava A K. Landslide hazard assessment using analytic hierarchy process (AHP): A case study of National Highway 5 in India. Ain Shams Engineering Journal, 2021, 13(3): 101626. https://doi.org/10.1016/j.asej.2021.10.021

[9] Thennavan E, Pattukandan Ganapathy G. Evaluation of landslide hazard and its impacts on hilly environment of the Nilgiris District-a geospatial approach. Geoenvironmental Disasters, 2020, 7(1): 1-14. https://doi.org/10.1186/s40677-019-0139-3

[10] Hejazi A, Rezaeimoghaddam M, Naseri A. Landslide hazard zoning using artificial neural network models and TOPSIS downstream of Sanandaj Dam. Hydrogeomorphology, 2020, 7(24): 65-82.

[11] Karami M. Combination AHP and Neural Network Model to landslide Hazard Zonation (Case Study city of Bijar). Journal of Engineering Geology, 2018, 12(1): 153-182.

[12] Lee S. Current and future status of GIS-based landslide susceptibility mapping: a literature review. Korean Journal of Remote Sensing, 2019, 35(1): 179-193.

[13] Lucchese L V, de Oliveira G G, Pedrollo O C. Attribute selection using correlations and principal components for artificial neural networks employment for landslide susceptibility assessment. Environmental monitoring and assessment, 2020, 192(2): 1-22. https://doi.org/10.1007/s10661-019-7968-0

[14] Nguyen H D, Pham V D, Nguyen Q H, et al. An optimal search for neural network parameters using the Salp swarm optimization algorithm: a landslide application. Remote Sensing Letters, 2020, 11(4): 353-362. https://doi.org/10.1080/2150704X.2020.1716409

[15] Ghorbanzadeh O, Blaschke T, Gholamnia K, et al. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. Remote Sensing, 2019, 11(2): 196. https://doi.org/10.3390/rs11020196

[16] Jacinth Jennifer J, Saravanan S. Artificial neural network and sensitivity analysis in the landslide susceptibility mapping of Idukki district, India. Geocarto International, 2021, 37(19): 5693-5715. https://doi.org/10.1080/10106049.2021.1923831

[17] Mallick J, Alqadhi S, Talukdar S, et al. Risk assessment of resources exposed to rainfall induced landslide with the development of GIS and RS based ensemble metaheuristic machine learning algorithms. Sustainability, 2021, 13(2): 457. https://doi.org/10.3390/su13020457

[18] Nayek P S, Gade M. Artificial neural network-based fully data-driven models for prediction of newmark sliding displacement of slopes. Neural Computing and Applications, 2021, 34(11): 9191-9203.