Vietnam Journal of Earth Sciences, 47(2), 182–200, https://doi.org/10.15625/2615-9783/22196

Vietnam Journal of Earth Sciences
https://vjs.ac.vn/index.php/jse

Prediction of safety factor for slope stability using machine learning models

Dam Duc Nguyen[1], Manh Duc Nguyen[2], Indra Prakash[3], Nguyen Van Huong[4], Hiep Van Le[1], Binh Thai Pham[1,*]

[1]*University of Transport Technology, 54 Trieu Khuc, Thanh Xuan, Hanoi, Vietnam*
[2]*University of Transport and Communications, Hanoi, Vietnam*
[3]*DDG(R) Geological Survey of India, Gandhinagar, Gujarat, 382010, India*
[4]*Hanoi University of Natural Resources and Environment, Hanoi, Vietnam*

Received 04 October 2024; Received in revised form 04 November 2024; Accepted 08 January 2025

ABSTRACT

Slope instability is a common geological hazard along mountainous roads in Vietnam, leading to significant damage to infrastructure, traffic disruptions, and loss of life. Predicting slope stability, typically quantified by the Factor of Safety (FS), is challenging due to the complex interactions between geotechnical, topographical, and environmental factors. This study aims to develop efficient and accurate models for predicting the FS of natural slopes using advanced machine learning techniques, including Gradient Boosting (GB), Support Vector Machine (SVM), Multi-layer Perceptron (MLP) Neural Networks, Random Forest (RF), and AdaBoost (AB). 371 slope stability cases were used to create a comprehensive database for model training. Both geotechnical and topographical parameters were considered in the FS prediction process. The performance and reliability of these models were evaluated using standard metrics such as R^2, MAE, and MSE. The results demonstrated that all models exhibited satisfactory prediction capabilities, with the optimized GB model achieving the highest accuracy (R^2 = 0.975, MAE = 0.079, and RMSE = 0.120). Additionally, SHAP analysis was employed to assess the importance of input variables in predicting the FS. The findings revealed that slope ratio (X1), slope height (X2), and the number of steeps (X3) were the most influential parameters in the FS prediction.

Keywords: Slope instability, soft computing, factor of safety, gradient boosting, Vietnam.

1. Introduction

Predicting slope stability safety factors is a critical concern in geotechnical engineering. It has significant implications for infrastructure safety, environmental sustainability, and disaster management (Cheng, 2003; Cho, 2007). Unstable slopes can lead to landslides, threatening human lives, property, and ecosystems (Zhang Fanyu et al., 2021). Accurately assessing slope stability's safety factor (FS) is essential for preventing such disasters (Collison et al., 1996). This makes it a key area of research. The complexity of the slope stability factors includes soil properties, water content, slope geometry, and external forces such as rainfall. These complexities complicate the predictive task (Kolapo et al., 2022). Since slope failures often occur

*Corresponding author, Email: binhpt@utt.edu.vn

unpredictably, enhancing predictive models is vital for better risk assessment. This is essential for informed decision-making in construction and land-use planning (Fell et al., 2008).

Analyzing slope stability involves soil, rock, and other materials with complex spatial variability and nonlinearity, making it challenging to study and predict their behavior. Traditional methods for determining slope stability primarily rely on laboratory and field tests, requiring considerable technical expertise to assess stability based on limited experiments (Moayedi et al., 2019). Additionally, conventional analytical techniques often simplify conditions, such as assuming homogeneous soil, which overlooks the inherent spatial variability in actual soils. Today, numerous computational tools enable more efficient slope stability analysis, including geotechnical software based on limit equilibrium methods (LEM), including methods such as the Fellenius method (Swedish Circle), modified bishop method, jambu method, Spencer method, and Morgenstern-price method and finite element methods (FEM) include methods such as linear finite element analysis (Linear FEM), nonlinear finite element analysis (Nonlinear FEM), limit state analysis, stress-strain method, and plasticity analysis (Hammouri et al., 2008). These methods use static analyses to estimate the safety factor of slopes. They calculate the balance between driving forces, like gravity and water pressure, and resisting forces, such as soil strength. LEM and FEM provide valuable insights into how slopes behave. However, they depend heavily on assumptions about soil properties and boundary conditions. This can lead to inaccuracies in complex, real-world situations. Additionally, these methods are often expensive and time-consuming, which limits their use in cases that need quick assessments or large-scale analyses (Liu et al., 2020; Liu et al., 2019). Therefore, the general challenge in further developing slope stability analysis methods is creating a reliable general design tool to assess slope stability accurately.

Machine learning (ML) models have recently become famous for predicting slope stability. They are effective because they can handle complex, nonlinear relationships between variables and learn directly from data (Lin et al., 2024). Many studies have used soft computing (SC) techniques, which are computational methods designed to model complex soil, rock, and other natural conditions affecting slope stability. These include artificial neural networks (ANN) (Kumar et al., 2024; Meng et al., 2021) and support vector machines (SVM) (Wang et al., 2023) to address geotechnical problems like slope failure prediction. Ensemble learning methods, such as Random Forest (RF) (Xie et al., 2022), Gradient Boosting (GB) (Wang et al., 2020), and AdaBoost (AB) (Lin et al., 2022), often perform better than traditional single models. They improve accuracy and capture interactions between variables. However, most previous studies have focused on individual models or small datasets. The comparative performance of different ensemble and boosting models is still poorly understood. This gap shows the need for thorough evaluations of these advanced techniques in predicting slope stability.

This study develops and evaluates several machine learning models for predicting slope stability safety factors, including Gradient Boosting (GB), Random Forest (RF), AdaBoost (AB), Multi-Layer Perceptron (MLP), and Support Vector Machine (SVM). The novelty of this study lies in its application of multiple state-of-the-art machine learning models to predict slope stability, offering a robust, data-driven approach to assessing the Factor of Safety in real-world conditions.

Unlike traditional methods that rely on empirical formulas and often fail to account for the complexity of natural slopes, this study leverages the power of machine learning to integrate and analyze a wide range of geotechnical and topographical parameters. Furthermore, while most previous studies have focused on individual models or small datasets, the comparative performance of different ensemble and boosting models in predicting slope stability remains underexplored. This gap underscores the importance of conducting thorough evaluations of these advanced techniques. The findings of this study provide a more accurate, timely, and reliable tool for predicting slope stability, which can be used for proactive risk management and mitigation in mountainous regions. By identifying the key factors influencing slope stability, this research also offers valuable insights that can guide the design of safer infrastructure and inform policy decisions related to land use and environmental conservation. The models are evaluated using multiple performance metrics, such as root mean square error (RMSE), mean absolute error (MAE), and R-squared (R^2). The findings will aid in selecting the most effective machine-learning approach for slope stability assessment.

2. Materials and Methods

2.1. Data used

The data utilized in this study pertains to the newly planned construction of the Tuyen Quang - Ha Giang expressway, which has a total length of nearly 105 km. The expressway begins at the intersection of the Tuyen Quang - Phu Tho Expressway with QL 2D, located in Nhu Khe Commune, Yen Son District (Tuyen Quang), and terminates at Tan Quang Commune, Bac Quang District (Ha Giang). Of this, 77 km is within Tuyen Quang province, passing through Yen Son and Ham Yen districts, while the remaining 27.5 kilometer stretch traverses Ha Giang Province, specifically Bac Quang District (Fig. 1).

The data of 371 slopes of the Tuyen Quang - Ha Giang expressway were constructed and analyzed to train and validate the predictive models. This data includes two variables: the FS of the slopes (dependent variable - output) and factors that affect the FS of the slopes (independent variables - inputs). Out of these, a set of the affecting factors include dimensional factors of slopes (height, slope ratio, and the Number of steeps) and geotechnical factors of the slopes (thickness of soils, unit weight (γ), cohesion (c), and internal friction angle (φ)). A description of the variables used in the analysis is given below and in Table 1:

2.2.1. Factor of safety (FS)

FS values of the slopes were calculated using Bishop's theory embedded in the GEO-Studio software. In this regard, the slip surface is assumed to be a circular cylindrical arc, dividing the sliding mass into several slices, and the stability factor (FS) is calculated as the ratio between the resisting force and the driving force (Malik et al., 2020). Table 1 shows that the values of the FS range from 0.532 to 3.32.

2.2.2. Slope ratio (X1)

The slope ratio is an important index used to describe the slope of a land surface. It is typically calculated by dividing the height of the slope by its corresponding horizontal length. The slope ratio is often expressed as a ratio or as an angle (Wang et al., 2022). A higher slope ratio usually indicates a steeper terrain with a higher probability of instability, while a lower coefficient suggests a less steep and more stable slope. This study analyzes the slope ratio with 7 different slopes including 0.7, 0.75, 1, 1.25, 1.5, 1.75, 2.

2.2.3. Slope height (X2)

The height of a slope affects not only its ability to accumulate soil but also other factors such as slope angle, water pressure, erosion, and soil consistency (Yuan et al., 2021). Taller slopes tend to accumulate more soil mass, resulting in more significant pressure on the underlying soil. This can increase the probability of instability, especially when dealing with weaker soils like clay. Additionally, as the height of the slope increases, so does its slope angle, leading to more vital shear forces, reducing the soil's bearing capacity, and increasing the risk of landslides.

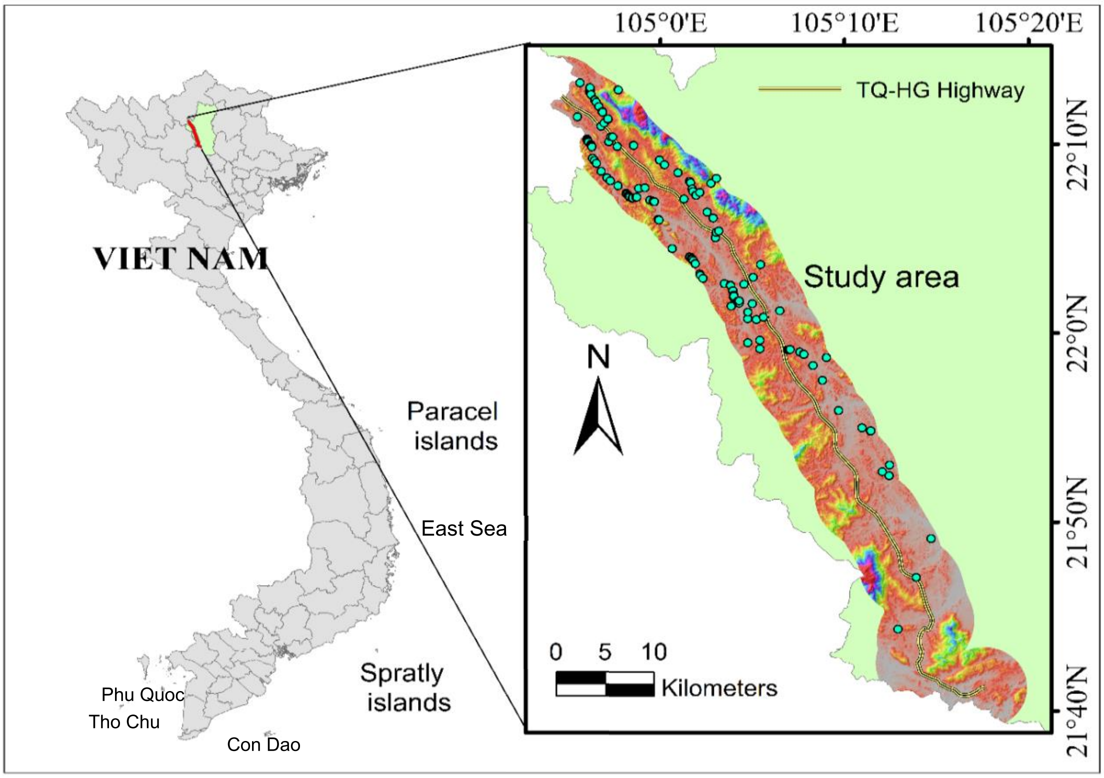

Figure 1. Study area of Tuyen Quang - Ha Giang expressway

2.2.4. Number of steeps (X3)

The number of slope steeps is an essential factor in evaluating terrain stability. It is often used to assess the hazard potential of terrain gradients (Statham et al., 2020). High slopes are typically associated with a greater risk of instability, while slope ratios for more stable configurations are also considered. Generally, each slope has a height ranging from 6 to 12 meters.

2.2.5. Geotechnical factors

Each soil layer exhibits distinct characteristics reflecting its formation environment and structure. These parameters significantly influence the stability and behavior of each soil layer (Khanna et al., 2021). Understanding these characteristics for each soil layer helps geological experts and civil engineers assess and predict the stability of slopes in detail and accurately, thereby devising effective protective measures. The parameters of the soil layers used are as follows:

Unit weight (γ): The unit weight of the soil layer affects its overall density and mass. Soil layers with higher unit weights typically exert greater downward forces, which can impact stability and compaction (Chang et al., 2023).

Cohesion (c): Cohesion represents the internal strength of the soil layer and its resistance to shear forces. Soils with higher cohesion often demonstrate greater stability and shear strength, especially in cohesive soils like clay (Schjønning et al., 2020).

Internal friction angle (φ): The internal friction angle determines the soil's resistance to shear deformation (Rasti et al., 2021). Soils with higher phi *(φ)* values can withstand larger shear stresses before experiencing shear failure. This parameter is particularly crucial in granular soils such as sand and gravel.

Thickness of soils (h): The thickness of each soil layer contributes to the overall structure and geometry of the slope. Thicker layers can provide greater stability and resistance to strong erosion, whereas thinner layers may be prone to collapse and mass movement (Zhang Xuanchang et al., 2021).

In this study, we examined the correlation between the input and output variables (FS). Pearson correlation coefficient was used to measure the linear correlation between pairs of variables, with values ranging from -1 to 1. Values close to 1 or -1 indicate a strong correlation, while values near 0 indicate no correlation (Saccenti et al., 2020). From Fig. 2, it can be seen that variables slope ratio, slope height, and number of slope steeps have the highest correlation with the FS. Other variables also have a reasonable correlation with the FS. Therefore, all the input variables were used to build a predictive model for predicting the FS.

Table 1. Statistical analysis distribution of the inputs in this study

No.	Parameter	Notation	Unit	Mean	STD	Min	Max
1	Unit weight_layer 1	$\gamma 1$	kN/cm^3	12.105	8.712	0	19.99
2	Cohesion_layer 1	$c1$	kPa	15.864	11.408	0	25.9
3	Internal friction angle_layer 1	$\varphi 1$	°C	11.048	7.949	0	17.9
4	Thickness_layer 1	h1	M	4.64	4.605	0	15.81
5	Unit weight_layer 2	$\gamma 2$	kN/cm^3	15.507	7.043	0	19.99
6	Cohesion_layer 2	$c2$	kPa	26.697	15.314	0	77.3
7	Internal friction angle_layer 2	$\varphi 2$	°C	15.25	6.935	0	19.85
8	Thickness_layer 2	h2	m	6.173	4.731	0	18.59
9	Unit weight_layer 3	$\gamma 3$	kN/cm^3	12.147	10.533	0	27.14
10	Cohesion_layer 3	$c3$	kPa	18.851	16.057	0	36
11	Internal friction angle_layer 3	$\varphi 3$	°C	12.357	10.718	0	27.36
12	Thickness_layer 3	h3	m	3.217	3.298	0	11.5
13	Unit weight_layer 4	$\gamma 4$	kN/cm^3	4.272	9.014	0	27.14
14	Cohesion_layer 4	$c4$	kPa	5.926	12.414	0	36
15	Internal friction angle_layer 4	$\varphi 4$	°C	4.149	8.824	0	27.36
16	Thickness_layer 4	h4	m	1.208	3.082	0	16.43
17	Unit weight_layer 5	$\gamma 5$	kN/cm^3	12.207	13.443	0	27.2
18	Cohesion_layer 5	$c5$	kPa	16.302	17.944	0	36
19	Internal friction angle_layer 5	$\varphi 5$	°C	12.389	13.637	0	27.36
20	Thickness_layer 5	h5	m	4.845	8.233	0	32.17
21	Slope ratio	X1	-	1.25	0.501	0.5	2
22	Slope height	X2	m	20.942	10.286	4.43	54.88
23	Number of steeps	X3	-	2.618	1.286	0.55	6.86
24	FS	Y	-	1.553	0.471	0.532	3.32

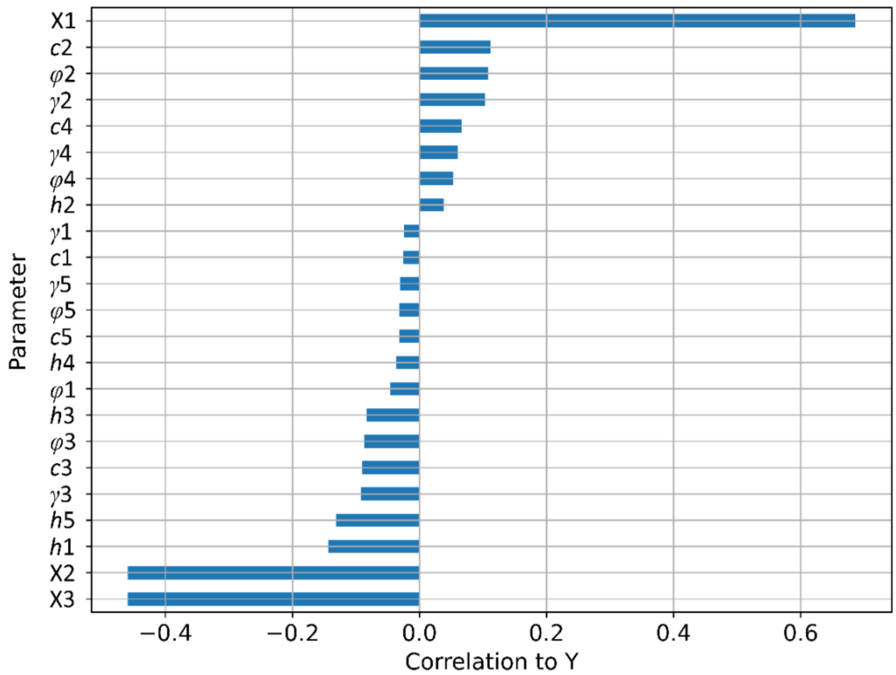

Figure 2. Correlation analysis between input variables and the FS

2.3. Methods used

In this study, five machine learning models are applied to predict the FS of slopes: Gradient Boosting (GB), Support Vector Machine (SVM), MLP Neural Networks (MLP), Random Forest (RF), and AdaBoost (AB). The basic algorithms are described below:

2.3.1. Gradient Boosting (GB)

This study employs the Gradient Boosting Regression (GB) model to minimize expectations and enhance model accuracy. GB is a rational choice due to its high performance and stability. Introduced by Friedman, the GB algorithm extends the boosting method to tackle regression problems (Friedman, 2001). This algorithm utilizes a gradient-based loss function to find the minimum of the objective function. GB has been widely adopted due to its efficient handling of noisy and unclean data. It supports various loss functions and provides robust predictive capabilities for nonlinear data (Friedman, 2001). One significant advantage of GB is its ability to avoid overfitting in decision tree learning by curbing the excessive growth of the trees from the outset.

GB works by building an ensemble of weak decision trees, with each tree trained to correct the errors of the preceding one. This process continues until the maximum number of trees is reached or the error is reduced to an acceptable level. Another strength of GB is its ability to incorporate new decision trees without retraining the entire model, thereby saving time and computational resources (Nguyen et al., 2021). The application of GB has demonstrated high effectiveness due to its ability to analyze and predict complex and diverse datasets accurately. This allows researchers to make reliable conclusions and valuable predictions across various fields, from genetics to disease research and pharmacology.

2.3.2. Support Vector Machine (SVM)

SVM has gained widespread adoption because of its capability to handle both linearly and non-linearly separable data effectively. It utilizes various kernel functions, including linear, polynomial, and radial basis function (RBF) kernels, which allow it to capture complex relationships within the data (Ghosh et al., 2001). A significant strength of SVM is its ability to prevent overfitting, especially in high-dimensional spaces, by maximizing the margin between different classes. The SVM algorithm operates by constructing one or more hyperplanes in a high-dimensional space that can be utilized for tasks such as classification and regression. The primary objective is to identify a hyperplane that separates data points into different classes with the maximum margin. For datasets that are not linearly separable, SVM applies kernel tricks to map the data into a higher-dimensional space where linear separation becomes possible (Pisner et al., 2020).

The effectiveness of SVM has been proven across various applications due to its precise classification and prediction capabilities for complex and varied datasets (Abdullah et al., 2021). This robustness enables researchers to draw reliable conclusions and make valuable predictions in diverse fields, such as image recognition, bioinformatics, and financial forecasting.

2.3.3. MLP Neural Networks (MLP)

MLP is a regression model utilizing a feedforward neural network to solve complex regression problems by learning from data to identify nonlinear relationships. The ability to process input data through multiple layers of interconnected neurons helps improve prediction accuracy and make more confident decisions based on existing information. The development of libraries such as sci-kit-learn has made the application of MLP regressors easier and more efficient, allowing researchers and professionals to exploit the potential of this model in real-world problems (Jin et al., 2022).

Extending the perceptron model, MLP includes three layers: the input layer, hidden layer, and output layer. Each layer comprises artificial neurons, which act as information processing units. In an MLP network, neurons compute a weighted sum of inputs, apply an activation function, and produce an output. The backpropagation algorithm involves feedforward and backward phases (Sekhar et al., 2020). During feedforward, input information traverses the network without altering parameters, while the network error, reflecting the difference between actual and desired output, is computed. In the backward phase, parameters (weights and biases) are adjusted to minimize this error.

2.3.4. Random Forest (RF)

RF is a machine learning model based on the Random Forest algorithm. This algorithm is part of ensemble models, where multiple decision trees are built and combined to create a predictive model. RF is widely used to predict numerical values such as house prices, sales volumes, and personal income.

Each decision tree in the RF is constructed based on a subset of the training data sampled according to the bootstrap sampling principle (Ali et al., 2012). This ensures that each decision tree is trained on a different dataset, reducing the risk of overfitting and increasing the diversity of the model. RF also supports various hyperparameters that can be tuned to optimize the model's performance, including the number of trees, the depth of each tree, and the number of features used in building each tree.

During prediction, RF computes the average of forecasts generated by each decision tree in the ensemble. This helps minimize bias and improve the accuracy of predictions. RF is often a robust choice for regression problems when the data has many features and complex correlations (Ben Ishak, 2016).

2.3.5. AdaBoost (AB)

AB is a regression model in machine learning built on the foundation of the AdaBoost algorithm. It's part of ensemble models, where multiple weak learners are combined to form a strong predictive model. AB is used to predict numerical values and is a popular choice in regression problems (Rabbani et al., 2023). The algorithm works by training multiple weak regressors, often simple decision stumps, on different subsets of the data. Each weak regressor focuses on predicting data points that previous models have misclassified. The predictions from these weak regressors are then combined to produce the final prediction of AB.

AB automatically adjusts the weights of each weak regressor based on their performance (Shrestha et al., 2006). Those with better performance are given higher weights in the final prediction combination. This helps improve prediction accuracy and minimize errors.

By combining multiple weak regressors, AB often produces robust and accurate predictions, especially in regression problems with complex and nonlinear data.

2.4. Validation indicators

Validation indicators, also called evaluation metrics or performance measures, are essential for evaluating the quality and efficacy of machine learning models. These metrics offer insights into how well a model generalizes to unseen data and facilitate the comparison of different models or the tuning of hyperparameters (Hicks et al., 2022). Several common validation indicators used across various machine learning tasks include:

Mean Absolute Error (MAE): MAE quantifies the average absolute variance between predicted and actual values. It provides a straightforward interpretation of the average error magnitude.

$$MAE = \frac{1}{n}\sum_{k=1}^{n}\left[y_k - y_k^{\wedge}\right] \qquad (1)$$

Root Mean Squared Error (RMSE): RMSE is the square root of MSE and offers an interpretable measure of the average error in the same units as the target variable.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(y_k - y_k^{\wedge}\right)^2} \qquad (2)$$

R-squared (R^2): R^2 signifies the proportion of variance in the dependent variable predictable from the independent variables. Ranging from 0 to 1, higher R^2 values denote superior model performance.

$$R^2 = 1 - \frac{\sum_{k=1}^{n}\left(y_k - y_k^{\wedge}\right)^2}{\sum_{k=1}^{n}(y_k - \bar{y})^2} \qquad (3)$$

where n represents the size of the training dataset, y_k is the actual observed values, $y_k^{\wedge}$ the predicted values and $\bar{y}$ is the mean of the observed values.

2.5. SHapley Additive exPlanations (SHAP)

SHAP is a technique for explaining model predictions in machine learning. It relies on cooperative game theory and Shapley values to provide a theoretical framework for determining the importance of features for model predictions (Fryer et al., 2021). Moreover, SHAP is model-agnostic and can be applied to any machine learning model, including complex ones like deep neural networks. It offers a rational and theoretically sound approach to identifying the importance of features, ensuring that the contributions of features are consistent and unbiased (Lundberg et al., 2017). SHAP values provide comprehensive explanations of feature importance, helping users understand the overall behavior of the model.

In summary, SHAP is a powerful technique for explaining model predictions by quantifying the importance of features based on principles of cooperative game theory. SHAP offers visualization techniques to effectively represent these values, such as summary, force, and dependence plots. These visualizations help users understand which features significantly impact the model's prediction for a selected data point (Yin et al., 2024).

3. Results and discussion

3.1. Validation and comparison of the ML models

In training the models, the grid search cross-validation method was used to tune and select the optimal hyperparameters for the models (Table 2). This method involves systematically searching through a predefined set of hyperparameters, evaluating the model's performance using cross-validation, and selecting the set of hyperparameters that yields the best results. In grid search, a grid of hyperparameters is defined, and the model is trained and evaluated for each possible combination of these hyperparameters. The models were validated using various statistical indexes such as MAE, RMSE, and R^2 on both training and testing datasets. While validation of the models using a training dataset shows the goodness of fit of the models with the data used, validation of the models using a testing dataset shows the predictive capability of the models. From Table 3, it can be observed that the GB model achieved the lowest values of MAE on both training (0.04) and testing (0.079) datasets, and the AB received the highest values of MAE on both training (0.141) and testing (0.158) datasets. With RMSE, the GB model also received the lowest values on both training (0.059) and testing (0.100) datasets, the MLP had the highest value (0.183) on the training dataset, and the MLP and AB had the highest value on the testing dataset (0.200). Figures 3 and 4 show the performance of the models using R^2. It can be seen that the GB model has the highest value of R^2 (0.984), while the MLP model received the lowest value of R^2 (0.843) on the training dataset. The GB model also received the highest R2 (0.939) with the testing dataset, while the AB received the lowest R^2 (0.830). Figures 5 and 6 show the distribution of errors produced while training and testing datasets in the models.

Table 2. The models' hyper-parameters used in this study

No.	Hyper-parameter	Model				
		GB	SVM	MLP	RF	AB
1	n_estimators	100	-	-	200	100
2	learning_rate	0.1	-	-	-	0.1
3	max_depth	3	-	-	10	-
4	min_samples_split	-	-	-	5	-
5	min_samples_leaf	-	-	-	2	-
6	c	-	10	-	-	-
7	kernel	-	poly	-	-	-
8	hidden_layer_sizes	-	-	50	-	-
9	activation	-	-	relu	-	-
10	loss	-	-	-	-	linear
11	solver	-	-	adam	-	-

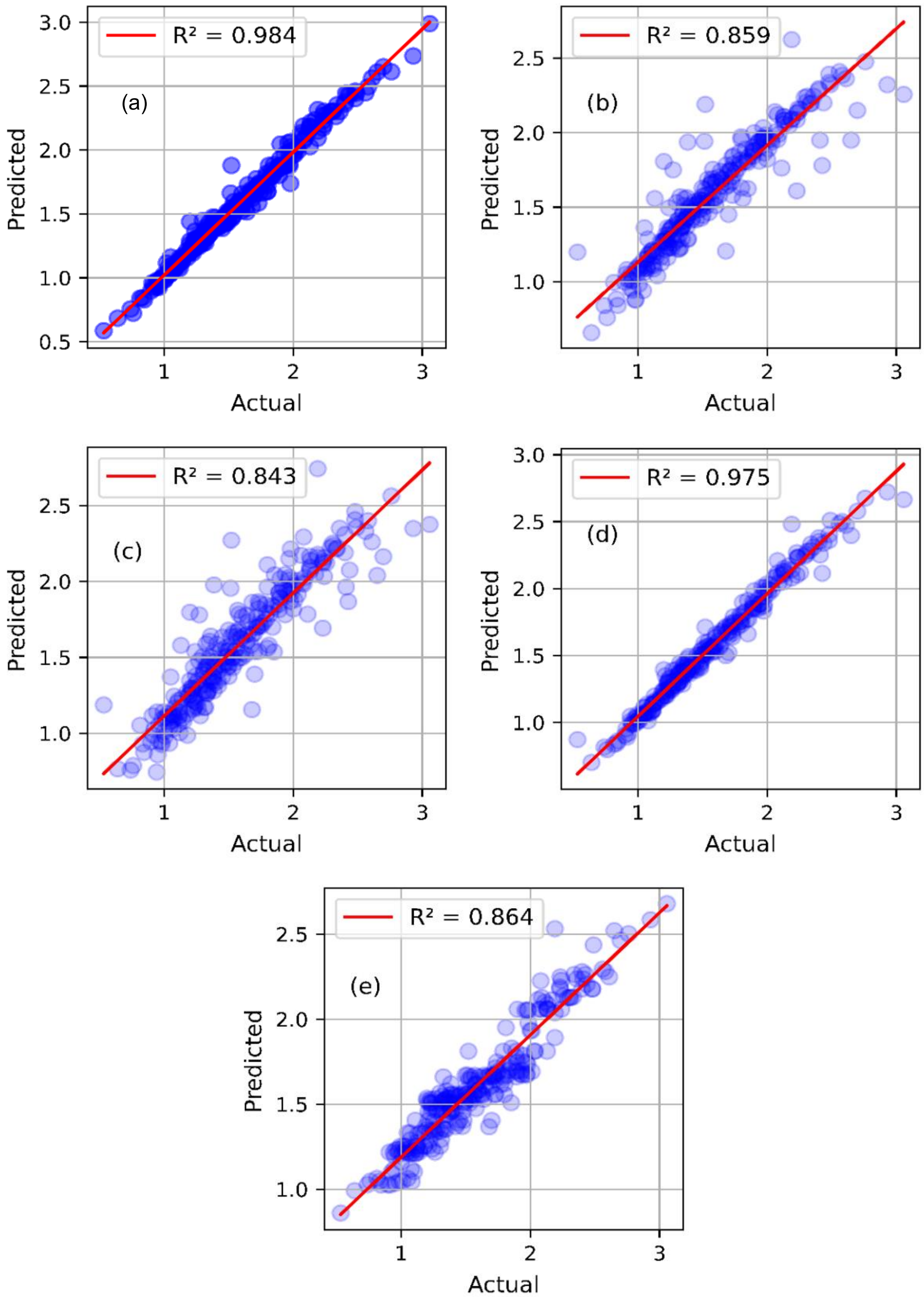

Figure 3. R^2 values for the training: a) GB; b) SVR; c) MLP; d) RF; e) AB

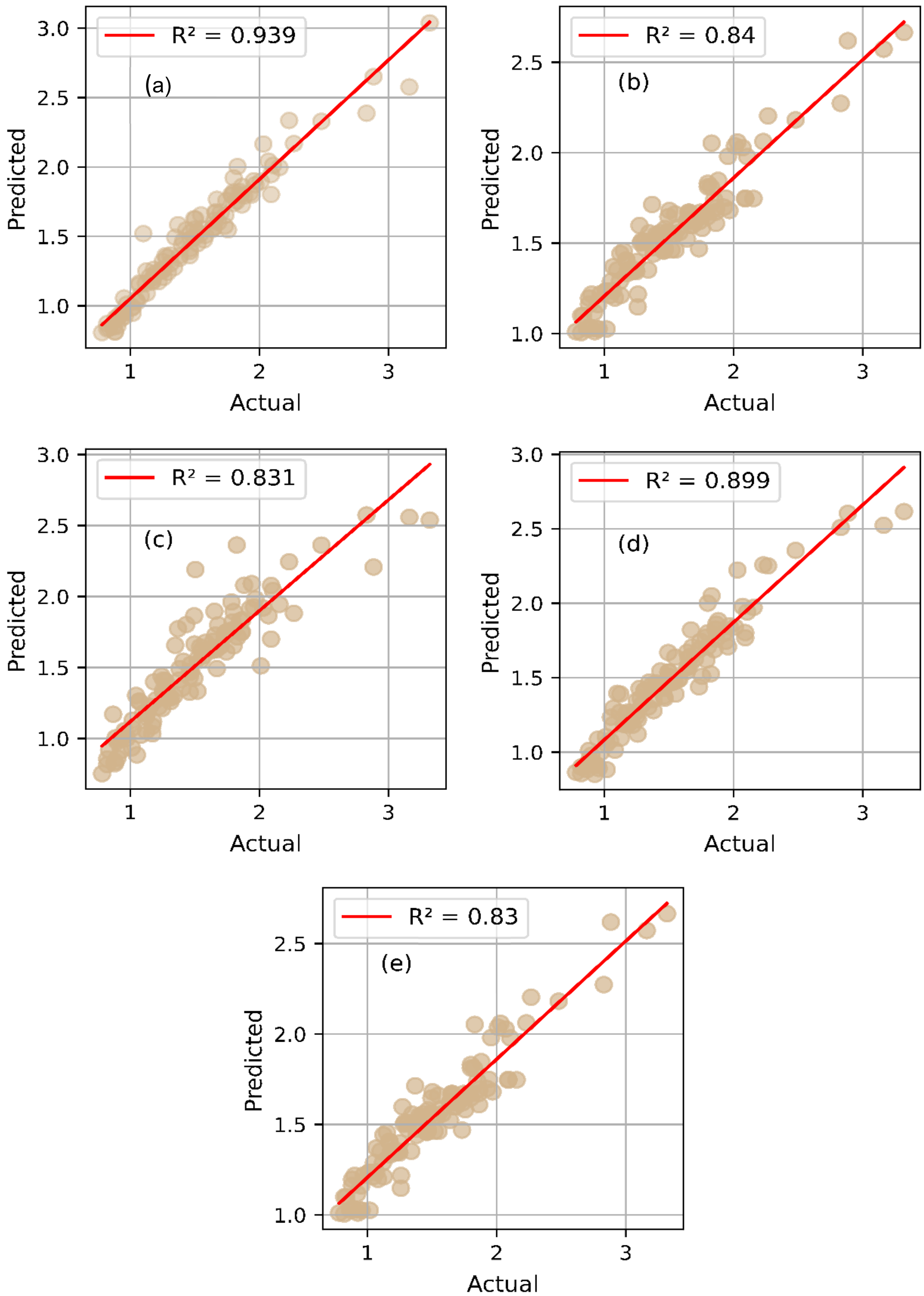

Figure 4. R^2 values for the testing: a) GB; b) SVR; c) MLP; d) RF; e) AB

Table 3. Validation of the models using MAE and RMSE indicators

No.	Parameter	Training	Testing
MAE			
1	GB	0.04	0.079
2	SVM	0.108	0.120
3	MLP	0.121	0.130
4	RF	0.047	0.108
5	AB	0.141	0.158
RMSE			
1	GB	0.059	0.100
2	SVM	0.173	0.194
3	MLP	0.183	0.200
4	RF	0.073	0.108
5	AB	0.170	0.200

The analysis of results showed that the GB model outperformed other models, such as SVM, RF, MLP, and AB, for predicting the FS of slope stability in this study. It is reasonable as the GB model is a sophisticated ensemble learning approach that systematically corrects the shortcomings of single models through an iterative process, particularly in the context of predicting the safety factor of slope stability, where nonlinear relationships dominate, and interactions among multiple variables are complex (Karir et al., 2022).

GB has a decisive advantage due to its iterative process (Bentéjac et al., 2021). Unlike RF, which averages predictions from multiple independent decision trees, GB builds trees one after another. Each new tree is trained to fix the errors made by the previous tree. This step-by-step approach helps GB identify complex patterns and nonlinear relationships among input variables, such as slope gradient, soil composition, and moisture levels. On the other hand, RF treats each tree separately and may not capture these relationships either. This iterative improvement is why GB often outperforms RF in complex scenarios, like predicting slope stability.

GB also effectively minimizes bias and variance, enhancing its predictive ability (Karir et al., 2022). In contrast, AdaBoost (AB) mainly focuses on reducing bias by adjusting the weights of misclassified data points in each iteration. However, it does not refine predictions as deeply as GB. AB combines weak learners, which may be less effective in problems with significant nonlinearity, such as slope stability modeling. GB reduces bias and manages variance through its gradual learning process, making it better at balancing underfitting and overfitting than AB.

Additionally, SVM can perform well in high-dimensional spaces but struggles with variable interactions in regression tasks unless finely tuned with suitable kernels. This tuning can be resource-intensive, and SVM's need for proper kernel selection makes it less flexible than GB, which automatically manages these complexities (Bentéjac et al., 2021).

Furthermore, GB focuses on the hardest-to-predict instances (Samaei et al., 2024). Each iteration gives more weight to data points where the previous model struggled. This focus is essential in slope stability prediction, where certain geotechnical conditions or outliers can be challenging to predict accurately.

Although MLP is powerful for modeling complex, nonlinear data, it can act as a "black box" and is more prone to overfitting without proper regularization (Lapuschkin, 2019). Research shows that GB excels with moderate-sized and noisy datasets, making it suitable for slope stability predictions. MLP can perform well in large-scale and complex environmental modeling, but only if sufficient data is available.

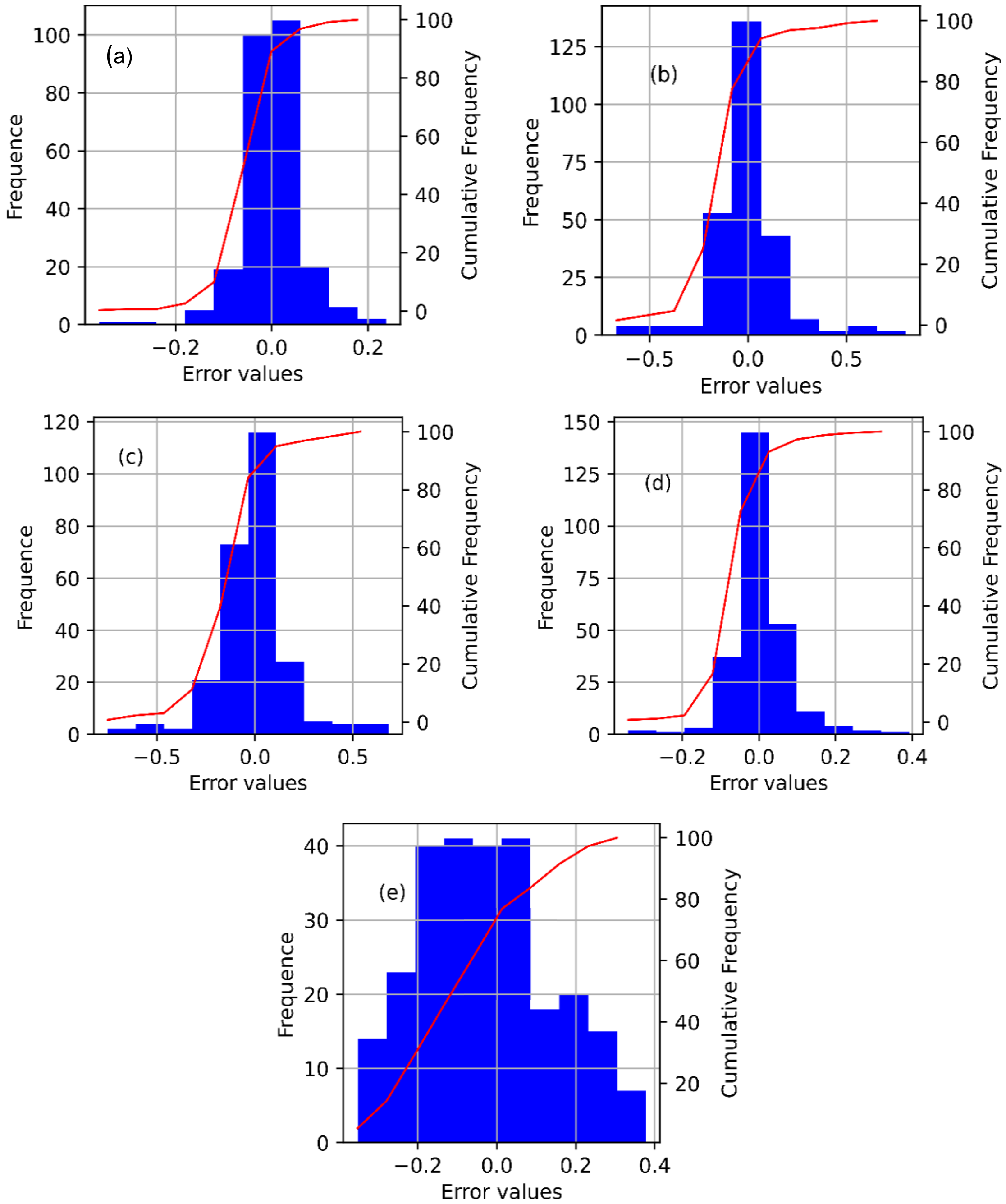

Figure 5. RMSE values with a) GB; b) SVR; c) MLP; d) RF; e) AB

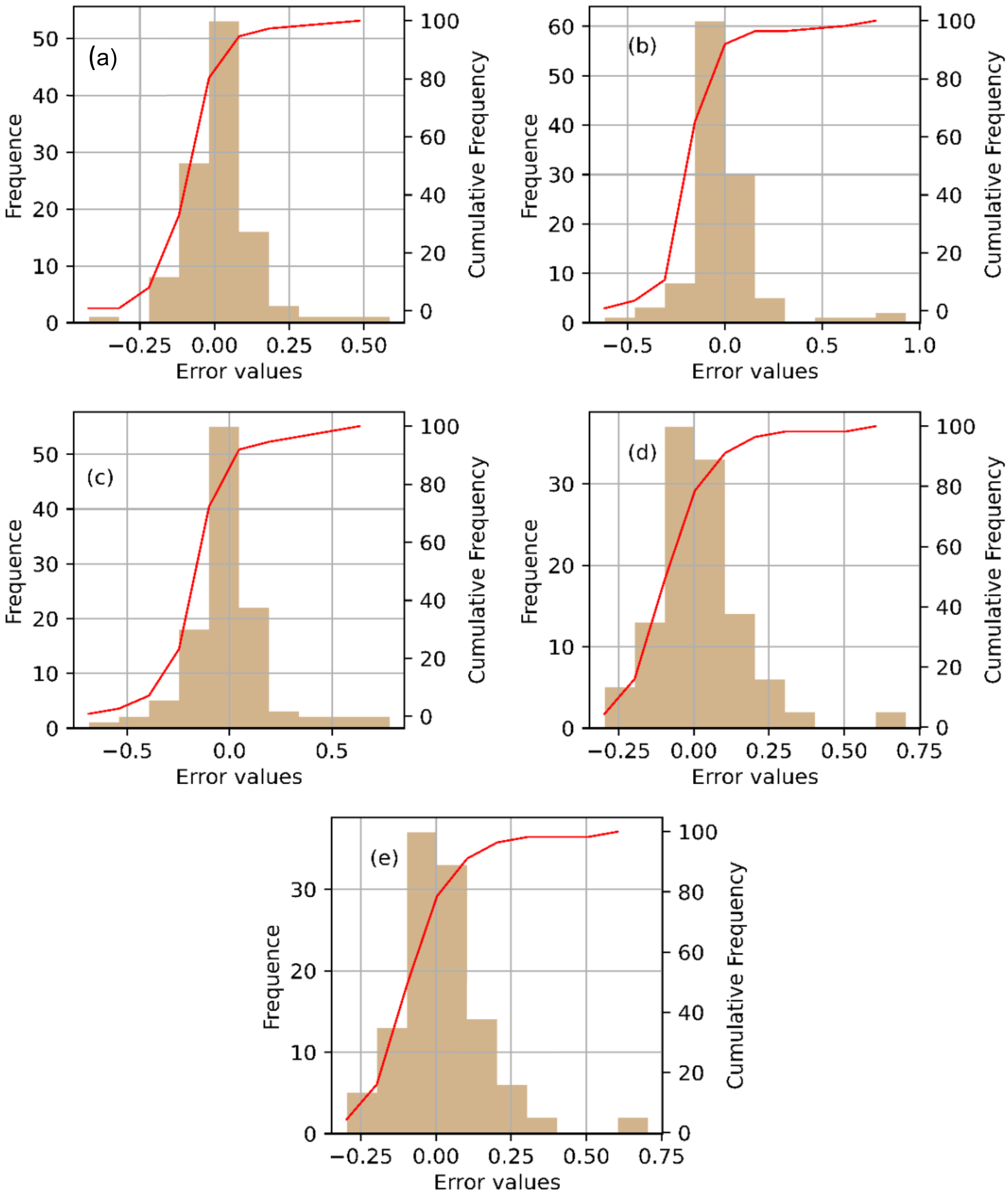

Figure 6. RMSE values with a) GB; b) SVR; c) MLP; d) RF; e) AB

3.2. SHAP analysis of the importance of input variables

The SHAP analysis evaluated the importance of input variables on the predictive capability of the best model GB (Figs. 7 and 8). It shows the significant influence of three critical parameters, including X1, X2, and X3, on predicting the FS of slope stability. X1 emerges as the most critical factor, with a mean SHAP value of 0.28, indicating that it has the most decisive impact on the model's output. This result is expected X1 directly influences the balance between driving and resisting forces acting on a slope. Steeper slopes are inherently more prone to instability due to the increased gravitational force acting along the slope surface, which explains the dominant influence of this feature in determining the safety factor. X2 also plays a significant role in slope stability, with a mean SHAP value of 0.13. Taller slopes generally carry more potential energy, which increases the driving forces that can lead to slope failure. The greater the height, the more critical the prediction becomes, as taller slopes may experience more significant deformations or failure mechanisms under certain conditions. This explains why X2 is another essential feature in the model, contributing significantly to the stability prediction. X3, with a mean SHAP value of 0.11, also plays an important role, though slightly less than slope ratio and height. The number of steeps refers to the abrupt changes in slope angle along the slope profile, which can introduce localized weaknesses or points of failure. A higher number of steeps can create areas of concentrated stress, increasing the likelihood of slope instability. The model's recognition of this variable's impact reflects the practical importance of accounting for complex slope geometries when assessing stability. However, The remaining variables exhibit minimal influence on the model's output, with SHAP values close to zero, suggesting that they either do not have a significant relationship with slope stability or their effects are negligible compared to the more dominant features.

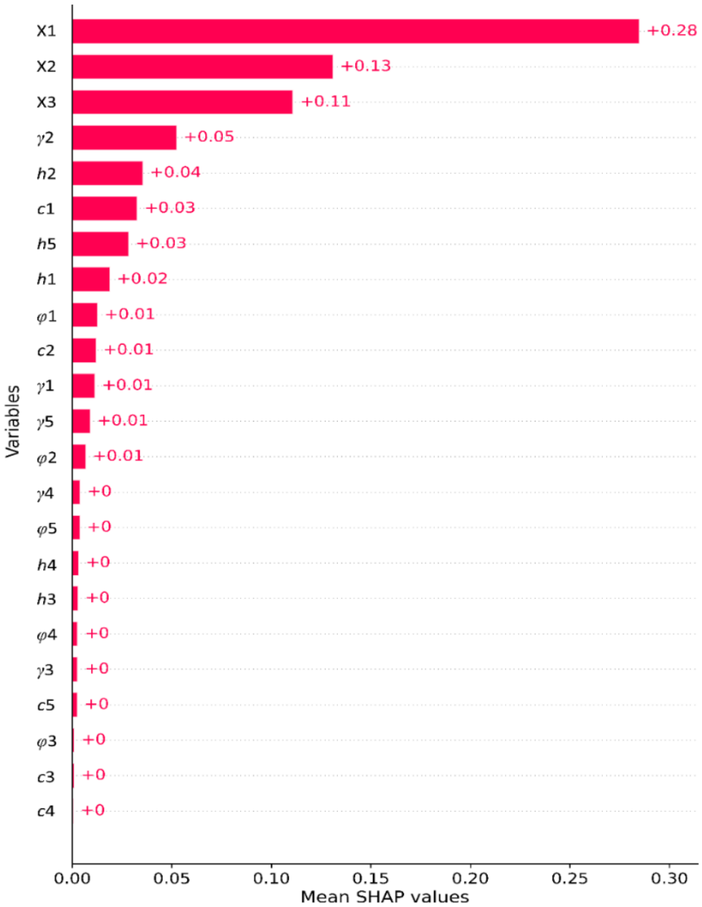

Figure 7. The partial SHAP dependence plot highlights the impact of the three representative features and the most influential result

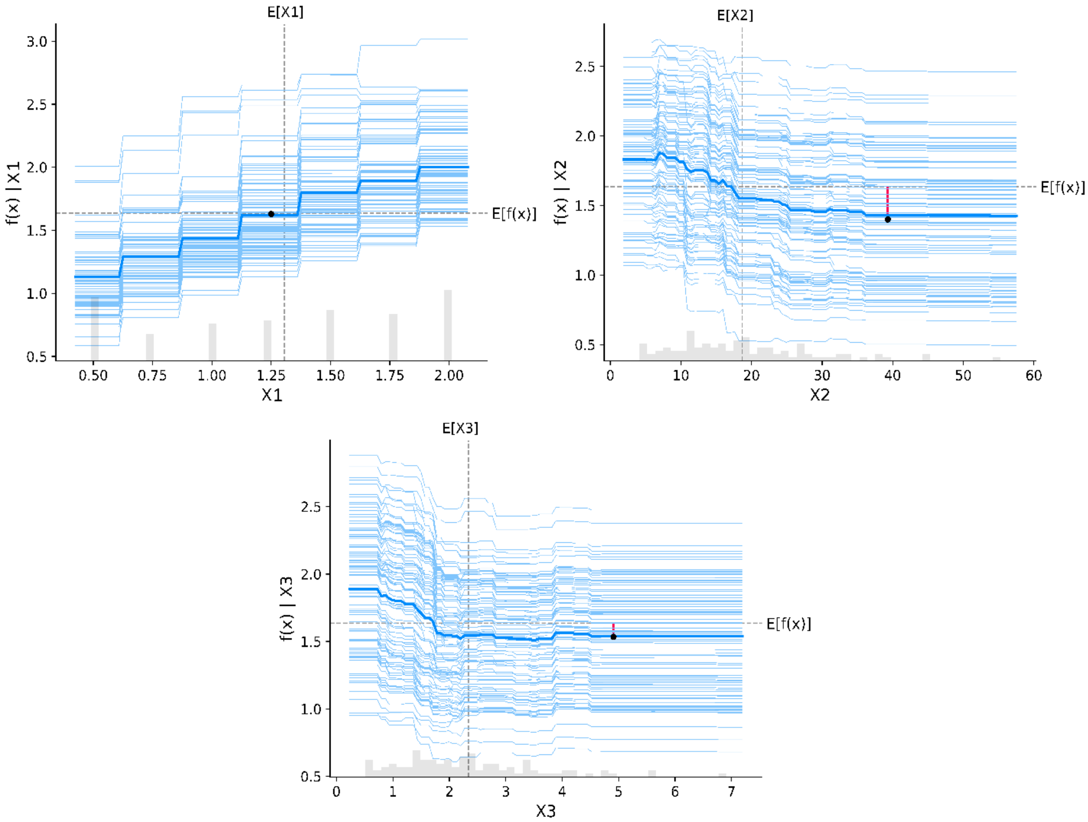

Figure 8. Importance of variables (X1, X2, and X3) with GB using SHAP analysis

In general, these results highlight the critical role of the top three features (X1, X2, and X3) in determining slope stability while also suggesting that lower-ranked features could potentially be excluded from future models to streamline the prediction process without compromising accuracy. Therefore, it can be stated that steeper and taller slopes with more abrupt changes in angle are inherently more susceptible to failure.

4. Conclusions

This study evaluated the performance of several machine learning models, including Gradient Boosting (GB), Random Forest (RF), AdaBoost (AB), Multi-layer Perceptron (MLP) Neural Networks, and Support Vector Machine (SVM), in predicting the Factor of Safety (FS) for slope stability. The results demonstrated that the GB model outperformed the other models in terms of prediction accuracy, with an R^2 of 0.975, MAE of 0.079, and RMSE of 0.120. GB's ability to iteratively refine its predictions made it particularly effective at capturing complex interactions between geotechnical and topographical variables, contributing to its superior performance. On the other hand, AB exhibited the weakest performance, struggling to adequately model the intricate relationships inherent in the slope stability data.

These findings underscore the importance of selecting the appropriate machine-learning model for slope stability analysis. While RF and SVM provided acceptable predictions, GB's capacity to model variable interactions and focus on challenging cases made it the most robust and precise option for this study.

However, the study has several limitations. The relatively small dataset, consisting of 371 slope stability cases, may not fully encompass the range of geotechnical conditions influencing slope behavior, potentially limiting the generalizability of the models. A more extensive and diverse dataset, including additional geotechnical and environmental parameters, would likely improve the performance and robustness of the models. Furthermore, while GB performed well, its complexity and susceptibility to overfitting remain concerning, particularly with limited data. Future work should focus on fine-tuning model parameters, such as learning rates and tree depth, to further enhance prediction accuracy and prevent overfitting.

We recommend expanding the dataset to incorporate a broader spectrum of geotechnical, environmental, and topographical variables for future research. Integrating real-time data like weather conditions and rainfall could significantly improve the models' predictive accuracy. Additionally, exploring hybrid models that combine machine learning with traditional geotechnical analysis, or investigating deep learning architectures like Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), may offer more comprehensive insights into the complex dynamics of slope stability, ultimately leading to more reliable predictions of the FS.

Acknowledgements

This study is funded by the project No. NVCC11.02/22-23.

References

Abdullah D.M., Abdulazeez A.M., 2021. Machine learning applications based on SVM classification a review. Qubahan Academic Journal, 1(2), 81–90.

Ali J., Khan R., Ahmad N., Maqsood I., 2012. Random forests and decision trees. International Journal of Computer Science Issues (IJCSI), 9(5), 272.

Ben Ishak A., 2016. Variable selection using support vector regression and random forests: A comparative study. Intelligent Data Analysis, 20(1), 83–104.

Bentéjac C., Csörgő A., Martínez-Muñoz G., 2021. A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review, 54, 1937–1967.

Chang Z., Catani F., Huang F., Liu G., Meena S.R., Huang J., Zhou C., 2023. Landslide susceptibility prediction using slope unit-based machine learning models considering the heterogeneity of conditioning factors. Journal of Rock Mechanics and Geotechnical Engineering, 15(5), 1127–1143.

Cheng Y.M., 2003. Location of critical failure surface and some further studies on slope stability analysis. Computers and Geotechnics, 30(3), 255–267.

Cho S.E., 2007. Effects of spatial variability of soil properties on slope stability. Engineering Geology, 92(3-4), 97–109.

Collison A., Anderson M., 1996. Using a combined slope hydrology/stability model to identify suitable conditions for landslide prevention by vegetation in the humid tropics. Earth Surface Processes and Landforms, 21(8), 737–747.

Fell R., Corominas J., Bonnard C., Cascini L., Leroi E., Savage W.Z., 2008. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Engineering Geology, 102(3-4), 85–98.

Friedman J.H., 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics, 1189–1232.

Fryer D., Strümke I., Nguyen H., 2021. Shapley values for feature selection: The good, the bad, and the axioms. Ieee Access, 9, 144352–144360.

Ghosh J., Nag A., 2001. An overview of radial basis function networks. Radial basis function networks 2: new advances in design, 1–36.

Hammouri N.A., Malkawi A.I.H., Yamin M.M., 2008. Stability analysis of slopes using the finite element method and limiting equilibrium approach. Bulletin of Engineering Geology and the Environment, 67, 471–478.

Hicks S.A., Strümke I., Thambawita V., Hammou M., Riegler M.A., Halvorsen P., Parasa S., 2022. On evaluation metrics for medical applications of artificial intelligence. Scientific Reports, 12(1), 5979.

Jin H., Kim Y.-G., Jin Z., Rushchitc A.A., Al-Shati A.S., 2022. Optimization and analysis of bioenergy production using machine learning modeling: Multi-layer perceptron, Gaussian processes regression, K-nearest neighbors, and Artificial neural network models. Energy Reports, 8, 13979–13996.

Karir D., Ray A., Bharati A.K., Chaturvedi U., Rai R., Khandelwal M., 2022. Stability prediction of a natural and man-made slope using various machine learning algorithms. Transportation Geotechnics, 34, 100745.

Khanna R., Dubey R., 2021. Comparative assessment of slope stability along road-cuts through rock slope classification systems in Kullu Himalayas, Himachal Pradesh, India. Bulletin of Engineering Geology and the Environment, 80(2), 993–1017.

Kolapo P., Oniyide G.O., Said K.O., Lawal A.I., Onifade M., Munemo P., 2022. An overview of slope failure in mining operations. Mining, 2(2), 350–384.

Kumar N., Kumari S., 2024. State-of-the-art advanced hybrid ANNs paradigm for assessment and prediction of slope stability. Multiscale and Multidisciplinary Modeling, Experiments and Design, 1–27.

Lapuschkin S., 2019. Opening the machine learning black box with layer-wise relevance propagation. Dissertation, Berlin, Technische Universität Berlin.

Lin M., Zeng L., Teng S., Chen G., Hu B., 2024. Prediction of stability of a slope with weak layers using convolutional neural networks. Natural Hazards, 1–25.

Lin S., Zheng H., Han B., Li Y., Han C., Li W., 2022. Comparative performance of eight ensemble learning approaches for the development of models of slope stability prediction. Acta Geotechnica, 17(4), 1477–1502.

Liu S., Su Z., Li M., Shao L., 2020. Slope stability analysis using elastic finite element stress fields. Engineering Geology, 273, 105673.

Liu X., Wang Y., Li D.-Q., 2019. Investigation of slope failure mode evolution during large deformation in spatially variable soils by random limit equilibrium and material point methods. Computers and Geotechnics, 111, 301–312.

Lundberg S.M., Lee S.-I., 2017. A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 10, 7102.

Malik M.K., Karim I.R., 2020. Seepage and slope stability analysis of Haditha Dam using Geo-Studio Software. In: IOP conference series: materials science and engineering. IOP Publishing, 2, 022074.

Meng J., Mattsson H., Laue J., 2021. Three-dimensional slope stability predictions using artificial neural networks. International Journal for Numerical and Analytical Methods in Geomechanics, 45(13), 1988–2000.

Moayedi H., Tien Bui D., Kalantar B., Kok Foong L., 2019. Machine-learning-based classification approaches toward recognizing slope stability failure. Applied Sciences, 9(21), 4638.

Nguyen H., Vu T., Vo T.P., Thai H.-T., 2021. Efficient machine learning models for prediction of concrete strengths. Construction and Building Materials, 266, 120950.

Pisner D.A., Schnyer D.M., 2020. Support vector machine. In: Machine learning. Elsevier, 101–121.

Rabbani A., Samui P., Kumari S., 2023. Implementing ensemble learning models for the prediction of shear strength of soil. Asian Journal of Civil Engineering, 24(7), 2103–2119.

Rasti A., Adarmanabadi H.R., Pineda M., Reinikainen J., 2021. Evaluating the effect of soil particle characterization on internal friction angle. American Journal of Engineering and Applied Sciences, 14(1), 129-138.

Saccenti E., Hendriks M.H., Smilde A.K., 2020. Corruption of the Pearson correlation coefficient by measurement error and its estimation, bias, and correction under different error models. Scientific reports, 10(1), 438.

Samaei M., Alinejad Omran M., Keramati M., Naderi R., Shirani Faradonbeh R., 2024. Assessing the shear strength of sandy soil reinforced with polyethylene-terephthalate: an AI-based approach. Earth Science Informatics, 1–20.

Schjønning P., Lamandé M., Keller T., Labouriau R., 2020. Subsoil shear strength-Measurements and prediction models based on readily available soil properties. Soil and Tillage Research, 200, 104638.

Sekhar C., Meghana P.S., 2020. A study on backpropagation in artificial neural networks. Asia-Pacific Journal of Neural Networks and Its Applications, 4(1), 21–28.

Shrestha D.L., Solomatine D.P., 2006. Experiments with AdaBoost. RT, an improved boosting scheme for regression. Neural Computation, 18(7), 1678–1710.

Statham I., Francis S., 2020. Influence of scree accumulation and weathering on the development of steep mountain slopes. Hillslope Processes, 24, 5–267.

Wang G., Zhao B., Wu B., Zhang C., Liu W., 2023. Intelligent prediction of slope stability based on visual exploratory data analysis of 77 in situ cases. International Journal of Mining Science and Technology, 33(1), 47–59.

Wang L., Wu C., Tang L., Zhang W., Lacasse S., Liu H., Gao L., 2020. Efficient reliability analysis of earth dam slope stability using extreme gradient boosting method. Acta Geotechnica, 15, 3135–3150.

Wang W., Griffiths D., Wang S., 2022. Stability Analyses and Charts for Undrained Slopes with Linearly Increasing Strength and an External Water Surface. International Journal of Geomechanics, 22(8), 06022020.

Xie H., Dong J., Deng Y., Dai Y., 2022. Prediction model of the slope angle of rocky slope stability based on random forest algorithm. Mathematical Problems in Engineering, 1, 9441411.

Yin S., Li H., Sun Y., Ibrar M., Teng L., 2024. Data Visualization Analysis Based on Explainable Artificial Intelligence: A Survey. IJLAI Transactions on Science and Engineering, 2(2), 13–20.

Yuan B., Li Z., Su Z., Luo Q., Chen M., Zhao Z., 2021. Sensitivity of multistage fill slope based on finite element model. Advances in Civil Engineering, 1, 6622936.

Zhang F., Peng J., Huang X., Lan H., 2021a. Hazard assessment and mitigation of non-seismically fatal landslides in China. Natural Hazards, 106, 785–804.

Zhang X., Li Y., Liu Y., Huang Y., Wang Y., Lu Z., 2021b. Characteristics and prevention mechanisms of artificial slope instability in the Chinese Loess Plateau. Catena, 207, 105621.