PREDICTION OF ROCK SLOPE FAILURE BASED ON MULTIPLE MACHINE LEARNING ALGORITHMS

by

Mohammed MNZOOL *

Department of Civil Engineering, College of Engineering,
Taif University, Taif, Saudi Arabia

Original scientific paper
https://doi.org/10.2298/TSCI2406907M

Slope failures have the potential to seriously jeopardize access to sustainable development since they cause numerous casualties as well as disastrous effects on society and the economy. It is imperative to use precise operable computational designs in this case. This study examined the efficacy of five distinct machine learning models, namely support vector machines, decision trees, gradient boost machine learning, and random forest, in predicting the slope safety factors. This article's primary goal is to assess and improve the different machine learning-based analytical representations in relation factor of safety computations. The genetic algorithm mimics the processes of growth, hybridization, and mutagenesis found in the expected collection and inherent procedures to resolve the hyperparameters of machine learning algorithms. A total of 217 cases were collected in order to train and evaluate these models. Multiple convergence analysis is also used to study the independence of individual characteristics. The assessed methods' competence was assessed through the application of diverse performance assessment indicators. The various classifiers function satisfactorily for slope failure inquiry, according to the evaluation and comparison of the data. Random forest was found to be the best classification method for slope failure prediction, with an accuracy of 91%. .

Key words: *slope failure, rock slope, support vector machine, decision tree, random forest, genetic algorithms*

Introduction

Because it affects the safety of industrial engineering tasks, slope stability testing is among the most significant areas of geotechnical production [1-4]. Compared to other geotechnical engineering tasks, analysing slope stability is a harder and more difficult task that requires more complexity and effort. In many nations, slope instability is a complicated and expected process that results in significant emergencies and financial damages [5-7]. Hence, lowering the environmental danger of slope tragedies and guaranteeing the security of both individuals and assets depend heavily on developing a safe, dependable, and efficient classical for slope stability investigation, assessment, and estimation.

Landslips are common in mountainous areas worldwide. Landslides can result from human activity, dormant landslides, deforestation, and settlement migration [8, 9]. Anthropogenic activities include environmental damage, pollution, and illogical highways on slopes. Without considering the consequences, hill drilling may cause slope instabilities and failure.

* Author's e-mail: mdway@tu.edu.sa

Slope failure destroys towns and habitats. Slopes can be synthetic or natural. Thus, understanding slope stability changes is crucial to managing landslides.

Deteriorating rocks reduce slope shear resistance. Vegetation also causes slope failures due to mechanical factors like plant root systems and tree loads [10]. Investigation and prediction of slope collapse scenarios are difficult due to slope structure and cause complexity [11]. Landslides are classified by their causes and contributing factors. Recent research on slope failure analysis has used many influencing factors [12]. However, the hidden nature of the main contributing components makes exact forecasts of slope failures difficult to achieve. Despite these challenges, numerous scholars used various techniques to perform slope stability analyses across various nations. The methodologies encompass both qualitative and quantitative analysis as well as field research and analysis [13, 14].

Slope stability has been assessed and predicted using a variety of techniques. Existing literature research has already documented numerous evaluation techniques, including the distinctive line procedure, the boundary assessment method, the limit equilibrium method (LEM), and mathematical modelling [15]. Because of its straightforward application and analysis methodology, LEM is one of the most popular techniques for evaluating slope stability [16, 17]. Though commonly used in practice, LEM has certain intrinsic limitations. Furthermore, in situations where stratifications are non-homogeneous and anisotropic which typically involve geotechnical uncertainties (LEM) they are unreliable [18-20]. Research and analysis clarified the concept of unreliable slopes but provided no solid answer. The use of machine learning algorithms to examine the geomorphological and geological characteristics of hilly areas is currently popular [21]. Suitable outcomes for slope failure study and prediction were shown by machine learning techniques. The [22] showcased a range of supervised machine learning techniques. In [23] for the analysis and prediction of slope stability by incorporating six causal elements. To forecast slope collapses based on the safety factor while taking six causation elements into consideration, researchers used machine learning techniques. Dissimilar machine learning methods were used to build prediction models, which were then assessed and contrasted to provide predictions of slope failure. Bui *et al.* [24], used multivariate modeling and the neural network approach to examine slope stability [25]. The neural network arrangement's input specifications and predicted safety issue were geotechnical and geometrical [26]. Qian *et al.* [27] produced integrated models. Due to advances in science and technology, slope stability studies now use machine learning. The slope data available to assess and predict slope stability is used to examine the relationship between slope stability and its effect. Qi and Tang [28] used metaheuristic machine learning on multiple datasets to assess slope stability, and used an artificial immune algorithm based on the biological immune system's antigen recognition processes to predict slope stability perfectly. Scholars Hoang and Pham [29] proposed a meta heuristic-artificial neural network hybrid model to assess slope stability, they are examined the slope stability of open pit mines using the backpropagation neural network, naive Bayes, decision tree (DT), and support vector machine (SVM) models. Researchers [30] used an ANN to forecast the minimum safety factor for various soil slopes. In [31] determined the slope's safety factor by using a genetic algorithm to find the slope's essential slip surface and a spline curve.

Machine learning and genetic algorithm

In this work, slope stability was simulated and predicted using a machine learning method, and its hyperparameters were optimized using an evolutionary genetic algorithm. This section briefly presents the ideas of the various machine learning and genetic algorithms used in this article.

Machine learning

Machine learning approaches were chosen for this study. These four machine learning algorithms, SVM, random forest (RF), DT, and gradient boosting machine (GBM) were selected because they are superior at handling binary or multi-classification issues influenced by a variety of factors [32, 33]. All of them must choose sensible hyperparameters in order to improve the algorithm's predictive performance and industrial field applicability.

Support vector machine

The area of the SVM model was to identify the best decision boundaries across classes. Using this model has the benefit of its exceptional capacity to handle high dimensional and non-separable data. It is not possible to determine mathematically which outside forces may cause slope failures. The SVM looks for the line or border that most effectively divides both slope failure classes. The SVM divides the classes along several lines, but selects the line with the greatest margin as the final border. The support vectors are the different points that are situated along this line:

$$D = \left\{ \left(u_i + v_i \right) \right\}_{i=1}^{N} \tag{1}$$

where $v_i \in \{+1, -1\}$ is the number of results (u and v are taken into consideration in this study) and u_i stands for the slope failure analysis variables to be studied. The entire number of instances is denoted by N. When classes are linearly separable, the ideal decision boundary that divides the data:

$$f\left(u\right) = wu + b = 0 \tag{2}$$

where b is the indicates the bias value and w indicates weight. The SVM practices its kernel function handle non-linear classification, which may be stated:

$$f\left(u\right) = \operatorname{sign}\left(\sum_{j=1}^{n} \alpha_i v_i k(u_i, v_i) \right) + b \tag{3}$$

where $k\ (u_i, v_i)$ is a kernel function.

The SVM method is a binary classification method that divides the data along a hyperplane class. The fundamental goal of segmentation is to simultaneously convert the interval into a convex quadratic programming problem and maximize its size.

Random forest

The random forest RF model is essentially an improvement over the DT algorithm [33]. It works by generating numerous decision trees. The bootstrap resampling approach constantly extracts the initial training sample set N into m random samples. A fresh training sample set is created using the self-help sample set, and m arrangement trees are built to create an RF. The new data classification uses the score from the number of polls in the classification tree as its basis. The RF model has been selected due to its proven reliability and excellent accuracy across a wide range of data sets [34, 35]. It is also a popular model for classification challenges.

Decision tree

The DT model, a basic machine learning arrangement technique, learns basic decision rules using data elements to predict the value of an objective variable. The algorithm's implementation entails feature selection, DT generation, and DT trimming. The algorithm generates the DT recursively after first classifying it based on the attributes. In the meantime, it trims the unnecessary and redundant nodes from the created DT. The DT generation correlates with the region-specific optimal solution of the model, while the DT pruning correlates with the global optimal solution. The DT is used in the actual slope prediction and is readily recognized [35] employed the DT technique to examine and categorize the variables influencing the slope. The results demonstrate the DT's strong categorization effect.

Genetic algorithm

Hyperparameter settings greatly impact all three machine learning algorithms. Optimisation procedures often use algorithms and experience-based parameter settings. This produces a local optimal solution, which is bad. This work developed a smart hyperparameter optimisation method to address the issue. Some intelligent optimisation algorithms are simulated [35, 36], Metaheuristics, or intelligent optimisation algorithms, are random search algorithms based on artificial intelligence or physical processes. This type of novel algorithm usually does not need restrictions and objectives continuity.

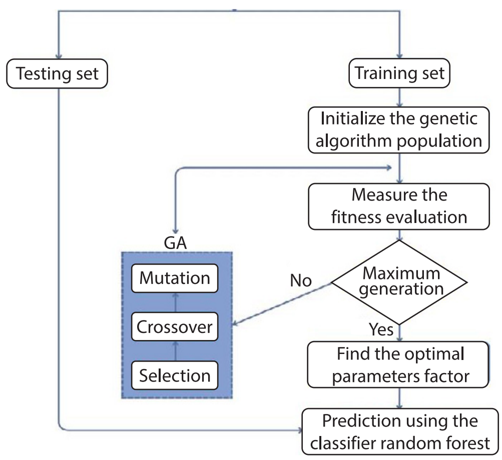

Figure 1. Hyperparameter optimization by the genetic algorithm

Professor Holland first suggested the genetic algorithm in 1969. It is a heuristic search system created based on population genetics and Darwin's theory of evolution by normal selection [36]. The programme mimics the processes of hybridization, mutation, and reproduction in natural collections natural genetics. The genetic process generates a random starting population with a given encoding length. High suitability individuals are chosen for the genetic procedure, while low fitness individuals are excluded. The population creates a new genetic group after meeting the end-point requirement. The genetic algorithm executes the most developed offspring. Figure 1 shows the genetic algorithm's hyperparameter optimisation phases. System integration is easy with genetic algorithm, a combinatorial optimisation method based on biological evolution. Genetic algorithm solves multivariable optimisation problems well.

Different mutation, selection, and cross-recombination processes prevent the optimisation process from missing the global or local optimum. Also suitable for machine learning algorithm hyperparameter optimisation. To predict soil liquefaction potentials, Wang *et al.* [27] optimised RF hyperparameters using a genetic algorithm. The SVM parameters are selected using several methods to identify and classify hazardous railroad cargo, with positive results.

Slope stability analysis dataset

A dataset for slope stability analysis comprises different properties and metrics that are important for evaluating the stability of slopes that are either created by humans or naturally. Geotechnical engineers utilize these datasets to comprehend and forecast how slopes will behave in various scenarios, tab. 1. Assessing a slope's stability or likelihood of failure which could result in landslide or other unstable problems is the main objective. Figure 2 show the correlation matrix that highlights parameters.

Table 1. Sample data of soil properties include cohesion, friction angle, and slope length

Soil type	Moisture content [%]	Cohesion [kPa]	Angle of internal friction [°]	Slope angle [°]	Slope length [m]	Safety factor
Clay	15	350	29	26	60	1.2
Sand	8	150	34	20	70	1.5
Silt	12	200	30	30	60	1.1
Peat	10	200	33	30	50	1.3
Loam	12	200	31	33	70	1.4
Loamy sand	9	175	24	35	60	1.1
Sandy loam	12	250	32	38	50	1.0
Silt loam	15	300	35	40	70	1.3
Sandy clay loam	20	500	38	53	80	1.9
Sandy clay	14	350	31	45	50	1.3
Clay loam	11	225	27	48	70	1.5
Silty clay loam	13	250	26	50	60	1.1
Silty clay	19	450	37	53	80	1.6

In this research, the dataset based on slope stability analysis is used to create machine leaning models. The dataset contains several instances. A list of the instances includes moisture content, cohesion, angle of internal friction, slope angle, slope length and safety factor. Usually, the dataset consists of several examples, each of which reflects a distinct place or situation. The stability of slopes is then evaluated and predicted using the data by doing conventional geotechnical investigations or training machine learning models. It is important to remember that slope stability is a complicated problem that is impacted by a number of natural, man-made, and hydrological variables. Thus, for reliable slope stability evaluations, a complete dataset that takes these different factors into account is essential as shown in fig. 3.

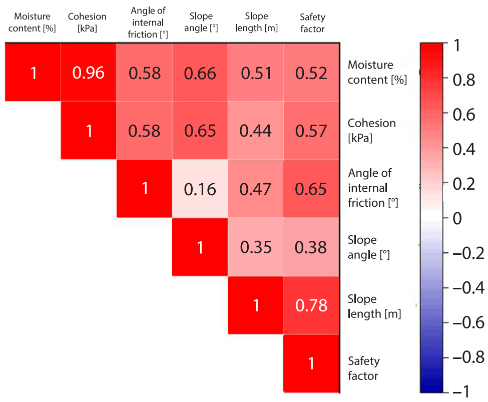

Figure 2. Correlation matrix between six input variables and one output variable in the dataset

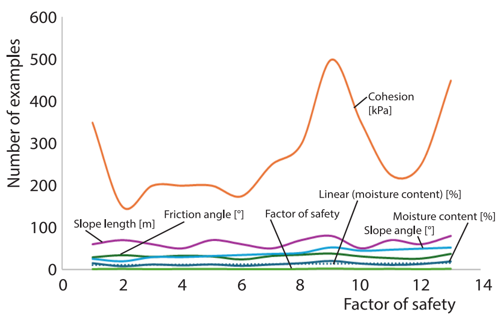

Figure 3. Visualization of slope stability dataset instances

Results and discussion

Result

The review of the approach and its metrics for performance are crucial requirements before implementing any model.

Here is a discussion of the several assessment measurements.

Evaluation metrics

The confusion matrix plays a crucial role in evaluating the performance of the model. The true positive (TP) indicates the real slope failure case that was correctly projected, the true negative (TN) indicates the actual non-slope failure instance that was correctly predicted, the false positive (FP) indicates the actual non-slope failure cases predicted, and the false negative (FN) indicates the actual slope failure cases correctly projected. contrasted with the false positive rate (FPR). The confusion matrix displays both genuine positive rates and false positive rates. The accuracy (ACC), sensitivity (TPR), and specificity (TNR) performance indicators can be defined using the previously specified factors. The following equations show how the aforementioned metrics are formulated:

$$\text{Specificity} = \left(\frac{TN}{N}\right) \tag{4}$$

$$\text{Sensitivity} = \left(\frac{TP}{P}\right) \tag{5}$$

$$\text{Accuracy} = \left(\frac{TN+TP}{N+P}\right) \tag{6}$$

Precision: The definition of accuracy is the fraction of positive slope failure cases that are accurate, or a metric of accuracy. If there are no false positive cases produced by the hypothesis, the model's precision will be equal to 1.0. The precision metric is defined as the fraction of real slope failures accurately expected. If a hypothetical scenario produces no false negative cases, the recall will be 1.0. It is able to be stated as It is able to be stated:

$$\text{Precision} = \left(\frac{TP}{TP+F\ P}\right) \tag{7}$$

Recall: Recall the term as the fraction of real slope failures that were accurately predicted, or a measure of perfection. If there are no false negative cases produced by the model, the recall equal to 1:

$$\text{Recall} = \left(\frac{TP}{TP+F\ N}\right) \tag{8}$$

Predicted accuracy values from sensitivity (TPR) and specificity (FPR) values. A well-performing model will always have a value that is close to or equal to 1, covering the determined area below the value. The optimal location for the TPR against FPR score graph is in the upper left corner, where the values are 100% TPR and 0% FPR. When the AUC value is 1, it indicates that every incidence of slope failure has been correctly categorized.

The study results are in fig. 4. To compare predicted outcomes, several assessments are explained and used. For comparability, the study fully explains and applies several assessments to predicted outcomes. Machine learning improves slope failure analysis and prediction. To predict slope failures, this study compared RF, SVM, gradient boost, and DT machine learning methods. After controlling for causal factors, the four classifiers-RF, DT, gradient boost, and SVM-performed well. Comparative data showed that the RF model was more precise than other classification methods, tab. 2. The RF is 91% accurate. The RF model has higher specificity, sensitivity, precision, and recall (0.84, 0.79, 0.59, and 0.91). There are several categories of

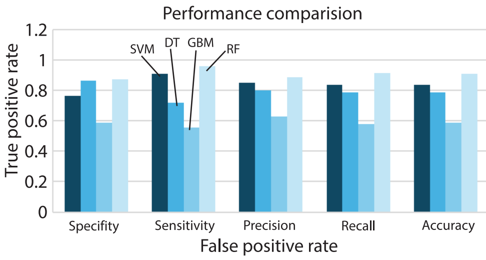

Figure 4. Performance comparison of SVM, DT, GB, and RF classifiers

outcomes. According to the initial rule, unstable slopes, 55° slope angles, and heavy precipitation increase slope collapse risk. Steep highway cuts cause slope failure. Slope weakness and erosion result from low degradation and excessive precipitation everywhere. The slope failure study noted these common features in many classes. Additionally, many slope failure-provoking criteria interact with the expected category. The RF scores higher than others. The RF model predicts slope failures best in the chosen study region based on its specificity, sensitivity, precision, recall, and correctness. By finding dropout patterns in a changing dataset, the study shows the value of machine learning classifiers. In fig. 4, all things considered, the optimal RF model outperformed the other four methods, achieving a 91% accuracy rate and higher TNR values. Due to the poor prediction of TNR by GBM (accuracy = 59%), the SVM algorithm's performance was lower than the RF model, despite its slightly higher AUC and TPR values. Given the realities of slope engineering issues, estimating an unstable slope incorrectly will result in significant property and human damage.

Table 2. Performance of SVM, DT, GB, and RF classifier

Performance measure	Results			
	SVM	DT	GBM	RF
Specificity	0.767	0.867	0.59	0.875
Sensitivity	0.909	0.72	0.558	0.963
Precision	0.85	0.8	0.63	0.89
Recall	0.84	0.79	0.58	0.915
Accuracy	84%	79%	59%	91%

As a result, accurate TNR estimation is more important in slope stability analysis than TPR estimation. This article suggests the best RF model for predicting slope stability.

The K-fold cross validation

Hyperparameter tuning always risks overfitting or selection deviation in the prediction model. Thus, model optimisation usually uses cross-validation. Cross-validation evaluates the model's predictive power on self-determining data sets, finds the hyperparameter rate

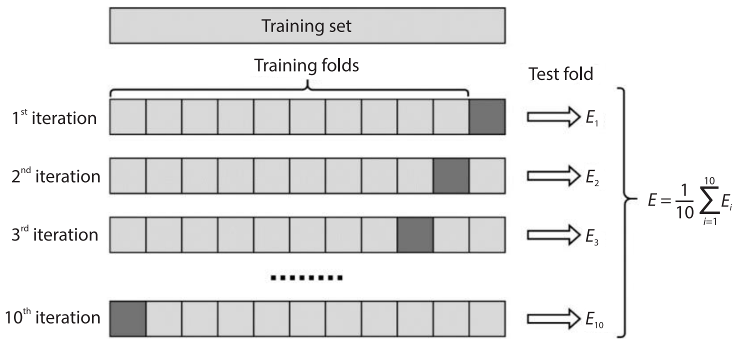

Figure 5. Diagram of the K-fold verification

that optimises simplification presentation, retrains the model across each training set, and uses the independent test set to evaluate its performance, fig. 5. The verification and test sets are complementary, and the K-fold divides the data into k-tiny chunks. This study used 10-fold cross-validation, the most common.

Discussion

When compared to conventional approaches like consulting subject experts through field research and analysis, machine learning algorithms demonstrate higher accuracy in analyzing and forecasting slope failures. To evaluate and compare in order to anticipate slope failures, the following machine learning techniques were used in this work: RF, SVM, DT, and GB. Slope stability study makes extensive use of specific machine learning techniques.

Figure 6 outcomes demonstrate that, when certain causal elements were taken into account, the three classifiers RF, DT, and SVM performed well. The RF model fared better with respect of accuracy when compared to other classifiers, according to an analysis of the comparative data. The RF model has an accuracy rate of 91%. The RF model's performance can be regarded as the best for predicting slope failures in the chosen study area based on the performance comparison on the evaluation metrics are specificity, sensitivity, precision, recall and accuracy.

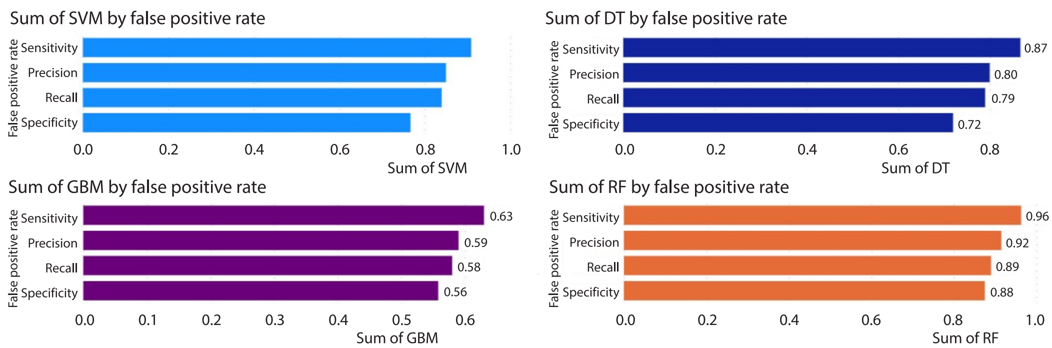

Figure 6. Performance accuracy comparison of SVM, DT, GB, and RF classifiers

Conclusion

In order to forecast unbalanced slopes and hidden interacting patterns associated with slope failures in the north, this study used machine learning approaches. All four of the classification models collectively provide strong prediction abilities for slope failures in the area. It is clear from the comparison and evaluation metrics results that the RF has a high degree of predictive power for slope failures. Therefore, we can utilize the model to assess and develop models that forecast slope failure. When choosing a highly capable model, the receiver operating characteristics curve is useful for comparison shopping. The stability of slopes influences the elements considered in the analysis. When real-time data is available, RF can be applied as an intellectual model to forecast slope failure trends. In additional parts of the state with comparable slope characteristics and a high risk of landslides and slope failures, this model can also be helpful for predicting and analyzing slope failures. Using optimization strategies is one way of enhancing the analyzed classifiers' prediction capability. Furthermore, not each slope constraint is appropriate for each classification model. As a result, choosing a categorization model requires careful consideration of all relevant triggering elements. Ultimately, slope stability analysis can effectively use machine learning algorithms. The disaster management team will also benefit from the model's timely risk identification and reduction.

Acknowledgment

The authors would like to acknowledge the Deanship of Graduate Studies and Scientific Research, Taif University for funding this work.

References

[1] Komadja, G. C., *et al.*, Geotechnical and Geological Investigation of Slope Stability of a Section of Road Cut Debris-Slopes Along NH-7, Uttarakhand, India, *Results in Engineering*, *10* (2021), 100227
[2] Ray, A., *et al*., Stability Prediction Of Himalayan Residual Soil Slope Using Artificial Neural Network, *Natural Hazards*, *103* (2020), 3, pp. 3523-3540
[3] Gao, W., Ge, S., A Comprehensive Review of Slope Stability Analysis Based on Artificial Intelligence Methods, *Expert Systems with Applications*, *239* (2023), 122400
[4] del Potro, R., Hurlimann, M., The Decrease in the Shear Strength of Volcanic Materials with Argillic Hydrothermal Alteration, Insights from the Summit Region of Teide Stratovolcano, Tenerife, *Engineering Geology*, *104* (2009), 1-2, pp. 135-143
[5] Suman, S., *et al*., Slope Stability Analysis Using Artificial Intelligence Techniques, *Natural Hazards*, *84* (2016), July, pp. 727-748
[6] Baghbani, A., *et al*., Application of Artificial Intelligence on Geotechnical Engineering: A State-of-the-Art Review, *Earth-Science Reviews*, *228* (2022), 103991
[7] ***, Application of Elasto-Plastic Analysis, pdf
[8] Van Eynde, E., *et al.*, Impact of Landslides on Soil Characteristics: Implications for Estimating Their Age, *Catena*, *157* (2017), Oct., pp. 173-179
[9] Sun, L., *et al*., Stability Analysis of Reservoir Slopes under Fluctuating Water Levels Using the Combined Finite-Discrete Element Method, *Acta Geotechnica*, *18* (2023), 10, pp. 5403-5426
[10] Feng, S., *et al*., Analytical Analysis of The Mechanical and Hydrological Effects of Vegetation on Shallow slope stability, *Computers and Geotechnics*, *118* (2020), 103335
[11] Raghuvanshi, T. K., Plane Failure in Rock Slopes – A Review on Stability Analysis Techniques, *Journal of King Saud University-Science*, *31* (2019) 1, pp. 101-109
[12] Zhang, W., *et al.*, Landslide Susceptibility Research Combining Qualitative Analysis and Quantitative Evaluation: A Case Study of Yunyang County in Chongqing, China, *Forests*, *13* (2022), 7, 1055
[13] Wang, X., H., *et al*., A Method for Slope Stability Analysis Considering Subsurface Stratigraphic Uncertainty, *Landslides*, *15* (2018), Dec., pp. 925-936
[14] Napoli, M. L., *et al*., A Stochastic Approach to Slope Stability Analysis in Bimrocks, *International Journal of Rock Mechanics and Mining Sciences*, *101* (2018), Jan., pp. 41-49
[15] Deng, D.-P., *et al*., Limit Equilibrium Method (LEM) of Slope Stability and Calculation of Comprehensive Factor of Safety with Double Strength-Reduction Technique, *Journal of Mountain Science*, *14* (2017), 11, pp. 2311-2324
[16] Lim, K., *et al*., Slope-Stability Assessments Using Finite-Element Limit-Analysis Methods, *International Journal of Geomechanics*, *17* (2017), 2, 06016017
[17] Cala, M., Flisiak, J., *Slope Stability Analysis with FLAC and Limit Equilibrium Methods, in FLAC and Numerical Modelling in Geomechanics 2001*, CRC Press, Boca Raton, Fla., USA, 2020, pp. 111-114
[18] Demir, S., Sahin, E. K. Random Forest Importance-Based Feature Ranking and Subset Selection for Slope Stability Assessment Using the Ranger Implementation, *Avrupa Bilim ve Teknoloji Dergisi*, *48* (2023), Feb., pp. 23-28
[19] Liu, Y., *et al.*, Rock-Soil Slope Stability Analysis by Two-Phase Random Media and Finite Elements, *Geoscience Frontiers*, *9* (2018), 6, pp. 1649-1655
[20] Mohammed, M., Wan, L., Slope Stability Analysis of Southern Slope of Chengmenshan Copper Mine, China, *International Journal of Mining Science and Technology*, *25* (2015), 2, pp. 171-175
[21] Barnes, T. J., *et al*., A Machine Learning Approach to the Geomorphometric Detection of Ribbed Moraines in Norway, *EGUsphere*, *12* (2024), 3, pp. 801-818
[22] Yang, H.-F., *et al*., Supervised Learning of Semantics-Preserving Hash Via Deep Convolutional Neural Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, *40* (2017), 2, pp. 437-451

[23] Tang, Y., *et al*., Integrating Principal Component Analysis with Statistically-Based Models for Analysis of Causal Factors and Landslide Susceptibility Mapping: A Comparative Study from The Loess Plateau Area in Shanxi (China), *Journal of Cleaner Production*, *277* (2020), 124159
[24] Achour, Y., Pourghasemi, H. R., How do Machine Learning Techniques Help in Increasing Accuracy of Landslide Susceptibility Maps, *Geoscience Frontiers*, *11* (2020), 3, pp. 871-883
[25] Goetz, J., *et al.*, Evaluating Machine Learning and Statistical Prediction Techniques for Landslide Susceptibility Modelling, *Computers and Geosciences*, *81* (2015), pp. 1-11
[26] Lin, Y., K. *et al.*, Prediction of Slope Stability Using Four Supervised Learning Methods, *Ieee Access*, *26* (2018), pp. 31169-31179
[27] Wang, C., *et al*., Machine Learning-Based Regional Scale Intelligent Modelling of Building Information for Natural Hazard Risk Management, *Automation in Construction*, *122* (2021), 10
[28] Qi, C., Tang, X., Slope Stability Prediction Using Integrated Metaheuristic and Machine Learning Approaches: A Comparative Study, *Computers and Industrial Engineering*, *118* (2018), Apr., pp. 112-122
[29] Hoang, N.-D., Pham, A.-D., Hybrid Artificial Intelligence Approach Based on Metaheuristic and Machine Learning for Slope Stability Assessment: A Multinational Data Analysis, *Expert Systems with Applications*, *46* (2016), Mar., pp. 60-68
[30] Qian, Z., *et al*., An Artificial Neural Network Approach to Inhomogeneous Soil Slope Stability Predictions Based on Limit Analysis Methods, *Soils and Foundations*, *59* (2019), 2. pp. 556-569
[31] Feng, X., *et al*., A Research on the Methods for Prediction of the Slope Stability of Open-pit Mine, *Proceedings*, 9th China-Russia Symposium, Coal in the 21st Century: Mining, Intelligent Equipment and Environment Protection, (COAL 2018), Atlantis Press, Amsterdam, The Notherlands, 2018
[32] Kibria, H. B., Matin, A., The Severity Prediction of The Binary and Multi-Class Cardiovascular Disease-A Machine Learning-Based Fusion Approach, *Computational Biology and Chemistry*, *98* (2022), 107672
[33] Chakraborty, A., Goswami, D., Prediction of Slope Stability Using M Linear Regression (MLR) and Artificial Neural Network (ANN), *Arabian Journal of Geosciences*, *10* (2017), 17, pp. 1-11
[34] Teng, X., Gong, Y., Research on Application of Machine Learning in Data Mining, IOP Conference Se*ries: Materials Science and Engineering*, *392* (2018), 6, 062202
[35] Hwang, S., *et al.*, Slope Failure Prediction Using a Decision Tree: A Case of Engineered Slopes in South Korea, *Engineering Geology*, *104* (2009), 1-2, pp. 126-134
[36] Huljanah, M., *et al.*, Feature Selection Using Random Forest Classifier for Predicting Prostate Cancer, *IOP Conference Series: Materials Science and Engineering*, *546* (2019), 5, 052031

Paper submitted: June 25, 2024
Paper revised: August 20, 2024
Paper accepted: October 17, 2024