STU SvF
SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA
FACULTY OF CIVIL ENGINEERING

Slovak Journal of Civil Engineering

Vol. 33, 2025, No. 4, 1-12
DOI: 10.2478/sjce-2025-0022

ASSESSMENT OF STABLE SLOPES THROUGH BPSO-DRIVEN ENSEMBLE MODELS

Saurabh Kumar ANURAGI [1], D.KISHAN [1]

Abstract

This study explores a hybrid approach that combines BPSO with ensemble machine learning techniques to improve predictive accuracy in assessments of slope stability. The methodology employs BPSO to optimize the selection of features that are critical to the prediction process. In addition, a grid search technique is utilized to fine-tune the hyperparameters of the ensemble models. The research evaluates the performance of three ensemble models: RF, XGBoost, and LightGBM. For the predictive analysis, six features identified as potentially influential were selected, including: slope height (H), pore water ratio (ru), unit weight (ϒ), cohesion (c), slope angle (β), and angle of internal friction (Φ). The effectiveness of the models was assessed using various performance metrics, including the AUC, kappa and accuracy of the predictions. The findings indicate that the hybrid approach, particularly the LightGBM model, significantly outperformed the other models, achieving an AUC of 0.871, a kappa of 0.658 and an accuracy rate of 0.832. This underscores the potential of the proposed hybrid method as a valuable tool for accurately predicting slope stability and mitigating risks associated with slope failures in engineering applications.

Address

[1] Dept. of Civil Engineering, Maulana Azad National Institute of Technology, Bhopal, India

[*] ***Corresponding author:*** saurabh.1399@gmail.com

Key words

- Random forest,
- Xgboost,
- Lightgbm,
- Slope stability,
- Machine learning.

1 INTRODUCTION

Slopes are defined as inclined surfaces that can range from gentle inclines, such as rolling hills, to steep escarpments like cliffs or mountainsides. The characteristics of these slopes are influenced by a variety of factors, including the geomorphological features of the landscape, the geological composition of the underlying materials, and the hydrological conditions present in the area. Slope engineering involves the systematic analysis and identification of potentially weak zones within a slope to mitigate the risk of slope failures, which can pose significant dangers to both human life and property. By understanding the underlying causes of slope instability, engineers and geologists can develop effective strategies to reinforce these areas and prevent collapses or landslides. The study of slope stability focuses on evaluating the effects of shear stress on different slopes. Areas experiencing higher shear stress tend to correspond with steeper slopes, which makes them more susceptible to failure. Conversely, gentler slopes typically exhibit lower shear stress, thereby resulting in greater stability and a reduced likelihood of collapse (Huang, 2014). Another crucial aspect of slope stability is the condition of the rock and soil materials involved. Steeper slopes face increased vulnerability to failure due to the dual pressures of gravitational forces and the inherent instability of the soil. Loose or saturated soil on steep inclines can lead to dangerous landslides if not properly managed (Nanehkaran et al., 2023).

Additionally, vegetation plays a significant role in slope stability. Areas with high vegetation density often experience increased soil stabilization, as roots bind the soil particles together, which reduces erosion and enhances cohesion. The presence of plant life not only helps to retain moisture in the soil but also contributes to the overall health of the landscape, which makes slopes less prone to failure.

Frequent slope failures have resulted in significant human and property losses; this prompts researchers to explore effective mitigation measures. To implement stabilization strategies, it is essential to conduct a comprehensive slope stability analysis. The complexity of slope engineering arises from the highly non-linear relationships among various influencing factors, thereby characterizing it as a non-linear dynamic system. Various traditional techniques, including the limit equilibrium method (Cheng et al., 2007; Liu et al., 2015), finite element method (Matsui and San., 1992; Griffiths and Lane., 1999; Zheng et al., 2005; Li., 2007), and discontinuous deformation analysis (Zheng et al., 2019; Zhuang et al., 2021), have been employed by researchers for assessing slope stability. However, these methods often exhibit limitations due to the intricate mechanisms involved and can impose a considerable computational burden (Dawson et al., 1999; Abramson et al., 2001; Duncan et al., 2014). Specifically, the stability calculation processes can be intricate and do not always align with the need for expedited slope design solutions. To address these challenges, appropriate techniques are being utilized to minimize potential losses by employing both analytical and numerical modelling, which aim to provide accurate and reliable predictions for slope stability.

Recent advancements in computational techniques have led many researchers to explore machine learning methods as viable alternatives to slope stability analysis. These approaches facilitate the modelling of complex relationships among various influencing factors, including soil geometry, soil properties, and hydrological conditions, through a data-driven framework (Anuragi., 2024). Researchers have effectively utilized these models to develop predictive tools that enhance the assessment of slope stability. Qi et al. (2018) conducted a comparative study focusing on slope stability prediction using various techniques, including support vector machines, Gaussian process classifiers, discriminant analysis, adaptive boosting decision trees, artificial neural networks, and k-nearest neighbor, with optimization of the models achieved through genetic algorithms. Similarly, Khajehzadeh et al. (2012) employed a gravitational search algorithm for predicting slope stability. Luo et al. (2021) introduced a PSO-CA (particle swarm optimization-cubist algorithm) hybrid model to predict the safety factor and compare its performance with that of support vector machines, CART, and k-nearest neighbors. Rukhaiyar et al. (2018) utilized a PSO-ANN hybrid technique for predicting safety factors, while Sakellariou et al. (2005) implemented neural networks for the same purpose. Xue et al. (2014) optimized a support vector machine using particle swarm optimization as a hybrid model for slope stability prediction. Nanehkaran et al. (2022) conducted a comparative analysis involving random forests, k-nearest neighbors, multilayer perceptrons, and support vector machines for safety factor prediction. Koopialipoor et al. (2019) performed a comparative study using genetic algorithms, artificial bee colonies, particle swarm optimization, and the imperialist competitive algorithm; they integrated them with artificial neural networks as hybrid models for predicting slope stability. Das et al. (2011) utilized a differential evolution neural network for safety factor prediction, while Lin et al. (2022) engaged in a comparative analysis of various techniques, including hist gradient boosting, AdaBoost, random forests, bagging, extra trees, gradient boosting machines, voting, and stacking for assessing slope stability. All the machine learning techniques previously mentioned contribute significantly to our understanding of the complexity and behavior of slopes. Despite these contributions, the complexity of the slope stability problem continues to present variations, which can be attributed to the limitations inherent in each model employed. Each machine learning model comes with its own strengths and weaknesses, which can affect its performance in accurately predicting slope stability. The primary objective remains to develop machine learning models that are not just accurate but also reliable for the analysis of slope stability. To achieve this, it is essential to explore and identify more advanced models that offer higher levels of precision and consistency in their predictions. By focusing on the pursuit of such sophisticated and high-accuracy models, we can aim for superior results that enhance our understanding and management of slope stability issues. This continued research and development in machine learning methods is crucial for more effective and practical applications in real-world scenarios related to slope stability.

The behaviour of slope stability is influenced by a variety of key factors that determine the overall stability of the soil. These factors include height (H), pore water ratio (ru), unit weight (ϒ), cohesion (c), slope angle (β), and the angle of internal friction (ϕ). Each of these parameters plays a significant role in assessing the shear strength of soil, which is crucial for understanding how well a slope can withstand various forces. Among these parameters, unit weight (ϒ), cohesion (c), and the angle of internal friction (ϕ) are particularly critical as they directly impact the soil's shear strength. The interplay among these factors creates a complex environment and makes it essential to thoroughly understand their interactions when assessing slope stability. The study conducted by (Lin et al., 2021) systematically examines these features by eliminating each one in turn, which allows for an evaluation of how variations in each parameter affect the factor of safety (FOS) predictions. This approach employs a hit-and-trial method for slope stability predictions, thereby offering insights into the relationship between individual parameters and overall slope stability. The present study aims to fill a notable gap in the existing literature by adopting the Binary Particle Swarm Optimization technique (BPSO). This method evaluates multiple machine learning models to identify the most crucial features that should be taken into account for a comprehensive analysis of slope stability. Furthermore, optimizing hyper-parameter settings is vital, as these parameters influence the performance and accuracy of the models developed. To achieve this, grid search optimization techniques are employed. This method systematically explores a predefined set of hyper-parameter values and evaluates the model's performance for each combination through cross-validation. By analyzing the results, we can identify the optimal hyper-parameter settings that maximize the model's accuracy and enhance its predictive capabilities. This thorough approach ensures that the final model is not just effective but also robust in handling various data scenarios.

2 MACHINE LEARNING TECHNIQUE

2.1 Random Forest

Random Forest is an ensemble learning method that employs a bagging approach to enhance the accuracy of predictions. Initially, it extracts multiple subsets of training data, upon which distinct decision trees are developed. The final prediction is derived by aggregating the outputs from these individual decision trees. Each tree is constructed using a random selection of independent variables and a distinct subset of observations, which helps to minimize correlation among the trees and thereby improve their collective predictive performance. A significant advantage of the Random Forest algorithm is its capability to efficiently manage a large number of input variables, as well as its effectiveness in capturing complex interactions among those variables (Belgiu and Drăguţ., 2016; Biau and Scornet., 2016; Puri et al., 2018; Pham et al., 2020).

2.2 XGBoost

XGBoost is a powerful machine learning algorithm that combines regression trees with gradient boosting for high predictive accuracy. It builds a series of decision trees with each one correcting the errors of the previous tree and improving the model's performance iteratively. The algorithm calculates residuals, i.e., the differences between the predicted and actual values at each stage to train the next tree to minimize these residuals. XGBoost enhances traditional gradient boosting with features like regularization (L1 and L2 penalties) to reduce a model's complexity and prevent overfitting. It efficiently handles missing data, sparse matrices, and supports parallel computation. The final model is an ensemble of all the trees, where each tree contributes based on its performance; this leads to accurate predictions. This versatility makes XGBoost a popular choice in classification, regression, and ranking tasks, especially with large datasets (Natekin and Knoll., 2013).

2.3 LightGBM

LightGBM is a highly efficient and precise machine learning technique designed for handling large datasets with intricate features. It operates based on the principles of gradient boosting and focuses on minimizing the residual errors from previous iterations while conducting training in a sequential manner. This method facilitates faster convergence and enhanced accuracy by prioritizing the splitting of leaves that exhibit the greatest reduction in losses. The implementation of a leaf-wise growth strategy, while potentially leading to imbalanced trees, is effectively addressed through sophisticated regularization techniques, including L1 and L2 penalties, maximum tree depth constraints, and minimum data requirements per leaf. To optimize computational efficiency and reduce memory usage, LightGBM discretizes continuous features into bins through a histogram-based decision tree algorithm. During the training process, it incorporates gradient-based one-side sampling (GOSS), a technique that reduces dataset size while preserving critical information. Furthermore, to enhance performance in high-dimensional data contexts, LightGBM employs Exclusive Feature Bundling (EFB), which combines sparse features with mutually exclusive non-zero values into singular features. This significantly accelerates processing without compromising the integrity of the data. The framework also supports parallel and distributed training; this makes it highly scalable for large datasets and well-suited for big data applications. With these innovative mechanisms, LightGBM stands out in terms of speed, scalability, and accuracy (Ke et al., 2017; Sun et al., 2020; Alshboul et al., 2024).

2.4 Binary Particle Swarm Optimization (BPSO)

Binary Particle Swarm Optimization (BPSO) is a sophisticated optimization algorithm, specifically designed for tackling binary optimization challenges. In these types of problems, the main objective is to identify the optimal binary string that either maximizes or minimizes a predetermined objective function. BPSO is inspired by the natural behavior of swarming systems, such as flocks of birds or schools of fish; it employs this heuristic approach to navigate a multi-dimensional search space effectively. In BPSO, each particle represents a potential solution and is characterized by a position vector in a binary space, where each element can only assume the values of 0 or 1. When the algorithm updates a particle's position, it involves flipping one or more bits in the particle's binary representation. This flipping action enables the particle to effectively traverse the various corners of the hypercube that defines the binary search space. Each modification of the bits can lead to the particle moving either closer to the optimal solution or further away, which makes the search process dynamic and exploratory. This continuous evolution of particle positions allows BPSO to search for the best possible solution within the constraints of binary optimization effectively (Khanesar et al., 2007).

3 MATERIALS AND METHODOLOGY

3.1 Data Collection and Preprocessing

The data utilized in the current study are sourced from (Lin et al., 2022) and contains documented slope cases drawn from published geotechnical literature, primarily in China and other parts of Asia. Six influential features have been identified as critical parameters for predicting slope stability, namely: pore water pressure ratio (ru), slope height (H), unit weight (ϒ), cohesion (c), slope angle (β), and the angle of internal friction (ϕ). The dataset consists of 444 samples, which have been categorized into stable (1) and unstable (0) classes, as illustrated in Fig. 1. The dataset includes both natural slopes, which are affected by rainfall and weathering, and man-made slopes, such as road cuts, excavations, and embankments. Stability outcomes were determined either from field observations of actual failures or from back-analyses of reported case histories. While the dataset provides a balanced sample of stable and unstable slopes across a range of geometric and geotechnical conditions, it is important to note that it is geographically biased toward Asian contexts and may underrepresent ordinary stable slopes due to its reliance on published failure cases. As such, the dataset is well-suited for developing and benchmarking machine learning models of slope stability; however, local calibration with region-specific data is recommended for broader applications. To enhance the analytical process, the dataset has been normalized. This normalization minimizes potential biases and improves the algorithm's convergence during training, thereby enhancing the model's ability to generalize unseen data samples. Furthermore, this approach facilitates the development of a highly accurate classification model capable of delivering precise results.

The violin plot illustrated in Fig. 2 provides a comprehensive visualization of the distribution and density of the dataset across the various features under consideration. Each violin plot corresponds to a specific feature, with its shape offering valuable insights into the underlying data distribution. The width of the violin at any given point is indicative of the density of the data at that value; the wider sections signify areas with a higher concentration of data points, while the narrower sections reflect regions with a lower density. The horizontal line within each violin denotes the median value of the respective feature; it serves as a reference for the central tendency and aids in the identification of symmetrical distribution or skewness in the data. Furthermore, the presence of tails at both ends of the violin provides a visual indication of the data's range, thus highlighting any potential outliers or extremes.

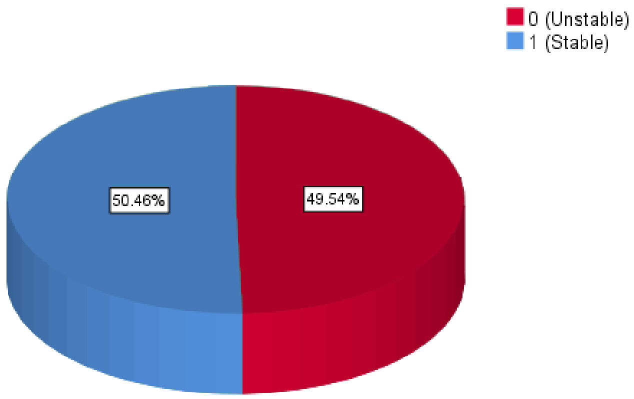


Fig. 1 *Dataset Pie Chart*

The variables ϒ, ϕ, β, and ru demonstrate a broad distribution, as evidenced by the diverse shapes of their corresponding violins. This broadness suggests a significant dispersion of data points across a wide range of values, which indicates a high variability in these features; this could be critical in influencing slope stability. In contrast, the variables c and H reveal a more narrowly shaped violin; this implies that their data points are densely clustered around specific values. This concentration indicates a lower degree of variability and more consistent values compared to the previously mentioned features. The narrow shape highlights a higher frequency of data points near the median, with a reduced occurrence of extreme values. This differentiation underscores that features like ϒ and β may exert a broader influence due to their variability, while features such as c and H may contribute more stable and consistent inputs to the model. These insights are vital for understanding the significance of each feature in the analysis of slope stability and for optimizing the overall performance of the model.

3.2 Model Development and Optimization

This study investigates the effectiveness of three ensemble models, i.e., Random Forest (RF), XGBoost, and LightGBM, in predicting slope stability through an analysis of six influential parameters: pore water pressure ratio (ru), slope height (H), unit weight (ϒ), cohesion (c), slope angle (β), and angle of internal friction (ϕ). The selection of features is conducted using the Binary Particle Swarm Optimization (BPSO) algorithm. The research methodology consists of two distinct stages, each designed to facilitate a comprehensive analysis of slope stability predictions. In Stage 1, the focus is on utilizing the Binary Particle Swarm Optimization (BPSO) algorithm to select the relevant features from the dataset. This feature selection process is crucial for enhancing the performance of the predictive models, as it identifies the most significant variables that contribute to slope stability. The selection of these features is based on their ability to provide meaningful insights and improve the evaluation of the model. Stage 2 involves a deeper evaluation of the selected models using the features identified in the first stage. During this stage, the models are rigorously tested and refined in order to enhance their accuracy in predicting slope stability. The process

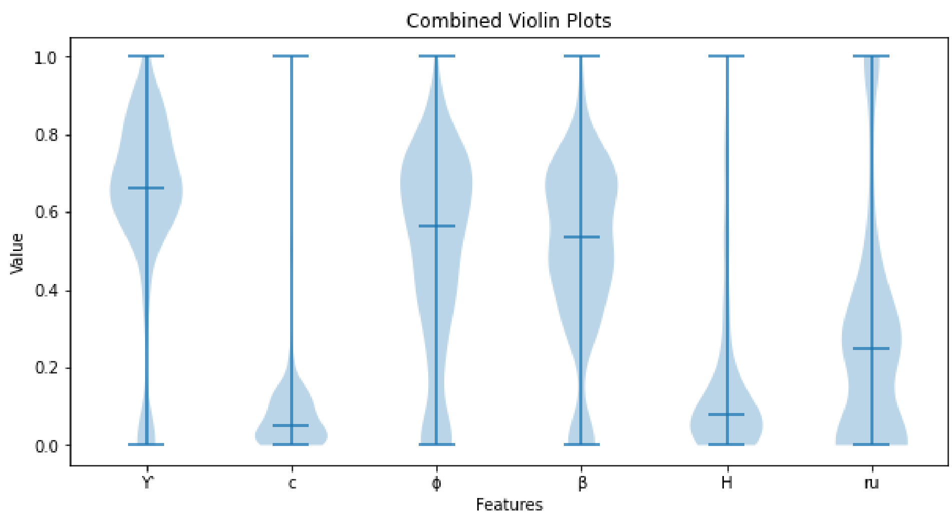


Fig. 2 *Violin Plots showing distribution of slope cases*

includes various assessment metrics and validation techniques to ensure that the predictions are robust and reliable. An illustrative representation of the methodology is shown in Fig. 3, which highlights the systematic approach taken throughout the research. The primary objective of this approach is to enhance the performance of the ensemble models by integrating two essential tasks: the selection of relevant feature subsets from the dataset and the optimization of the ensemble model parameters. Within the BPSO framework, each particle in the swarm represents a potential feature subset, which is indicated by a binary positional vector that denotes the presence (1) or absence (0) of specific features. The classification accuracy of each subset is assessed through a defined fitness function. Notably, the fitness function that achieves the highest degree of accuracy is utilized to evaluate solutions and update particle positions throughout the iterative process. This methodology aims to identify feature subsets that significantly contribute to precise classification outcomes. Upon identifying the most effective subset, the ensemble model is employed. The optimized features derived from BPSO are subsequently utilized to train and test the dataset (as presented in Table 1), with the goal of improving the accuracy of the classification through the tuned ensemble model. The combination of hyperparameters employed by BPSO is detailed in Table 2.

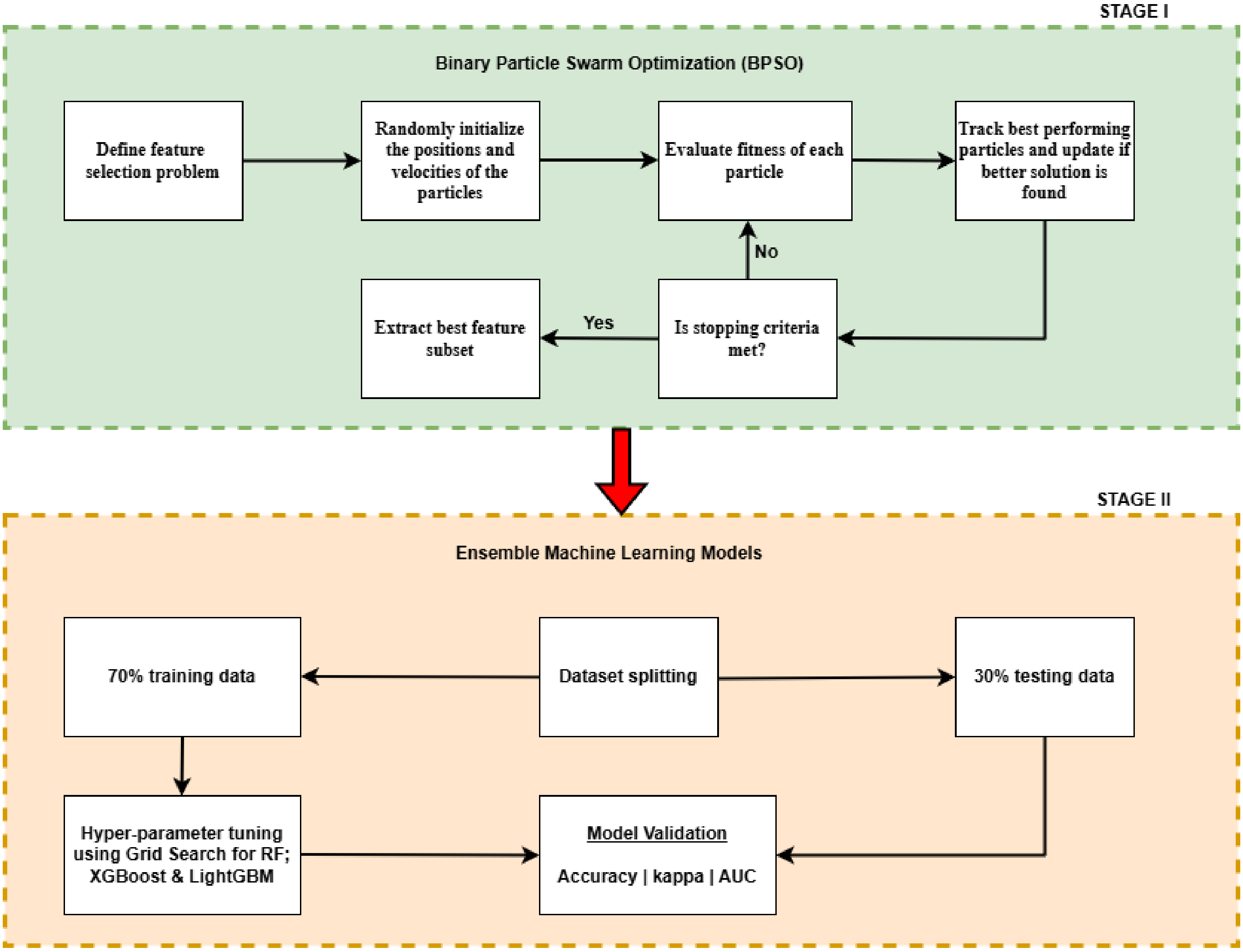


Fig. 3 *Methodology of the study*

Tab. 1 *Feature Optimization using BPSO*

Models	Features					
	Y	c	φ	β	H	ru
RF	1	1	1	1	1	0
XGBoost	1	1	1	0	1	1
LightGBM	1	1	1	0	1	1
0 = unselected feature			1 = selected feature			

Tab. 2 *Hyperparameters utilized for feature optimization using BPSO*

Parameters	Values
Acceleration coefficients (c1,c2)	[2,2]
Inertia Weight (w)	0.9
Number of dimensions (k)	7
Number of particles (n_particles)	50
Iterations (iter)	500
alpha	[1.0]

In supervised machine learning, it is essential to evaluate the performance of classification models on new, unseen data to assess their ability to generalize beyond the training examples. In this study, the dataset is strategically divided into two parts using a 70:30 ratio, where 70% of the data is designated for training the model, and the remaining 30% is reserved for testing its predictive capabilities. The training set, which constitutes the larger part of the data, is utilized not only for training the machine learning model but also for optimizing the hyperparameters, i.e., specific configurations that have a significant impact on the model's performance. Conversely, the testing set, which represents the smaller segment of the dataset, plays a critical role in evaluating the model's ability to generalize. This separation of data ensures that the model is exposed to a diverse range of instances during training, while the independent testing subset is used exclusively for rigorous assessment on new data. This methodological division is crucial in preventing any overlap that could lead to biased performance metrics, thereby validating the reliability of the model's predictions. The training phase for each kernel function involves a systematic exploration of various combinations of hyperparameters, as illustrated in Table 3. By employing the grid search technique, we can precisely evaluate each potential combination and identify the set of hyperparameters that yield optimal performance. This process is vital, as the selected parameters directly influence the model's ability to learn patterns from the training data. Once the optimal hyperparameters have been determined through the grid search, they are applied to the model when making predictions on the unseen data. This step is essential not only for assessing the model's accuracy but also for verifying its effectiveness in real-world applications. The overarching goal is to ensure that the model not only performs well on the training data but also maintains a high degree of predictive accuracy when encountering new instances, thereby demonstrating its robustness and practical utility.

Tab. 3 *Optimal hyperparameters of ensemble models using grid search*

Models	Hyperparameters	Optimal Values
RF	n_estimators = [1-200]	14
	max_depth = [1-10]	9
XGBoost	learning_rate = [0.001, 0.01,0.1]	0.1
	n_estimators = [1-200]	151
	depth = [1-10]	9
LightGBM	n_estimators = [1-200]	186
	learning_rate = [0.001, 0.01,0.1]	0.1

4 RESULTS AND DISCUSSION

The study employed three state-of-the-art ensemble models, i.e., Random Forest (RF), XGBoost, and LightGBM, to evaluate their effectiveness in predicting slope stability. These models are known for their ability to handle complex datasets and provide a high degree of predictive accuracy by combining multiple decision trees into a cohesive framework. To enhance the performance of these ensemble models, the Binary Particle Swarm Optimization (BPSO) algorithm is applied to identify the most relevant features for each model. This step ensures that only the most impactful features contribute to the prediction process, thereby reducing the computational complexity and improving the model's interpretability (Table 1). After the selection of the correct feature, grid search optimization is performed to fine-tune the hyperparameters of each ensemble model. This optimization ensures that the models are configured for maximum performance by systematically exploring combinations of hyperparameters such as tree depth, learning rate, and the number of estimators (Table 3). The performance of the classifiers is assessed using multiple metrics, including the Area Under the Curve (AUC), accuracy, and kappa scores. These metrics provide a comprehensive evaluation of the model's ability to distinguish between stable and unstable slopes. The results indicate the following AUC scores: RF achieved 0.809, XGBoost scored 0.794, and LightGBM outperformed the others with an AUC of 0.871 (Fig. 4). This demonstrates that LightGBM has superior discriminatory power in this specific application.

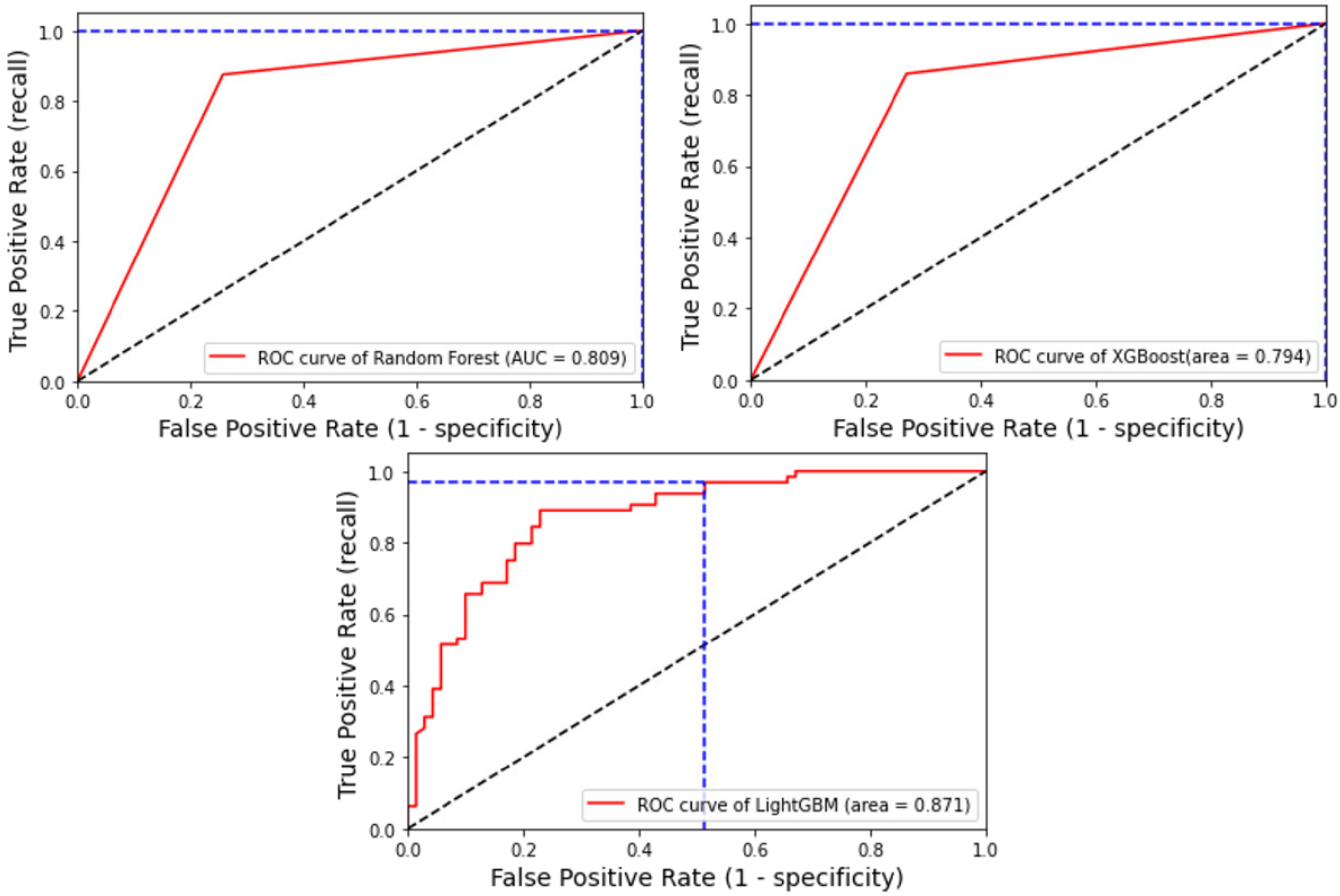


Fig. 4 *ROC curves of ensemble models*

Similarly, the accuracy scores reflect the overall proportion of correctly predicted instances, with RF scoring 0.838, XGBoost achieving 0.822, and LightGBM securing 0.832. While RF has a slight edge in accuracy, the differences are relatively small and showcase the competitiveness of all three models. The kappa scores, which measure agreement beyond coincidence, are 0.613 for both RF and XGBoost and 0.658 for LightGBM (Fig. 5).

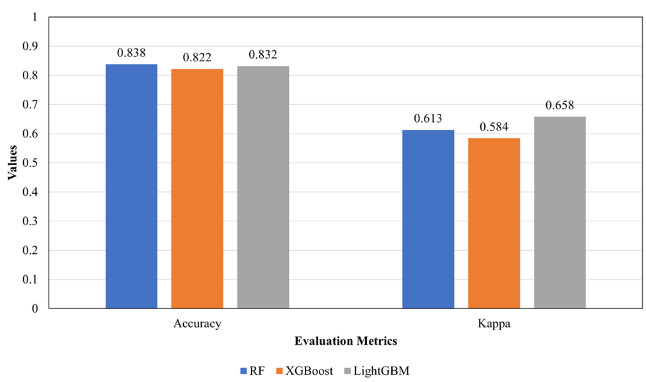


Fig. 5 *Evaluation metrics of ensemble models*

To validate these results, a confusion matrix analysis was conducted, which provides detailed insights into the performance of the classifications by reporting the number of misclassifications. The results of the confusion matrix analysis revealed that RF had 26 misclassifications and XGBoost had 28, while LightGBM had the fewest, with 23 misclassifications (Fig. 6). These findings align with the other metrics and confirm LightGBM is superior in performance compared to the other ensemble models in this study.

Ensemble Models	Actual values	Predicted Values 0	Predicted Values 1
RF	0	52	18
	1	8	56
XGBoost	0	51	19
	1	9	55
LightGBM	0	54	16
	1	7	57

Fig. 6 *Confusion matrix*

The sensitivity of all the models was further evaluated to assess their performance in accurately identifying the true positive class while limiting any false negatives. Sensitivity is a crucial metric in classification tasks, particularly in the context of slope stability analysis, as it plays a significant role in ensuring that high-risk zones or unstable slopes are accurately classified. This minimizes the potential for overlooking hazardous areas. The results of the analysis indicate the following sensitivity scores for the models: Random Forest

(RF) achieved a sensitivity of 0.875, XGBoost scored 0.859, and LightGBM recorded the highest degree of sensitivity at 0.89 (Fig. 7). These values were derived using Eq. 1, which defines sensitivity as the ratio of true positives (TP) to the sum of true positives (TP) and false negatives (FN).

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{1}$$

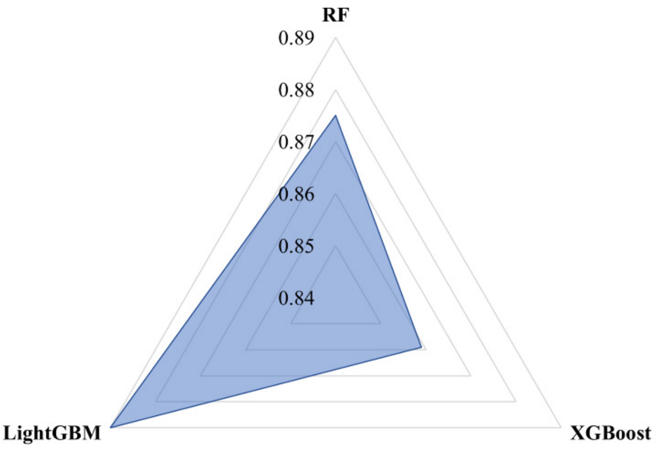


Fig. 7 *Sensitivity of ensemble models*

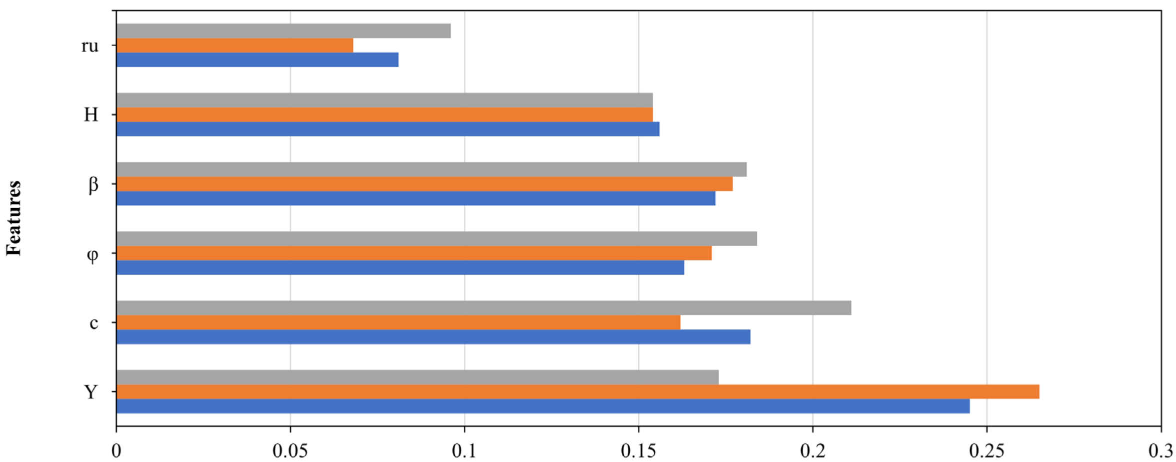
The marginally higher sensitivity value associated with LightGBM underscores its enhanced capability for accurately identifying positive cases, such as unstable slopes, in comparison to Random Forest (RF) and XGBoost. This performance differential demonstrates that LightGBM is particularly effective in capturing critical relationships between the features input and the target variable, which is vital for minimizing risks related to false negatives.

5 ANALYSIS OF THE IMPORTANCE OF FEATURE

The importance of feature represents a fundamental concept in the development of a machine learning model; it serves as a critical mechanism for understanding the contribution of the individual variables input to the predictive outcomes. This analytical approach enables researchers to systematically evaluate the relative significance of each feature within the decision-making framework of a given model. By quantifying the degree to which specific features affect model predictions, the importance of feature analysis provides valuable insights into the underlying relationships between the variables input and the target outcomes; it thereby enhances the model's interpretability and supports evidence-based feature selection processes in machine learning applications.

The feature importance resulting from LightGBM, XGBoost, and Random Forest (RF) are presented in Fig.8. Across all the models, the parameter Y demonstrates the highest degree of influence, with XGBoost (0.265) and RF (0.245) assigning substantially greater importance compared to LightGBM (0.173). This confirms Y as the dominant predictor across methods. Feature c is identified as the second most influential parameter in LightGBM (0.211), while RF (0.182) and XGBoost (0.162) assign it slightly lower but still considerable importance. Similarly, φ and β consistently rank as moderately important, with values ranging between 0.163 – 0.184 and 0.172 – 0.181, respectively. Feature H shows a consistent but lower degree of importance (0.154 – 0.156) across all three algorithms. Finally, ru is consistently the least significant feature, with its importance values ranging from 0.068 (XGBoost) to 0.096 (LightGBM).

Overall, the results highlight Y as the most critical variable, followed by the strength parameters (c, φ) and slope geometry (β, H), while the pore pressure ratio (ru) contributes minimally.



	Y	c	φ	β	H	ru
LightGBM	0.173	0.211	0.184	0.181	0.154	0.096
XGBoost	0.265	0.162	0.171	0.177	0.154	0.068
RF	0.245	0.182	0.163	0.172	0.156	0.081

Values

LightGBM XGBoost RF

Fig. 8 *Importance of feature analysis across the ensemble models*

The consistency across the models strengthens the robustness of these findings, although the variations in weighting reflect the algorithm-specific sensitivities.

6 SENSITIVITY ANALYSIS

The mitigation of slope failure necessitates a comprehensive assessment of feature sensitivity to identify the critical parameters that contribute to slope instability mechanisms. This evaluation of feature sensitivity is paramount for an accurate slope stability analysis and the development of optimal engineering interventions. The present study quantifies feature sensitivity across multiple classification models by analyzing their respective contributions to predictive performance under optimized hyperparameter configurations. This systematic sensitivity analysis provides crucial insights into the most influential geotechnical and environmental factors governing slope stability conditions, thereby facilitating evidence-based decision-making processes and enabling the implementation of proactive risk mitigation strategies to minimize potential slope failure. The methodology establishes a robust framework for understanding the relative importance of contributing factors in assessing slope stability, which is essential for advancing predictive modeling capabilities in geotechnical engineering applications.

Fig.9 illustrates the sensitivity of the six parameters input (Y, c, φ, β, H, ru), which were obtained using the three machine learning models, LightGBM, XGBoost, and Random Forest (RF). Among the parameters, Y consistently exhibits the highest sensitivity, with XGBoost (0.265) and RF (0.245) assigning notably greater influence compared to LightGBM (0.173). This confirms that Y is the most sensitive parameter and plays a dominant role in governing the model outcomes. The parameter c also shows a high degree of sensitivity, particularly in LightGBM (0.211), whereas XGBoost (0.162) and RF (0.182) reflect comparatively lower values. Both φ (0.163 – 0.184) and β (0.172 – 0.181) demonstrate a moderate degree of sensitivity, which highlights their stable but secondary contribution across the models. The parameter H exhibits a relatively lower degree of sensitivity (0.154 – 0.156), although it remains consistent across all the algorithms. In contrast, ru is the least sensitive factor, with values ranging from 0.068 to 0.096, suggesting only a marginal effect on the model predictions.

Overall, the sensitivity analysis indicates that the output is highly responsive to Y and moderately sensitive to c, φ, and β, while H and especially ru exert a weaker degree of influence. The convergence of the results across the three algorithms reinforces the robustness of these findings, although minor differences in the relative weightings highlight algorithm-specific sensitivities.

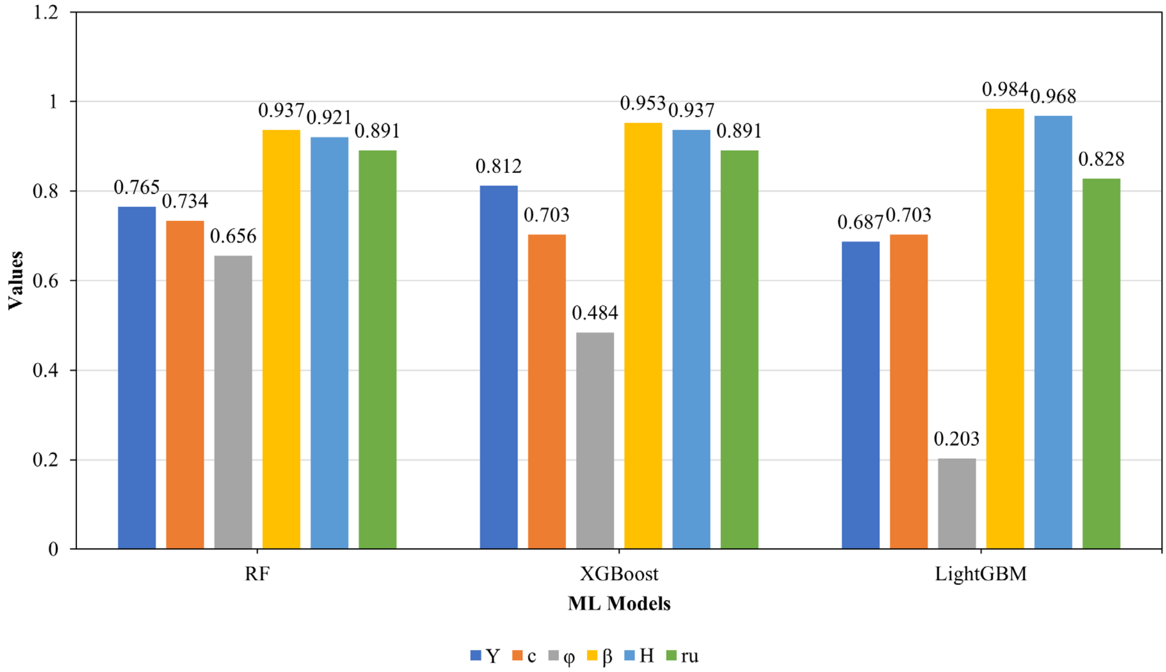


Fig. 9 *Sensitivity analysis of features across the ensemble models*

7 CONCLUSION

Overall The study utilized three ensemble models, namely, RF, XGBoost and LightGBM, which were integrated with BPSO for predicting slope stability. The findings highlight the importance of feature selection and hyperparameter optimization in enhancing a model's performance. The use of BPSO for feature selection not only streamlined the input data but also contributed to the improved efficacy of the models. Based on the study, the following conclusions can be drawn:

1. The BPSO algorithm employed across the three ensemble models resulted in the selection of distinct feature sets for each model. In the case of the Random Forest (RF) model,

out of six features, five were identified for inclusion: H, ϒ, c, β, and ϕ. For the XGBoost model, five features were similarly selected: ru, H, ϒ, c, and ϕ. Finally, the LightGBM model also identified the features ru, H, ϒ, c, and ϕ for its analysis.

2. The analysis of the ROC curve has provided valuable insights into the performance of the various models evaluated. The RF model achieved an auc of 0.809, which indicated a commendable level of classification ability. The XGBoost model demonstrated a slightly lower performance with an auc of 0.794, which still reflects a reasonable capacity for discriminating between the positive and negative classes. In contrast, the LightGBM model surpassed the others, attaining an auc of 0.871, which underscores its superior capability in accurately distinguishing between the classes within the dataset. These findings suggest that LightGBM is the most effective model among those assessed. LightGBM emerged as the most effective model in terms of AUC and kappa scores, which demonstrated its ability to handle the complex relationships among the features and provide more accurate predictions.

3. LightGBM achieved the highest accuracy rate of 0.832 and a kappa value of 0.658. These metrics underscore its effectiveness in addressing class imbalances and delivering reliable predictions when compared to RF and XGBoost. Furthermore, LightGBM recorded the fewest misclassifications, i.e., only 23 instances, thereby highlighting its precision in accurately classifying both the stable and unstable slopes.

4. The findings of this study have important implications for slope stability analyses. The impressive performance of LightGBM indicates its potential as an effective tool for slope stability predictions by facilitating early warning systems and risk assessments. Furthermore, the incorporation of advanced feature selection and optimization techniques ensures that the model is both efficient and interpretable, which enhances its value for practical applications in geotechnical engineering and disaster management.

5. The feature importance analysis using LightGBM, XGBoost, and Random Forest (RF) consistently demonstrated that Y is the most influential predictor; the analysis confirmed its dominant role in determining the model outputs. Features such as c, φ, and β were also shown to be highly important and reflected their strong geotechnical significance in stability assessments. In contrast, H was found to only had a moderate contribution, while ru consistently ranked lowest across all the algorithms, which suggested only a marginal influence on predictions. These results indicate that a model's performance is primarily governed by strength and geometric parameters, with the pore pressure ratio exerting only a limited direct effect.

6. The sensitivity analysis further reinforced these findings; it showed that the models are highly sensitive to variations in Y, moderately sensitive to c, φ, and β, and only weakly sensitive to H and ru. The convergence of the results across the three independent machine learning algorithms strengthens confidence in the robustness of the analysis. At the same time, the minor variations in the weightings highlight algorithm-specific sensitivities; this suggests that ensemble approaches or cross-model validations may provide more reliable insights. Collectively, the results emphasize the need to prioritize accurate estimations of dominant parameters (Y, c, φ, β) in practical applications, while recognizing the relatively lesser role of ru and H.

8 LIMITATIONS

Future research may enhance these findings by investigating additional ensemble techniques; it could incorporate real-time data for dynamic slope stability analyses and expand the study to encompass diverse geographical regions to assess the models' generalizability. Moreover, the integration of explainable AI methodologies could improve the interpretability of predictions, which could thereby assist stakeholders in making informed decisions regarding the management of slope stability.

REFERENCES

Abramson, L. W., Lee, T. S., Sharma, S., & Boyce, G. M. (2001). *Slope stability and stabilization methods. John Wiley & Sons.*

Ahangari Nanehkaran, Y., Pusatli, T., Chengyong, J., Chen, J., Cemiloglu, A., Azarafza, M., & Derakhshani, R. (2022). *Application of machine learning techniques for the estimation of the safety factor in slope stability analysis.* Water, 14(22), 3743.

Alshboul, O., Almasabha, G., Shehadeh, A., & Al-Shboul, K. (2024). *A comparative study of LightGBM, XGBoost, and GEP models in shear strength management of SFRC-SBWS.* In Structures (Vol. 61, p. 106009). Elsevier.

Belgiu, M., & Drăguţ, L. (2016). *Random forest in remote sensing: A review of applications and future directions.* ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31.

Biau, G., & Scornet, E. (2016). *A random forest guided tour.* Test, 25, 197-227.

Cheng, Y. M., Lansivaara, T., & Wei, W. B. (2007). *Two-dimensional slope stability analysis by limit equilibrium and strength reduction methods.* Computers and geotechnics, 34(3), 137-150.

Das, S. K., Biswal, R. K., Sivakugan, N., & Das, B. (2011). *Classification of slopes and prediction of factor of safety using differential evolution neural networks***.** Environmental Earth Sciences, 64, 201-210.

Dawson, E. M., Roth, W. H., & Drescher, A. (1999). *Slope stability analysis by strength reduction.* Geotechnique, 49(6), 835-840.

Duncan, J. M. (1996). *State of the art: limit equilibrium and finite-element analysis of slopes.* Journal of Geotechnical Engineering, 122(7), 577-596.

Duncan, J. M., Wright, S. G., & Brandon, T. L. (2014). *Soil strength and slope stability.* John Wiley & Sons.

Griffiths, D. V., & Lane, P. A. (1999). *Slope stability analysis by finite elements.* Geotechnique, 49(3), 387-403.

Huang, Y. H. (2014). *Slope stability analysis by the limit equilibrium method: Fundamentals and methods.* American Society of Civil Engineers.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye,Q., & Liu, T. Y. (2017). *Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems, 30.*

Khajehzadeh, M., Taha, M. R., El-Shafie, A., & Eslami, M. (2012). A *modified gravitational search algorithm for slope stability analysis. Engineering Applications of Artificial Intelligence,* 25(8), 1589-1597.

Khanesar, M. A., Teshnehlab, M., & Shoorehdeli, M. A. (2007). *A novel binary particle swarm optimization.* In 2007 Mediterranean conference on control & automation (pp. 1-6). IEEE.

Koopialipoor, M., Jahed Armaghani, D., Hedayat, A., Marto, A., & Gordan, B. (2019). *Applying various hybrid intelligent systems to evaluate and predict slope stability under static and dynamic conditions.* Soft Computing, 23, 5913-5929.

Li, X. (2007). *Finite element analysis of slope stability using a nonlinear failure criterion.* Computers and Geotechnics, 34(3), 127-136.

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.

Lin, S., Zheng, H., Han, C., Han, B., & Li, W. (2021). *Evaluation and prediction of slope stability using machine learning approaches.* Frontiers of Structural and Civil Engineering, 15(4), 821-833.

Liu, S. Y., Shao, L. T., & Li, H. J. (2015). *Slope stability analysis using the limit equilibrium method and two finite element methods.* Computers and Geotechnics, 63, 291-298.

Luo, Z., Bui, X. N., Nguyen, H., & Moayedi, H. (2021). *A novel artificial intelligence technique for analyzing slope stability using PSO-CA model.* Engineering with Computers, 37, 533-544.

Matsui, T., & San, K. C. (1992). *Finite element slope stability analysis by shear strength reduction technique.* Soils and foundations, 32(1), 59-70.

Nanehkaran, Y. A., Licai, Z., Chengyong, J., Chen, J., Anwar, S., Azarafza, M., & Derakhshani, R. (2023). *Comparative analysis for slope stability by using machine learning methods.* Applied Sciences, 13(3), 1555.

Natekin, A., & Knoll, A. (2013). *Gradient boosting machines, a tutorial.* Frontiers in neurorobotics, 7, 21.

Pham, B. T., Qi, C., Ho, L. S., Nguyen-Thoi, T., Al-Ansari, N., Nguyen, M. D., Lye, H.V., Le, H. V., & Prakash, I. (2020). *A novel hybrid soft computing model using random forest and particle swarm optimization for estimation of undrained shear strength of soil.* Sustainability, 12(6), 2218.

Puri, N., Prasad, H. D., & Jain, A. (2018). *Prediction of geotechnical parameters using machine learning techniques.* Procedia Computer Science, 125, 509-517.

Qi, C., & Tang, X. (2018). *A hybrid ensemble method for improved prediction of slope stability.* International Journal for Numerical and Analytical Methods in Geomechanics, 42(15), 1823-1839.

Rukhaiyar, S., Alam, M. N., & Samadhiya, N. K. (2018). *A PSO-ANN hybrid model for predicting factor of safety of slope.* International Journal of Geotechnical Engineering, 12(6), 556-566.

Sakellariou, M. G., & Ferentinou, M. D. (2005). *A study of slope stability prediction using neural networks.* Geotechnical & Geological Engineering, 23, 419-445.

Saurabh Kumar Anuragi, D.Kishan, S.K.Saritha. (2024). *Comparative Analysis of Ensemble Learning Approaches for Slope Stability Prediction, SSRG International Journal of Civil Engineering, 11 (5), pp. 168-180.* https://doi.org/10.14445/23488352/IJCE-V11I5P116.

Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). *A comparison of random forest variable selection methods for classification prediction modeling.* Expert systems with applications, 134, 93-101.

Sun, X., Liu, M., & Sima, Z. (2020). *A novel cryptocurrency price trend forecasting model based on LightGBM.* Finance Research Letters, 32, 101084.

Xue, X., Yang, X., & Chen, X. (2014). *Application of a support vector machine for prediction of slope stability.* Science China Technological Sciences, 57, 2379-2386.

REFERENCES

Zheng, F., Leung, Y. F., Zhu, J. B., & Jiao, Y. Y. (2019). *Modified predictor-corrector solution approach for efficient discontinuous deformation analysis of jointed rock masses.* International Journal for Numerical and Analytical Methods in Geomechanics, 43(2), 599-624.

Zheng, H., Liu, D. F., & Li, C. G. (2005). *Slope stability analysis based on elasto-plastic finite element method.* International Journal for Numerical Methods in Engineering, 64(14), 1871–1888. https://doi.org/10.1002/nme.1406

Zhuang, X., Zheng, F., Zheng, H., Jiao, Y. Y., Rabczuk, T., & Wriggers, P. (2021). *A cover-based contact detection approach for irregular convex polygons in discontinuous deformation analysis.* International Journal for Numerical and Analytical Methods in Geomechanics, 45(2), 208-233.