Journal of Energy & Environment

Journal homepage: journal.uniten.edu.my/jee

ISSN: 1985-7462

Classification of slope stability based on Artificial Neural Network and Naive Bayes

Ashanira Mat Deris [1], Badariah Solemon[1]

[1] Institute of Infrastructure Energy, Universiti Tenaga Nasional, Jln IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia

Keywords

slope stability prediction
Artificial Neural Network
Naive Bayes

Abstract

Estimating the slope stability is a crucial and critical process as the stability of soil slopes not only depends on the geological factors, but also depends on the physical and topography factors. Due to its challenging process, this study attempts on the prediction of slope stability using machine learning (ML) methods which are Artificial Neural Network (ANN) and Naive Bayes (NB) classifier using the historical slope cases worldwide. The prediction models were developed based on six input factors namely "unit weight, internal friction angle, cohesion, slope angle, slope height and pore pressure ratio" and factor of safety (FOS) as the output factor. The slope data was collected from the previous studies and divided into 70% training and 30% testing datasets for both models. The classification process of ANN and NB were implemented using python programming and the result shows that ANN prediction model gives better prediction result with accuracy of 95%, compared to NB with 84% of accuracy. The prediction of slope stability is one of the critical interests during the slope design process. Hence, this study may served as a benchmark study for the application of ANN and NB machine learning methods in predicting slope stability.

1. INTRODUCTION

The geotechnical engineering process such as estimating the slope stability is quite challenging because of its complexity and it deals with the natural data such as strength of the soil, the pressure of the pore water and the geometry of the slope. To measure the stability of slopes, FOS is used to determine whether slope is stable or failed. The FOS is generally described as the ratio of the current soil strength to the needed minimum shear strength to prevent the failure of its structure [1].The higher value of FOS shows that the slope is stable, while the lower value of FOS shows that the slope fails. Precise prediction of FOS, including the stability and efficiency of slope is not a simple task as it is difficult to determine mechanical properties, the degree of influence and the relationship between the FOS factors. Furthermore, the slope stability assessment often involved with many sources of uncertainties factors [2]. Based on the reviews, it was found that many researchers has studied slope stability and its associated FOS using various method including limit equilibrium method (LEM), finite equilibrium method (FEM), gravity increase method (GIM), strength reduction method (SRM), rigid element method, distinct element method, etc. [3]–[8]. Among these methods, LEM and FEM was found to be the most widely used and successful methods for the analysis of slope stability and determining the FOS value.

LEM is also well-known as "method of slices", uses the principle of slicing the failure mass to determine the FOS. It deals with the assumptions on the distribution of inter-slice force. The slicing methods consist of the "Fellenius [9], Bishop[10], Morgenstern and Price [11] , Spencer [12], and Janbu [13]" method. FEM, which is an alternative way, is an adaptable tool for analyzing stability of slope by alleviating assumptions on slope failure surface as required in LEM [14]. The advantages of LEM and FEM have made them the most preferred conventional method for solving slope stability analysis by the previous researchers [15]–[17]. However, these conventional methods have their own limitations such as LEM does not guarantee the effectiveness in handling various geometry or material variation with the large number of assumption [18]. Furthermore, FEM often criticized for its intensive computational power required and having lack computational efficiency for small probability levels [19].

Hence, the introduction of intelligent computational methods such as ML approaches serves as the alternative tools to analyze the slope stability, including the prediction of slope safety [20]. The interest of using ML in geotechnical engineering has been growing for the last few decades.

This is because by using ML, the prior knowledge of a particular model form is not required and the methods also possess flexible capability for the nonlinear modeling [21]. The most common ML methods used for estimating slope safety includes support vector machine (SVM), artificial neural network (ANN) and Naïve Bayes (NB) algorithm. SVM is an efficient ML method based on the Structural Risk Minimization (SRM) concept to construct decision planes to establish decision boundaries. SVM has been widely used in many fields including slope stability analysis [22].

ANN is a powerful ML method that imitates the biological neurons of the human brain to construct a solution for a problem. It has demonstrated the ability to predict a complex model and has been commonly used by previous researchers to evaluate slope stability. Bui et al [22] employed several ML methods including ANN, SVM, multi linear regression (MLR) and Gaussian process regression (GPR) to predict the FOS against the slope failure. 630 slope cases dataset with four input factors namely cohesion, setback distance ratio, slope angle and applied surcharge was divided into training and testing datasets. The prediction result was validated and it was found that ANN model through multilayer perceptron (MLP) network give the highest success rate with R2 is 0.9939. Ray et al, [23] predicted FOS of soil slope in Himalayan Region using ANN models. Two ANN models with the different number of input factors were develop to predict the FOS value. The first model consists of eleven input factors while second model consists of eight input factors, which are the significant factors obtained from the correlation analysis. The result found that both ANN models give good performance of prediction. However, the model with the significant factors has outperformed the model with all the factors listed by giving the higher R2 value. ANN was proven to be an excellent prediction tool by Mamat et al. [24], Qian et al. [25], and Chakraborty and Goswami [26].

NB is one of the commonly used ML methods that use a probability model described based on the Bayes theory to estimate the probability of a new observation belongs to a predefined group [27]. NB is applied in various fields including slope stability prediction. Feng et al [20] has employed NB classifier to predict the FOS value for the circular failures of landslides. The model was developed from six input factors namely cohesion, slope height, slope angle, unit weight, friction angle and pore water pressure, with the incomplete 69 slope cases dataset. The result found that the NB classifier model is capable to predict the slope stability with high accuracy and applicability. Tsangaratos and Ilia [27] has compared NB classifier with logistic regression in landslide susceptibility analysis for slope cases in Epirus, Greece. Seven were used as the input factors derived from 116 sites. The models were analyzed and compared based on the area accuracy and under the curves (AUC) area value. The result found that NB has outperformed logistic regression by giving the higher accuracy and AUC with 87.50% and 0.875, respectively. He et al [14] develop ML models namely NB, radial basis function (RBF) classifier and RBF network for landslide spatial modeling of Longhai, China. 14 conditioning factors was used as an input parameters from 16 of landslides were divided into 70% training and 30% testing datasets. The result shows that NB model give the acceptable prediction result with area under the receiver operating characteristic (AUROC) is 0.872.

From the review of the previous researchers, it shows that ANN and NB are able to affectively predict the slope stability by giving the good prediction results. In the current work, the ANN and NB are applied to predict the slope stability using classification model.

2. METHODOLOGY

A. Artificial Neural Network

ANN is one of the widely used supervised ML methods that inspired by the biological neurons of human brain [28]. The behaviour and the structure of biological neurons are adapted into the ANN model so that the model has the ability to learn, to adapt of changes and also imitate the human thought practice with little interaction from human [29]. The basic structure of ANN consists of input layer, hidden layer and output layer. Each layer contains of neurons that are connected to each other through weight and bias. Multilayer perceptron (MLP) is the most widely used ANN for solving prediction problem, either classification or regression based. The MLP is essentially used to identify the mathematical relationships between various variables, with the consideration of one or more activation functions [30]. Suppose the neurons have n inputs (for this case n is 6 that refers to the six slope factors) and computes the output y based on Equation (1) as follows:

$$y = f(\sum_{i=0}^{n} w_i p_i + b) \tag{1}$$

where f is the activation function, p_i is the input at i-th, w_i is the weight at i-th and b is the bias. Figure 1 shows the architecture of ANN.

This study implements MLP with back propagation neural network due to its suitability and efficiency of the network. Six neurons in the input layers refers to the six slope factors (γ, c . φ, β, H and ru). The hit and trial process specifies the hidden layers and the neurons of the hidden layers. Figure 1 shows the architecture of ANN.

B. Naive Bayes

NB is one of the supervised machine learning methods based on conditional independence assumption. NB classifier is capable to learn and to predict the output from the incomplete information. The structure of NB is pre-defined where no learning structure is needed. NB also proven to be suitable and efficient when dealing with large number of datasets. Furthermore, only a little amount of training data is required to predict the necessary parameters [20].

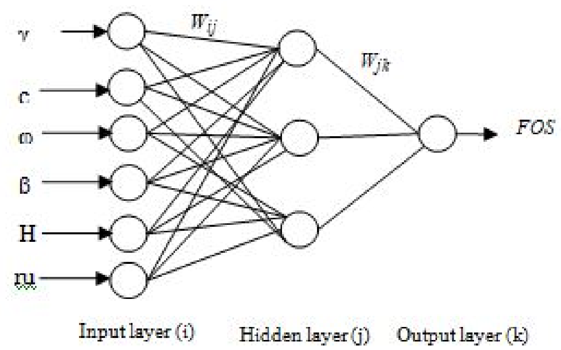

Figure 1. Architecture of ANN

Suppose x_i is the factors affecting slope stability for n attributes where $i = 1, 2, \ldots, n$ and y_j is the two outcomes of the slope stability; stable or failed where stable = 1 and failed = 0. NB estimates the probability $P(y_j/x_i)$ for all possible output class given as in Equation (2) as follows [14]:

$$\begin{aligned} y &= \arg\max P(y_i) \quad \textstyle\prod_{i=1}^{n} P(x_i/y_i) \\ y_i &= \{stable, failed\} \end{aligned} \tag{2}$$

where $P(y_i)$ is the initial probability of y_i, $p(x_i|y_i)$ is the posterior probability, calculated as follows:

$$p(x_i|y_i) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} \tag{3}$$

where σ is the standard deviation of x_i and μ is the mean.

C. Case Study

To develop prediction model of ANN and BN, 148 slope cases were collected from previous studies of [31]–[34]. Six factors affecting slope stability were investigated namely "unit weight, internal friction angle, cohesion, slope angle, slope height and pore pressure ratio" and FOS as the output factor which is classified as 1 for "stable" and 0 for "failed". Table 1 shows the basic statistic of the slope cases.

Table 1. Basic statistical of the slope cases

Slope parameters	Statistic		
	Minimum	Maximum	Standard Deviation
Unit weight (γ)	13.97	31.3	4.0193
Cohesion, (c)	4.95	300	47.1177
Internal friction angle (φ)	0	45	10.9554
Slope angle (β)	16	59	10.1382
Slope height (H)	3.6	511	138.2752
Pore water pressure ratio (ru)	0	45	3.6844

To develop ANN and NB prediction models, the whole datasets need to be divided into two new subsets which are training and testing data. This is to ensure the generalization capability of the datasets. Basically, the training data is used to train the model and tuning the hyper-parameters while testing data is used to test its generalization capacity for the prediction function. This study divides the datasets into 70:30 where 70% of the data is used for training and 30% is used for testing.

3. RESULT AND DISCUSSION

The developed prediction models utilized statistical analysis criteria based on accuracy of the prediction result and receiver operating characteristic curve (ROC) value based on the confusion matrix that are consists of true positive, false positive, true negative and false negative rates to estimate the performance of the classifiers. The accuracy of the prediction result is given by Equation (4) as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where TP is "true positive", refers to the number of correct positive prediction, TN is "true negative", refers to the number of correct negative prediction, FP is "false positive", refers to the number of wrong positive prediction and FN is "false negative", refers to the wrong negative prediction. Confusion matrix summarised the actual and prediction result of a classifier. Table 2 shows the confusion matrix for ANN and NB models.

Table 2. Prediction result of ANN and NB

		Actual		Accuracy (%)	ROC
ANN		+	-	95	0.9545
Predictive	+	15	0		
	-	2	20		
NB				84	0.8350
Predictive	+	23	2		
	-	5	15		

From Table 2, the confusion matrix shows that for ANN, compared with the actual slope cases, there are 15 stable slope cases are correctly predicted and no slope case is wrongly predicted while 20 failed slope cases are correctly predicted and 2 slope cases are wrongly predicted. For NB, 23 stable slope cases are correctly predicted and 2 slope cases are wrongly predicted while 15 failed slope cases are correctly predicted and 5 slope cases are wrongly predicted. ANN model gives higher accuracy and ROC value with 95% and 0.9545, compared to NB with 84% and 0.8350 respectively.

4. CONCLUSION

In this study, 148 slope cases were collected from previous studies are used to develop prediction model based on ANN and NB classifier with six input factors; "cohesion, slope height, slope angle, friction angle, unit weight and pore pressure ratio". A three layer feed forward back propagation network with 6-3-1 architecture was selected for ANN model. The dataset is divided into 70% for training and 30% of testing dataset. The prediction result of ANN gives 95% accuracy. For NB, the algorithm was employed based on the probabilities of the slope stability with the same six input parameters and percentage division for training and testing datasets. The prediction result of NB gives 84% accuracy. The conclusions are as follows:

i. ANN and NB show the capabilities to predict the slope stability using the classification approach.
ii. The comparison between ANN and NB was found that ANN performed better than NB for slope stability prediction.

ACKNOWLEDGMENTS

This project was funded by UNITEN under Fundamental Research Grant Scheme FRGS / 1 / 2018 / STG08 / UNITEN / 02 / 1.

REFERENCES

[1] S. Rawat and A. K. Gupta, "Analysis of a nailed soil slope using limit equilibrium and finite element methods," *International Journal of Geosynthetics and Ground Engineering*, vol. 2, no. 4, Oct. 2016. DOI: 10.1007/s40891-016-0076-0.

[2] J. A. Abdalla, M. F. Attom, and R. Hawileh, "Prediction of minimum factor of safety against slope failure in clayey soils using artificial neural network," *Environmental Earth Sciences*, vol. 73, no. 9, pp. 5463–5477, Oct. 2014. DOI: 10.1007/s12665-014-3800-x.

[3] A. K. Verma, T. N. Singh, N. K. Chauhan, and K. Sarkar, "A hybrid FEM–ANN approach for slope instability prediction," *Journal of The Institution of Engineers (India): Series A*, vol. 97, no. 3, pp. 171–180, Aug. 2016. DOI: 10.1007/s40030-016-0168-9.

[4] D.-p. Deng, L. Li, and L.-h. Zhao, "Limit equilibrium method (LEM) of slope stability and calculation of comprehensive factor of safety with double strength-reduction technique," *Journal of Mountain Science*, vol. 14, no. 11, pp. 2311–2324, Sep. 2017. DOI: 10.1007/s11629-017-4537-2.

[5] S. Wu, L. Xiong, and S. Zhang, "Strength reduction method for slope stability analysis based on a dual factoring strategy," *International Journal of Geomechanics*, vol. 18, no. 10, p. 04 018 123, Oct. 2018. DOI: 10.1061/(asce)gm.1943-5622.0001249.

[6] C.-m. Hu, Y.-l. Yuan, Y. Mei, X.-y. Wang, and Z. Liu, "Modification of the gravity increase method in slope stability analysis," *Bulletin of Engineering Geology and the Environment*, vol. 78, no. 6, pp. 4241–4252, Oct. 2018. DOI: 10.1007/s10064-018-1395-2.

[7] J. Singh, H. Banka, and A. K. Verma, "Analysis of slope stability and detection of critical failure surface using gravitational search algorithm," in *2018 4th International Conference on Recent Advances in Information Technology (RAIT)*, IEEE, Mar. 2018. DOI: 10.1109/rait.2018.8389049.

[8] H. Sarfaraz and M. Amini, "Numerical simulation of slide-toe-toppling failure using distinct element method and finite element method," *Geotechnical and Geological Engineering*, vol. 38, no. 2, pp. 2199–2212, Dec. 2019. DOI: 10.1007/s10706-019-01157-9.

[9] W. Fellenius, "Calculation of stability of earth dam," in *Transactions. 2nd Congress Large Dams, Washington, DC, 1936*, vol. 4, 1936, pp. 445–462.

[10] A. W. Bishop, "The stability of earth dams," Ph.D. dissertation, Unversity of London, 1952.

[11] N. R. Morgenstern and V. E. Price, "The analysis of the stability of general slip surfaces," *Géotechnique*, vol. 15, no. 1, pp. 79–93, Mar. 1965. DOI: 10.1680/geot.1965.15.1.79.

[12] E. Spencer, "A method of analysis of the stability of embankments assuming parallel inter-slice forces," *Géotechnique*, vol. 18, no. 3, pp. 384–386, Sep. 1968. DOI: 10.1680/geot.1968.18.3.384.

[13] N. Janbu, "Slope stability computations," in *Embankment-dam Engineering*, R. C. Hirschfeld and S. J. Poulos, Eds. New York: John Wiley and Sons Inc., 1973. DOI: 10.1016/0148-9062(75)90139-4.

[14] L. He, Q. Tian, Z. Zhao, X. Zhao, Q. Zhang, and J. Zhao, "Rock slope stability and stabilization analysis using the coupled DDA and FEM method: NDDA approach," *International Journal of Geomechanics*, vol. 18, no. 6, p. 04 018 044, Jun. 2018. DOI: 10.1061/(asce)gm.1943-5622.0001098.

[15] A. P. Dyson and A. Tolooiyan, "Prediction and classification for finite element slope stability analysis by random field comparison," *Computers and Geotechnics*, vol. 109, pp. 117–129, May 2019. DOI: 10.1016/j.compgeo.2019.01.026.

[16] M. M. Sazzad, F. I. Rahman, and M. A. A. Mamun, "Effects of water-level variation on the stability of slope by LEM and FEM," in *Proceedings of the 3rd International Conference on Civil Engineering for Sustainable Development*, Feb. 12, 2016, pp. 953–959.

[17] D.-Q. Li, T. Xiao, Z.-J. Cao, K.-K. Phoon, and C.-B. Zhou, "Efficient and consistent reliability analysis of soil slope stability using both limit equilibrium analysis and finite element analysis," *Applied Mathematical Modelling*, vol. 40, no. 9-10, pp. 5216–5229, May 2016. DOI: 10.1016/j.apm.2015.11.044.

[18] S. Suman, S. Z. Khan, S. K. Das, and S. K. Chand, "Slope stability analysis using artificial intelligence techniques," *Natural Hazards*, vol. 84, no. 2, pp. 727–748, Jul. 2016. DOI: 10.1007/s11069-016-2454-2.

[19] S. Y. Liu, L. T. Shao, and H. J. Li, "Slope stability analysis using the limit equilibrium method and two finite element methods," *Computers and Geotechnics*, vol. 63, pp. 291–298, Jan. 2015. DOI: 10.1016/j.compgeo.2014.10.008.

[20] X. Feng, S. Li, C. Yuan, P. Zeng, and Y. Sun, "Prediction of slope stability using naive bayes classifier," *KSCE Journal of Civil Engineering*, vol. 22, no. 3, pp. 941–950, Mar. 2018. DOI: 10.1007/s12205-018-1337-3.

[21] Y. Alimohammadlou, A. Najafi, and C. Gokceoglu, "Estimation of rainfall-induced landslides using ANN and fuzzy clustering methods: A case study in saeen slope, azerbaijan province, iran," *CATENA*, vol. 120, pp. 149–162, Sep. 2014. DOI: 10.1016/j.catena.2014.04.009.

[22] Bui, Moayedi, Gör, Jaafari, and Foong, "Predicting slope stability failure through machine learning paradigms," *ISPRS International Journal of Geo-Information*, vol. 8, no. 9, p. 395, Sep. 2019. DOI: 10.3390/ijgi8090395.

[23] A. Ray, V. Kumar, A. Kumar, R. Rai, M. Khandelwal, and T. N. Singh, "Stability prediction of himalayan residual soil slope using artificial neural network," *Natural Hazards*, vol. 103, no. 3, pp. 3523–3540, Jul. 2020. DOI: 10.1007/s11069-020-04141-2.

[24] R. C. Mamat, A. M. Samad, A. Kasa, S. F. M. Razali, A. Ramli, and M. B. H. C. Omar, "Slope stability prediction of road embankment on soft ground treated with prefabricated vertical drains using artificial neural network," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 2, p. 236, Jun. 2020. DOI: 10.11591/ijai.v9.i2.pp236-243.

[25] Z. G. Qian, A. J. Li, W. C. Chen, A. V. Lyamin, and J. C. Jiang, "An artificial neural network approach to inhomogeneous soil slope stability predictions based on limit analysis methods," *Soils and Foundations*, vol. 59, no. 2, pp. 556–569, Apr. 2019. DOI: 10.1016/j.sandf.2018.10.008.

[26] A. Chakraborty and D. Goswami, "Prediction of slope stability using multiple linear regression (MLR) and artificial neural network (ANN)," *Arabian Journal of Geosciences*, vol. 10, no. 17, Sep. 2017. DOI: 10.1007/s12517-017-3167-x.

[27] P. Tsangaratos and I. Ilia, "Comparison of a logistic regression and naïve bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size," *CATENA*, vol. 145, pp. 164–179, Oct. 2016. DOI: 10.1016/j.catena.2016.06.004.

[28] M. G. Sakellariou and M. D. Ferentinou, "A study of slope stability prediction using neural networks," *Geotechnical and Geological Engineering*, vol. 23, no. 4, pp. 419–445, Aug. 2005. DOI: 10.1007/s10706-004-8680-5.

[29] A. M. Zain, H. Haron, and S. Sharif, "Prediction of surface roughness in the end milling machining using artificial neural network," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1755–1768, Mar. 2010. DOI: 10.1016/j.eswa.2009.07.033.

[30] H. Moayedi, A. Osouli, H. Nguyen, and A. S. A. Rashid, "A novel harris hawks' optimization and k-fold cross-validation predicting slope stability," *Engineering with Computers*, Jul. 2019. DOI: 10.1007/s00366-019-00828-8.

[31] N. K. Sah, P. R. Sheorey, and L. N. Upadhyaya, "Maximum likelihood estimation of slope stability," *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, vol. 31, no. 4, p. 210, Aug. 1994. DOI: 10.1016/0148-9062(94)92314-0.

[32] K.-p. Zhou and Z.-q. Chen, "Stability prediction of tailing dam slope based on neural network pattern recognition," in *2009 Second International Conference on Environmental and Computer Science*, IEEE, 2009. DOI: 10.1109/icecs.2009.55.

[33] J. Li and F. Wang, "Study on the forecasting models of slope stability under data mining," in *Earth and Space 2010*, American Society of Civil Engineers, Mar. 2010. DOI: 10.1061/41096(366)77.

[34] Y. Xiaoming and L. Xibing, "Bayes discriminant analysis method for predicting the stability of open pit slope," in *2011 International Conference on Electric Technology and Civil Engineering (ICETCE)*, IEEE, Apr. 2011. DOI: 10.1109/icetce.2011.5776304.