

Review

Machine Learning in the Stochastic Analysis of Slope Stability: A State-of-the-Art Review

Haoding Xu , Xuzhen He * , Feng Shan, Gang Niu and Daichao Sheng

School of Civil and Environmental Engineering, University of Technology Sydney, Ultimo, NSW 2007, Australia; haoding.xu@student.uts.edu.au (H.X.); feng.shan@student.uts.edu.au (F.S.); gang.niu@student.uts.edu.au (G.N.); daichao.sheng@uts.edu.au (D.S.)

* Correspondence: xuzhen.he@uts.edu.au

Abstract: In traditional slope stability analysis, it is assumed that some “average” or appropriately “conservative” properties operate over the entire region of interest. This kind of deterministic conservative analysis often results in higher costs, and thus, a stochastic analysis considering uncertainty and spatial variability was developed to reduce costs. In the past few decades, machine learning has been greatly developed and extensively used in stochastic slope stability analysis, particularly used as surrogate models to improve computational efficiency. To better summarize the current application of machine learning and future research, this paper reviews 159 studies of supervised learning published in the past 20 years. The achievements of machine learning methods are summarized from two aspects—safety factor prediction and slope stability classification. Four potential research challenges and suggestions are also given.

Keywords: slope stability; factor of safety; slope stability classification; machine learning; geotechnical engineering; uncertainty; reliability analysis



Citation: Xu, H.; He, X.; Shan, F.; Niu, G.; Sheng, D. Machine Learning in the Stochastic Analysis of Slope Stability: A State-of-the-Art Review. *Modelling* **2023**, *4*, 426–453. <https://doi.org/10.3390/modelling4040025>

Academic Editor: Luca Lenti and Salvatore Martino

Received: 14 August 2023

Revised: 26 September 2023

Accepted: 29 September 2023

Published: 1 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Landslides are sudden and serious disasters that can cause significant damage to nearby facilities, resulting in economic and casualty losses. Therefore, the evaluation of slope stability is a critical prerequisite for disaster prevention and mitigation. In numerical simulations, slope stability is presented as a factor of safety (*FOS*), which is obtained using deterministic analyses such as limit analysis [1], finite element limit analysis [2], displacement-based finite element analysis combined with the strength reduction method [3,4], or finite element analysis with the gravity increasing method [5,6]. In site investigation and analysis, slope stability is typically classified using empirical formulas or expert judgments.

Slope stability analysis concerns earth materials such as soil, rock, and other materials. Soil is one of the engineering materials that has the most complex physical, mechanical, and chemical behaviors and is made of three phases. Due to geological action and stress history, soils exhibit complex spatial variability and anisotropy, which makes studying soils and predicting their behavior difficult. In geotechnical engineering, two main approaches have traditionally been used to study the mechanical behavior of geotechnical materials: (1) empirical methods, such as laboratory tests and site investigations, and (2) numerical and analytical methods. The cost of experiments and field tests tends to rise with the number of tests conducted, which directly impacts the precision of these assessments. Consequently, significant engineering expertise is necessary to gauge slope stability based on a limited set of experiments. Traditional analytical techniques often make simplifications, such as assuming uniform soil conditions, and thereby overlook the inherent spatial variability in real-world soil [7].

The stability of a slope depends on the material strength and geometric and hydraulic parameters associated with it [8–10]. Some studies use the deterministic inversion method

to estimate soil parameters and study slope stability, and the key to this method is to find a set of parameters that will lead to slope failure [11,12]. However, considering uncertainty in soil parameters, the parameters deduced using the deterministic method may not necessarily represent the real situation, and there may be multiple critical combinations of these parameters [9]. Therefore, it is necessary to consider the use of probabilistic methods for inversion (probabilistic back analysis).

With the development of statistical methods, geotechnical academics started to consider the use of random field theory to represent the spatial variability in soil [13] and quantify the safety margin of slopes using failure probability or the equivalent reliability index from a probabilistic perspective. The brute-force Monte Carlo simulation (MCS) method has gained popularity in reliability analysis due to its simplicity, flexibility, and ease of use. Theoretically, the Monte Carlo method can handle all problems. For example, Monte Carlo-based probabilistic back analysis can be used to perform parameter inversion, risk assessments, and sensitivity analyses [9,14]. However, when the sample size is large (i.e., for small probability events or when the probability density function (PDF) is needed), the brute-force Monte Carlo approach is time-consuming and computationally intensive [15–17]. The response surface method was then proposed and widely used [18–21]. It approximates the limit state function with a polynomial expression using the function value at a specific point. This type of analytic function replaces the exact limit state function in Monte Carlo simulations. As a result, the number of calculations required to assess the reliability of structural systems can be significantly reduced. However, RSM requires a prior specification of a suitable fitting function (usually a polynomial function) to be specified in advance [22]. Since real-world problems are often very complex, polynomial estimates may not perform well in providing a good representation of the objective function [23].

On the other hand, technological developments have made site investigation data more readily available to engineers. It is necessary to take into account these actual measured parameters in the estimation of parameters. When prior knowledge about parameters such as geometry, strength, and hydraulics is obtained, Bayesian inference can combine this prior knowledge to obtain the posterior probability distribution of parameters, which can be used for risk assessment, such as calculating the probability of slope instability [24]. The Bayesian method can also be used for time prediction of slope stability by combining historical data and new observational data, thereby reducing the uncertainty in parameters and predicting slope stability at a certain point in the future [25]. However, the high non-linearity of the model and the exponentially increased calculation and sample number prevent rapid *FOS* prediction and accurate classification [24].

In recent years, the machine learning (ML) method has gained increasing attention from geotechnical researchers as a promising approach for studying slope stability. ML-aided stochastic analyses have been successfully applied in many cases [26–32]. Research using machine learning to aid in slope stability assessment has been conducted in various countries. These studies often focus on predicting safety factors or categorizing slope safety based on data gathered from field tests. A large number of slope classifications have been conducted using field data collected globally. For instance, the data in a study by Zhang originate from Yunyang County, Chongqing, China [10], while the data in a study by Zhu are sourced from the South Pars Special Zone, Assalouyeh, Southwest Iran [32]. ML models in geotechnical reliability analysis can capture complex relationships between input and output data and construct high-dimensional nonlinear functions to directly predict the target output. Machine learning-generated surrogate models work well with variability. They can take into account the variability in data, that is, the parameter variability in soil and slopes, the variability in the model, such as different initialization values or model structures or hyperparameters, and the variability in evaluation and verification, that is, different evaluation indicators. For slope stability problems, the inputs are usually slope boundary parameters and soil mechanical parameters, and the output is the *FOS* or slope stability classification. Compared with traditional probabilistic or analytical methods (such as the brute-force Monte Carlo simulations or polynomial fitting), ML methods have several

advantages, such as high accuracy, not requiring any initial assumptions between input and output (the relationship between input and output is a high-order nonlinear function), being able to handle large datasets, and being capable of handling incomplete data [33]. By using ML techniques, geotechnical engineers can avoid the limitations of traditional methods, such as the need for prior assumptions or high computational requirements, and can accurately predict slope stability.

This paper examines 159 publications in the past 20 years that have used supervised ML methods in slope stability problems. The stability analysis method is briefly introduced in Section 2. Section 3 provides an overview of several commonly used ML algorithms. Section 4 outlines the different applications of ML in slope stability. Additionally, Section 5 offers some future research perspectives, and the main conclusion from this review is summarized in Section 6.

2. Brief Overview of Methods for Slope Stability Analysis

This section provides an overview of methods for slope stability analysis. These methods can be broadly categorized into four groups: empirical methods, the limit equilibrium method (LEM), the finite element method (FEM), and other numerical methods. Figure 1 illustrates the distribution of different methods within the papers examined in this review. The LEM is used in approximately 50% of all studies, which is much more than other methods and largely due to its early development, ease of use, and high computing speed. The LEM is followed by the FEM and empirical methods. Due to its powerful calculation ability and wide applicability, the FEM has gradually become the most popular numerical method. Compared with the LEM, the FEM not only considers the constitutive model of soil but can also represent stress and deformation outside the sliding surface. However, numerical methods are simplified solutions to practical problems and are sometimes unable to explain complex practical situations. Therefore, empirical methods are essential for slope stability problems as they can directly obtain the *FOS* or slope stability classification, which can be used as independent and dependent variables in reliability analysis. In addition, some other numerical methods are also used in slope stability analysis, such as the finite difference method, the limit analysis method, and mesh-free methods such as the smoothed-particle hydrodynamics (SPH) method [34–37]. The LEM and the FEM will be briefly introduced in the following subsection.

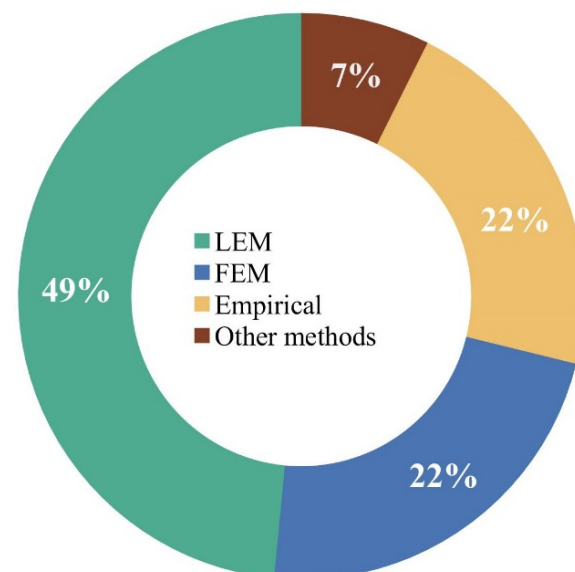


Figure 1. Proportion of different methods in slope stability (LEM: limit equilibrium method; FEM: finite element method).

2.1. The Limit Equilibrium Method (LEM)

A failure surface is required to be assumed in the LEM when solving a problem, which is a sliding line in the 2D case, usually a plane, circular, or logarithmic spiral. Based on this assumption, the stability problem is transformed into finding the most dangerous positions for the failure surface. In addition, assumptions are made about the stress distribution on the failure surface to obtain the overall equilibrium equation, which can then be solved using simple statics. Since the LEM assumes static conditions for the failure of slopes, constitutive behaviors and the related deformation are ignored in the calculations [38]. The *FOS* is defined as the ratio of soil resisting shear strength (τ_f) and the shear stress on the assumed failure surface (τ) [39]. If the plastic behavior of the soil follows the Mohr-coulomb failure criterion, the *FOS* is:

$$FOS = \frac{\tau_f}{\tau} = \frac{\sigma' \tan \phi' + c'}{\tau} \quad (1)$$

where σ' is the effective stress, ϕ' is the effective friction angle, and c' is the effective cohesion. The minimum *FOS* on the assumed failure surface is obtained using iterative calculations [40].

The direction, which is perpendicular to the cross-section, is usually considered insignificant for the calculation of *FOS* [6]. The commonly used two-dimensional LEMs are the Bishop simplified method and the ordinary method of slices (Swedish circle method, Petterson, Fellenius, Spencer, Sarma, etc.). These methods primarily differ in the inter-stripe force assumptions and whether all three equilibrium conditions are satisfied (e.g., force equilibrium in the horizontal and vertical direction and moment equilibrium condition). The Spencer and Sarma methods satisfy all requirements and therefore are called rigorous methods, which give more accurate calculations than other non-rigorous methods [41,42].

2.2. The Finite Element Method (FEM)

The FEM is a representative of the mesh-based method and is probably the most widely used numerical method [43–45] in geotechnical problems. Compared with LEMs, FEMs can consider the constitutive behavior of soil and are not required to assume the specific failure surface. Two methods were presented for slope stability analysis combined with FEM: the strength reduction method (SRM) and the gravity increasing method (GIM) [5,6].

In the SRM, the parameters of the soil are reduced using a reduction factor until the failure happens, and the *FOS* is calculated as the reciprocal of this factor using iterations. Taking the Mohr–Coulomb model as an example, in the i th iteration, the reduced strength parameters are:

$$c_i = \frac{c}{FOS_i} \quad (2)$$

$$\tan \phi_i = \frac{\tan \phi}{FOS_i} \quad (3)$$

The SRM and the LEM usually give very similar results for homogenous slopes [46]. However, the SRM is sometimes sensitive to nonlinear algorithms and flow rules. In addition, the SRM cannot determine failure surfaces, which may be only slightly less critical than the SRM failure surface [47].

In the GIM, the calculated gravity increases gradually until the slope becomes unstable. Therefore, the GIM aims to obtain the limit of gravity, which is represented by the acceleration of gravity [48]. The *FOS* is defined as the ratio between the failure gravity (g_f) and the real-world gravity (g_i):

$$FOS = \frac{g_f}{g_i} \quad (4)$$

3. Brief Overview of Machine Learning (ML) Methods

The applications of ML methods in slope stability analysis in the last two decades (2003–2022) are shown in Figure 2. The bars represent the annual number of studies, and the black dashed line is the trendline. The statistical data were obtained from the Web of Science database. Using ‘machine learning’ and ‘slope stability analyses’ as search keywords, and after checking the relevance between papers and the topic, a total of 159 publications from 2002 to 2022 were obtained. The number of publications has grown dramatically from 2019 to 2022. This shows that in recent years, researchers have paid more attention to the use of ML in slope stability problems. It is important to note that a single publication can use multiple ML methods, resulting in a larger number of methods used than the number of publications reporting those studies.

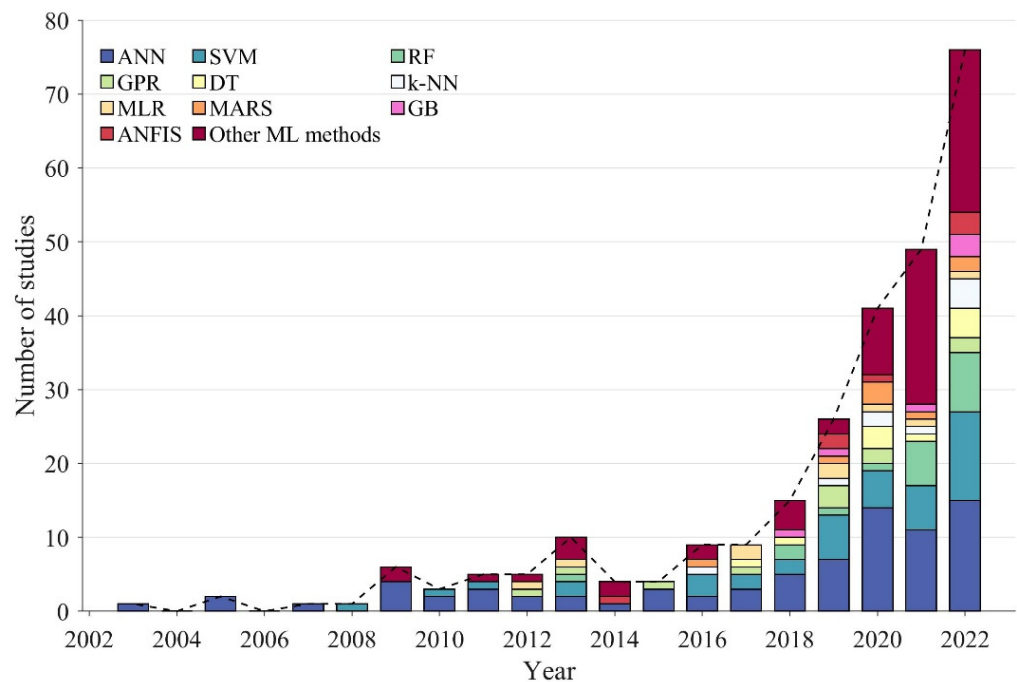


Figure 2. Annual number of studies about slope stability using ML, where the dashed line is the trendline.

3.1. Artificial Neural Networks (ANNs), Adaptive Neuro-Fuzzy Inference Systems (ANFISs), and Deep Neural Networks (DNNs)

Artificial neural networks (ANNs) are a commonly used ML method to study and approximate highly non-linear relationships between inputs and outputs without any assumptions or mathematical functions [49]. A large number of artificial neurons make up an ANN, each of which can receive and process ‘signals’ and transmit them to adjacent neurons. ‘Signals’ are real numbers transmitted between neurons with a specific activation function. Neurons adjust their weights with network learning and training. The ANN structure contains an input layer, hidden layers, and an output layer [31]. The different structures and connections among neurons determine the form of an ANN. The two ANN structures used in most of the examined studies include multilayer perceptron [50–52] and extreme learning machine [53,54]. ANNs have been widely used in the field of machine learning and can perform a wide variety of tasks, such as classification [32] and regression tasks [31]. Multilayer feedforward networks are typical universal approximators [55]. ANNs can also handle time series forecasting [56,57] and collaborative information filters [58]. In this review, the term ANN primarily denotes the conventional neural network architectures encompassing the most basic structure. These networks are characterized as having only fully connected layers (also known as ‘Dense layers’) as their hidden layers.

They operate without loops, lack inherent memory, and consider only the current input for any given operation.

Adaptive neuro-fuzzy inference systems (ANFISs) are a machine learning method that combines ANNs with Takagi–Sugeno fuzzy inference systems [59–61]. The key feature is a set of fuzzy rules, which are conditional statements of the form ‘if x is A then y is B ’, used to approximate the input–output relationship of a system. An ANFIS trains a model with gradient descent and backpropagation algorithms with roughly five steps [59]:

- (1) The input data are transformed into fuzzy sets using membership functions.
- (2) The firing strength of each rule is generated from the input and rules.
- (3) Normalized firing strengths are calculated using weighted averaging.
- (4) Consequent parameters are adjusted to optimize the parameters and weights.
- (5) All incoming signals are summed up to obtain the overall output.

The primary advantage of ANFIS is the ability to represent nonlinearity and structured knowledge [59]. However, ANFIS requires a large dataset to train the model, and selecting appropriate input–output data is much more crucial. Moreover, the number of fuzzy rules increases exponentially with the amount of input data, which greatly increases the computation cost and may affect the performance of a model [62].

Deep neural networks (DNNs) are extended from the concept of ANNs, and DNNs are probably the most successful deep learning method. The term ‘deep’ can refer to a structure of more than one hidden layer in NNs. A convolutional neural network (CNN) is a commonly used deep learning model, which is widely used in image and pattern recognition [63] in the computer field, and has been extended to other fields, such as foundation bearing capacity problems [15,64] and slope stability problems [65–67] in geotechnical engineering. Hidden layers in CNNs usually contain convolutional layers, pooling layers, and fully connected layers. CNNs effectively capture the topology of images and build features that automatically link features to image classification. Also, there are some variants of CNNs, such as locally connected networks [15,66]. Classic DNN structures also include recurrent neural networks (RNNs), transformers, etc. RNNs can be used to process language models [68] and speech recognition [69]. Long short-term memory (LSTM) is a popular variant of RNNs. The recently very popular ChatGPT from OpenAI is based on the transformer architecture [70]. DNNs usually have many layers and millions of parameters, which can help models fit the training data closely. However, overcomplicated model structures can lead to overfitting problems (the generalization of the trained model is poor). This problem is usually avoided using regularization or trimming connectivity (dropout, etc.). In this review, the term DNN pertains to more complex neural architectures. These can comprise hidden layers of varying structures, such as convolutional or pooling layers, and they may possess functionalities beyond those of basic ANNs. It is crucial to note that not all hidden layers in DNNs are fully connected.

3.2. Support Vector Machine (SVM)

Support vector machine (SVM) was initially developed to finish the binary classification task [71]. This method is now developed as a supervised learning method and can deal with classification and regression problems [31]. The key to SVM lies in the concept of identifying an optimal hyperplane. In classification tasks for linearly separable data, SVM maps training samples into an n -dimensional space and seeks to maximize the margin between the two categories. For regression tasks, SVM aims to find the most fitting hyperplane that contains the most data points. The ϵ -insensitive loss function is introduced to consider an acceptable error range ϵ . This loss function defines a region of width of 2ϵ , where the difference between the predicted and true values is ignored. Specifically, if the difference between the predicted and actual values is less than ϵ , the loss is 0; otherwise, the loss is the difference between the difference and ϵ . Quadratic programming methods are used to solve hyperplanes both in classification and regression problems. When handling non-linear data, SVM uses the kernel trick to map data into a higher-dimensional space, facilitating the determination of an optimal hyperplane in this newly transformed feature domain [31].

Popular kernel methods include the radial basis function, periodic kernel, polynomial kernel, exponential squares kernel, etc. SVM has been widely used in many fields, such as text and hypertext classification [72], bioinformatics [73], and anomaly detection [74].

3.3. Gaussian Process Regression (GPR)

Gaussian process regression (GPR) is a non-parametric supervised learning Bayesian approach that models the relationship between input variables and output variables [75,76]. It is used for both regression and classification problems.

The Gaussian distribution can be used to describe the distribution of random variables, while the Gaussian process is a generalization of Gaussian distribution, which is used to describe the distribution of functions. A Gaussian process can be defined by its mean function $m(x)$ and the covariance function $k(x, x')$ in the function space:

$$m(x) = E(f(x)) \quad (5)$$

$$k(x, x') = E((f(x) - m(x))(f(x') - m(x'))) \quad (6)$$

The Gaussian process then can be represented as:

$$f(x) = GP \sim (m(x), k(x, x')) \quad (7)$$

Usually, for notation simplicity, we treat the mean function as zero at this time [77,78].

Consider a dataset A with n observations ($A = \{(x_i, y_i) | i = 1, 2, 3, \dots, n\}$), where x_i is an M -dimensional input vector and y_i is the output scalar. The relationship between input data and output data is assumed to be [75,77,79]:

$$y = f(x) + \varepsilon \quad (8)$$

where $f(x)$ represents the arbitrary regression function and ε represents the Gaussian noise that follows an independent, identically distributed Gaussian distribution with zero mean and variance of σ_n^2 ($\varepsilon \sim N(0, \sigma_n^2)$). The input and output data are defined as two matrices $X = [x_1, x_2, \dots, x_n]$ and $Y = [y_1, y_2, \dots, y_n]$. The function collection $f = [f(x_1), f(x_2), \dots, f(x_n)]^T$ is assumed to follow a Gaussian process ($p(f|X) = N(0, K)$), where K is the matrix of covariance function $k(x, x')$:

$$K(X, X) = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \cdots & k(x_n, x_n) \end{bmatrix} \quad (9)$$

The prediction outputs (training outputs y and test outputs y^*) are jointly distributed as a multivariate normal distribution [78]. The joint distribution of the observed target value y and the predicted value y^* at the test location under a priori conditions can be written as:

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X, X) & K(X, X^*) \\ K(X^*, X) & K(X^*, X^*) \end{bmatrix}\right) \quad (10)$$

GRP is then used to compute the predictive distribution of the function values y^* at test points $X^* = [x_1^*, x_2^*, \dots, x_n^*]$:

$$p(y^* | X^*, X, y) \sim N(\bar{f}^*, \text{cov}(f^*)) \quad (11)$$

$$\bar{f}^* = K(X^*, X)K(X, X)^{-1}y \quad (12)$$

$$\text{cov}(f^*) = K(X^*, X^*) - K(X^*, X)K(X, X)^{-1}K(X, X^*) \quad (13)$$

Bayesian optimization is one of the most famous applications of Gaussian process regression [80,81]. When evaluating the objective function is very expensive (such as tuning deep learning hyperparameters), Bayesian optimization methods efficiently find the optimal value by building a Gaussian process model of the objective function to select the next evaluation point.

3.4. Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB)

A **decision tree (DT)** is a supervised learning method utilized in statistics, data mining, and machine learning. It uses a tree-like structure for both regression and classification problems [82,83]. Nodes in a DT that have outgoing edges are referred to as internal nodes, while others are known as leaf nodes (leaves), and branches connect the nodes in DT models. DTs can be categorized into two types: classification trees and regression trees. Classification trees split data into two subsets based on class labels, repeating this process until a stopping criterion is met. Regression trees are another type of DT used in machine learning to solve regression problems. They are used to predict continuous output variables, unlike classification trees, which are used to predict a discrete set of values. Regression trees split data into two subsets based on output data and repeat this process until a stopping criterion is met. DTs have many advantages, including simplicity in understanding and explanation [84]. However, DTs can be highly sensitive. Small changes in the input data can lead to significant effects on trees and final predictions [84]. Due to its reliance on the greedy algorithm, DTs may sometimes fail to return a globally optimal result [85]. The DT method can also build over-complex trees for specific problems, leading to poor generalization from training data (overfitting).

A **random forest (RF)** is an ensemble learning method used for solving classification, regression problems, and other tasks. During training, the RF builds a large number of DTs, which are then used to make predictions [86,87]. In classification tasks, the RF output is the class selected by most of the DTs, while in regression problems, the mean or average prediction value from each DT is returned as the RF result. One of the main advantages of RFs is the ability to overcome the overfitting problem that often occurs in DTs. As the number of trees in an RF increases, the generalization error decreases and stabilizes at a limit value. This means that the performance of an RF will not decrease due to overfitting [86]. However, one of the disadvantages of RFs is losing the intrinsic interpretability of decision trees [88]. Additionally, RFs may not accurately predict the extreme values in continuous variables [89,90].

Gradient boosting (GB) is a popular machine learning algorithm for solving classification and regression problems using a series of weak prediction models, usually in the form of decision trees [91,92]. GB iteratively adds new weak learners to the previous model. At each iteration, the model calculates the negative gradient of the loss function from the prediction values of previous learners. It then trains a new learner with this information to focus on the low-accuracy samples from previous learners. Typically, GB uses mean squared error (MSE) as the loss function [93]. Compared with the RF method, which is also built from DTs, GB is often found to perform better [91,92,94]. This is because GB can effectively reduce bias and variance in the model, leading to better accuracy in prediction. However, GB can be computationally expensive and prone to overfitting if the number of iterations is too high.

Decision trees have demonstrated high accuracy in disease prediction, especially heart disease prediction [95]. Random forest and gradient-boosting classifiers perform very well in a credit-scoring context and are able to cope comparatively well with pronounced class imbalances in these data sets [96].

3.5. K-Nearest Neighbor (k-NN)

The **k-nearest neighbor (k-NN)** method is a popular machine learning algorithm used for regression and classification problems. It assumes that similar samples are often located close to each other in the feature space [97]. For each sample point, the k-NN algorithm

finds the k nearest samples in the dataset and calculates the distances between them and the previous sample points. The input data consist of the samples with the smallest distances satisfying k . In classification problems, the output is the class membership (labels). A sample is classified based on the majority class among the k nearest samples. If $k = 1$, the sample is classified to the nearest class. In regression problems, the output value is the average value of the k nearest samples [98]. One advantage of the k -NN method is its simplicity and fast training process. Additionally, the k -NN algorithm can eliminate noise to some extent [99]. However, the value of k is a sensitive parameter, and a smaller k may lead to overfitting, while a larger k may lead to underfitting. Additionally, the computational cost can be expensive, especially for large datasets [99]. Therefore, it is necessary to carefully select hyperparameters to ensure performance. The application scenarios for the k -NN method include image recognition [100,101], text classification [102,103], and recommendation system [104].

3.6. Multilinear Regression (MLR) and Multivariate Adaptive Regression Splines (MARS)

Multilinear regression (MLR), also known simply as multiple regression, is a simple machine learning method used to model the relationship between two or more predictor variables and a response variable. The model is shown in Equation (14) [105]:

$$y = \alpha_0 + \alpha_1 \times x_1 + \alpha_2 \times x_2 + \dots + \alpha_n \times x_n + \varepsilon \quad (14)$$

where y is the dependent variable (output data), α_0 is the intercept or constant term, and α_1 to α_n are the coefficients of the independent variables (input data) x_1 to x_n , respectively. The aim of MLR is to estimate the values of the α coefficients to minimize the sum of squared errors between the predicted values and the actual values of the dependent variable. The disadvantage of MLR is that this method can only consider the linear relationship between the input and output data.

Multivariate adaptive regression splines (MARS) is a nonlinear regression algorithm used for classification and regression problems. MARS models the relationship between multiple inputs and one output using a series of piecewise linear functions called basis functions [106,107]. Taking the linear basis function as an example, other basis functions are similar. The linear basis function $B_i(x)$ can be defined as three types:

- (1) Constant values T (the intercept).
- (2) A hinge function:

$$B_i(x) = \max(0, x - T) \quad (15)$$

- (3) A function of two or more hinge functions.

MARS model M can be expressed as a sum of basis functions:

$$M = \sum_{i=1}^n t_i B_i(x) \quad (16)$$

where t_i is the constant coefficient. The model starts with a single basis function and adds new functions iteratively. The algorithm selects the optimal location for each hinge in a greedy forward stage-wise manner [108]. MARS has the advantage of generating a simple model that captures complex relationships between inputs and outputs using a small number of basis functions. However, MARS may be sensitive to the choice of the initial basis function and the stopping criteria. MARS is often used in credit scoring [109,110], and species distribution models [111] and also has applications in time series analysis [112].

3.7. Some Other Machine Learning Methods

Sparse polynomial chaos expansion (SPCE) is based on polynomial chaos expansion and does not require assumptions of the performance function [113–115]. Polynomial chaos expansion in the random FEM uses a series of coefficients to present the response

of the system. However, the computation costs for high-dimensional problems are very prohibitive. A step-wise regression technique is used in SPCE to eliminate the non-significant polynomials [114]. There are some variants of SPCE, such as adaptive sparse polynomial chaos [113,115].

A **Bayesian network (BN)** is a probabilistic graphical model that represents a set of random variables and their conditional dependencies using a directed acyclic graph and the probability distribution for each node [116,117]. BNs combine the knowledge of graphs and probability statistics and can be used to solve uncertainty problems with logical reasoning. BNs can implement sequential reasoning (such as estimating the probability of failure) and also allow reversal inference (such as evaluating the sensitivity of factors). Bayesian inference can be used to estimate or update the probability of certain nodes in the network when other nodes are observed. In addition, BNs also show better prediction accuracy on small samples.

3.8. The Hyperparameter Optimization (HPO) Algorithm in Machine Learning

In machine learning, hyperparameters (such as learning rates in ANNs) must be set before training the ML model [118]. Hyperparameter optimization (HPO) is an automated hyperparameter tuning technique aimed at obtaining the best model structure. Grid search is an easy-to-understand HPO method [119]. However, its efficiency is greatly affected by the number of hyperparameters to optimize and the values chosen on the grid. The random search method was proposed as an alternative method to speed up the search process [120]. More advanced HPO methods have also been proposed, which can be called metaheuristic-based optimization algorithms [89]. These algorithms optimize a problem based on a given heuristic function or a cost measure and find a good or acceptable solution within a reasonable amount of time and memory (e.g., particle swarm optimization and ant colony optimization).

4. Application of ML Methods in Slope Stability

As the global population continues to grow, mountainous regions are becoming increasingly densely populated. As a result, more and more infrastructures are being built in close proximity to slopes, which are highly susceptible to landslides. These landslides not only pose a significant risk to human life but also result in significant economic losses [31]. Although analytical and numerical methods have been developed and applied in recent years, most of them focus on deterministic analysis, such as the aforementioned LEM and FEM. Due to geological effects such as weathering, transportation, and stress history, geotechnical engineering materials are not homogeneous but subject to certain spatial variability [121,122].

To improve the accuracy of slope stability analysis, random field theory is used to represent the spatial variability in geotechnical engineering materials [123]. In the early stage, brute-force MCSs were used to estimate slope stability. However, this method requires a large number of samples for the estimation of small probability events. Therefore, some advanced methods were proposed to approximate numerical simulations, such as the response surface method [19,124–126]. These methods essentially fit independent and dependent variables, such as material parameters and the slope *FOS*, to polynomial functions and quickly estimate large numbers of samples [125,127]. Recently, machine learning methods were gradually applied to slope stability problems due to better performance in handling complex problems.

Figure 3 shows the distribution of slope stability problems analyzed in 159 works from the literature retrieved from the Web of Science database. The machine-learning-aided slope stability analysis mainly concerns two problems—prediction of the *FOS* and slope stability classification. Other problems include searching for limit failure surface, slope deformation prediction, establishing failure criterion, and so on.

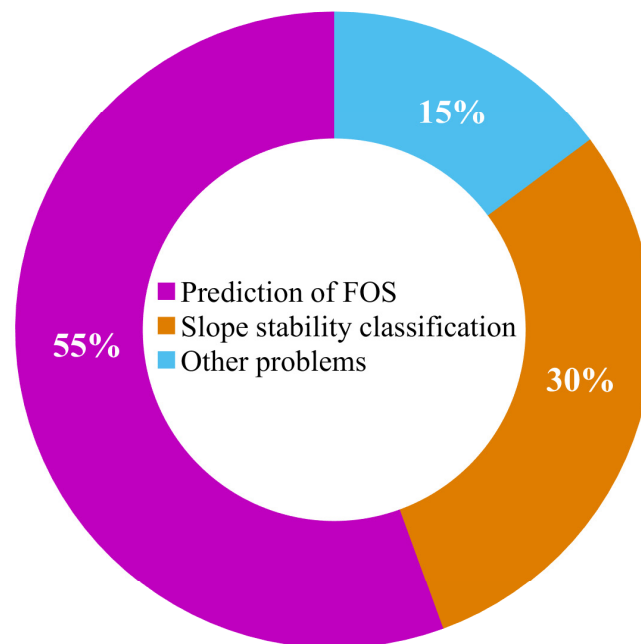


Figure 3. Proportion of slope stability problems in the various studies included in this review.

4.1. Performance Evaluation Metrics in Regression Problems Using Machine Learning

Evaluation metrics are required to assess the performance of ML models. For the use of machine learning surrogate models, it is common to think of all evaluation data points as interpolations. Common evaluation metrics for regression problems can be roughly divided into two categories: the first category reflects the overall fitting degree of the model, such as the coefficient of determination (R^2), while the second category, known as error parameters, directly expresses the difference between the predicted and observed values. The second category can be further divided into relative and absolute differences, such as the mean absolute percentage error (MAPE) and root mean square error (RMSE), respectively.

The **coefficient of determination R^2** is a statistical measure that represents the proportion of the variance in the dependent variable that can be explained by the independent variable(s) included in the model. In other words, it reflects the overall fit of the model. R^2 can be defined as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^n (y_i - y_i^*)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (17)$$

where y_i represents the i th observed value, y_i^* represents the i th predicted value, and \bar{y} represents the mean value of the observed values. R^2 increases with an improvement in model performance. If the predicted value exactly matches the observed value, $R^2 = 1$. However, it is important to note that R^2 can be misleading if used alone, as it does not provide information about the predictive ability or accuracy of the model.

Three metrics are introduced for evaluating the absolute difference between the predicted values and the observed values: mean squared error (MSE), root mean square error (RMSE), and mean absolute error (MAE).

Mean squared error (MSE) is a widely used evaluation metric in machine learning to measure the performance of a regression model. It is calculated as the average of the squares of the differences between the predicted and observed values, and it is always a non-negative value that decreases as the model performance improves. MSE can be defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2 \quad (18)$$

Root mean square error (RMSE) is calculated as the square root of the average of the squared differences between predicted values and actual values. RMSE is preferred over MSE when the output values have different dimensions, as RMSE maintains the same dimension as the output values. RMSE is always a positive value that decreases as the model performance improves. RMSE can be defined as:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2} \quad (19)$$

Mean absolute error (MAE) reflects the mean magnitude of the errors without considering their direction. RMSE and MAE have the same dimension, but RMSE is typically larger than MAE due to the squaring of errors in the calculation of RMSE. The squaring of errors in RMSE places greater emphasis on large errors and can lead to a larger overall value. Additionally, unlike RMSE, MAE is directly proportional to the absolute value of the error, meaning that each error influences MAE in proportion to its magnitude, regardless of its direction [128]. MAE can be defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i^* - y_i| \quad (20)$$

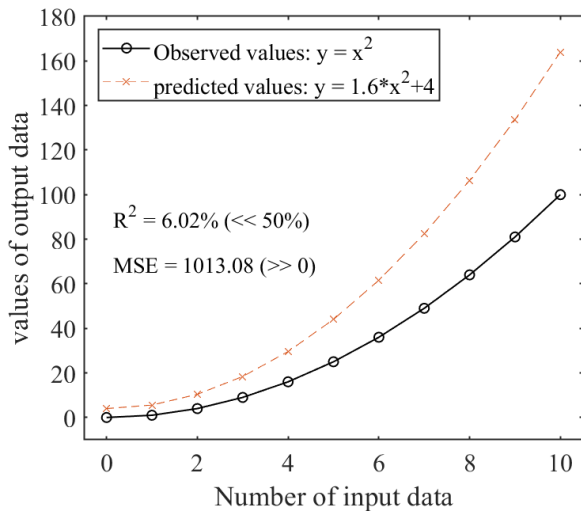
Mean absolute percentage error (MAPE) is one of the relative errors between the predicted values and the observed values. As a dimensionless evaluation metric, MAPE can simply compare the performance of different models on different datasets, and it is intuitive and easy to understand. Compared with RMSE, MAPE is less susceptible to outliers and is more robust. However, MAPE is not suitable for cases where the observed values are zero or close to zero because MAPE tends to infinity. MAPE can be defined as:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i^* - y_i}{y_i} \right| \quad (21)$$

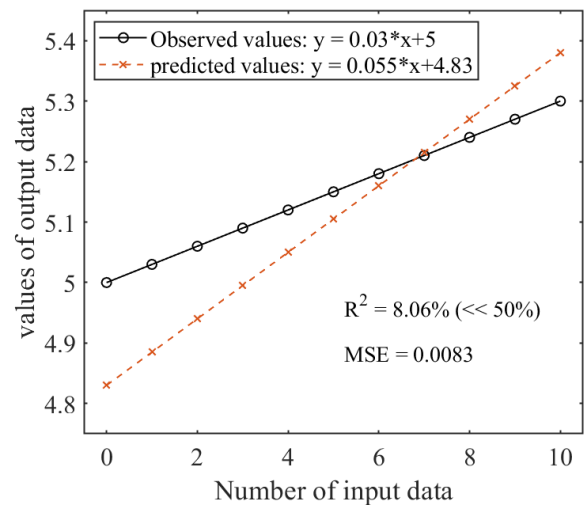
Figure 4 illustrates the impact of the coefficient of determination and error parameters on the evaluation of ML models. Several simple functions are chosen to represent observed and predicted values in four particular cases. These four cases are presented in Figure 4, respectively:

- (a) Low coefficient of determination and high error parameters. In this case, both the overall prediction and the individual predictions are unreliable, which may be due to outliers, such as wrong predictions of boundary values. In addition, non-linear relationships, heteroscedasticity, high noise, and overfitting/underfitting can also lead to this situation.
- (b) Low coefficient of determination and low error parameters. In this case, individual predictions are accurate, but the overall prediction is poor. One reason can be a low slope of the fitting function. Limited by the definition of R^2 , when the slope of the linear fitting function is low (the function value is close to the average value), even if the prediction accuracy is high, the calculation result of R^2 tends to be 0. The model fails to account for the variability in these data. Similarly, non-linear relationships, heteroscedasticity, high noise, and overfitting/underfitting can also lead to this situation.
- (c) High coefficient of determination and high error parameters. In this case, the overall prediction is accurate, but the individual predictions are unreliable. This may be due to the use of a linear relationship to fit a non-linear function (incorrect fitting relationships). Uncertainty in the data, heteroscedasticity, outliers, and overfitting/underfitting can also lead to this situation. For example, if the data have large uncertainties, even if the model can explain some of the variance, the MSE may still be large due to the inherent uncertainty of the data.

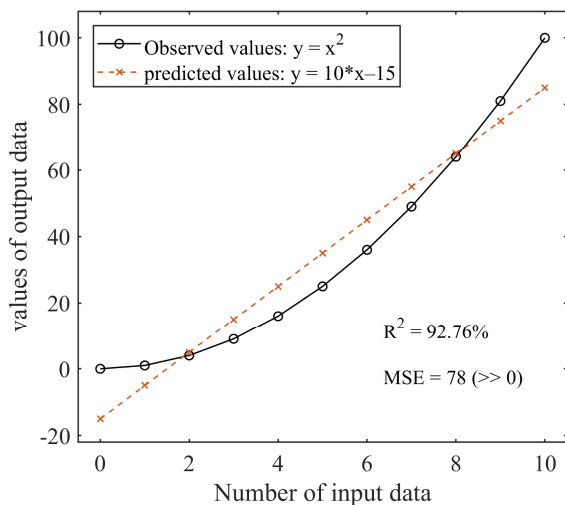
(d) High coefficient of determination and low error parameters. In this case, both the overall prediction and the individual predictions are good.



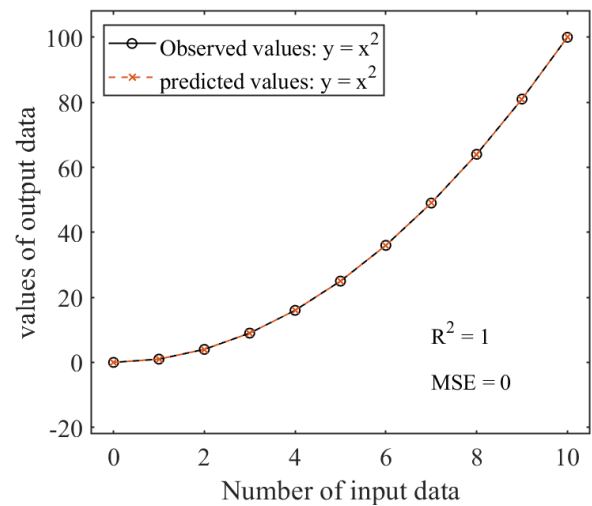
(a)



(b)



(c)



(d)

Figure 4. The importance of the coefficient of determination and error parameters, (a) Low coefficient of determination and high error parameters, (b) Low coefficient of determination and low error parameters, (c) High coefficient of determination and high error parameters (d) High coefficient of determination and low error parameters.

Cases (a)–(c) indicate poor performance of ML models in either overall or individual predictions, while Case (d) shows high accuracy in both overall and individual predicted values. Therefore, only one type of metric can lead to biased estimations of the ML model. It is recommended to consider both the coefficient of determination and error parameters together to accurately evaluate the performance of a model. Outliers and high noise are often eliminated during data processing and detailed information can be found in Shan’s work [129]. Overfitting/underfitting are also common problems in machine learning. Data augmentation, regularization, cross-validation, and early stopping help reduce overfitting problems, while increased model complexity and tuning hyperparameters can fix the underfitting problem.

4.2. Failure Probability of Slopes

In slope stability analysis, the safety margin of slopes is usually represented by the failure probability P_f from the perspective of probability. The limit state function can describe slope stability theoretically. A failure happens when the condition of $G(\Psi) \leq 0$ is satisfied, where Ψ represents the material parameters commonly regarded as random variables or random fields. The failure probability can be expressed as [130,131]:

$$P_f = P[G(\Psi) \leq 0] = \int_{G(\Psi) \leq 0} f(\Psi) d\Psi \quad (22)$$

where $f(\Psi)$ represents the joint probability density function of Ψ . In slope stability analysis, an FOS smaller than 1 usually represents an unstable state. So, $G(\Psi)$ can be expressed as:

$$G(\Psi) = FOS(\Psi) - 1 \leq 0 \quad (23)$$

However, the failure probability cannot be evaluated using the multiple integral in Equation (22) since $f(\Psi)$ is usually unavailable due to the spatial variability in materials in geotechnical problems. Therefore, approximation or simulation methods are needed to replace $f(\Psi)$. MCS is often used to approximate failure probability because it is easy to understand and operate. The failure probability can be obtained using:

$$P_f = \frac{1}{N_{MCS}} \sum_{i=1}^{N_{MCS}} [FOS_i < 1] \quad (24)$$

where FOS_i represents the i th realization of the FOS . $[FOS_i < 1]$ is the Iverson bracket used to judge the state of slopes. $[FOS_i < 1] = 1$ when the FOS_i is less than 1; otherwise, $[FOS_i < 1] = 0$.

The reliability index β is used to evaluate the system reliability without an exact probability distribution function [132]:

$$\beta = \frac{\mu_z}{\sigma_z} = \frac{E[G(\Psi)]}{\sigma[G(\Psi)]} \quad (25)$$

where $\mu_z = E[g(x)]$ and $\sigma_z = \sigma[G(\Psi)]$ are the mean and standard deviation of the performance function, respectively. It represents the distance between mean values of $G(\Psi)$ and 0 (assumed as the failure point) with standard deviations. $G(\Psi)$ is assumed to be normally distributed in Equation (25).

The relationship between failure probability and the reliability index is [77]:

$$P_f = \Phi(-\beta) = 1 - \Phi(\beta) \quad (26)$$

where Φ is the standard normal cumulative distribution function.

In slope stability analysis, uncertainty is caused by many factors, including the nature of the geological material, soil parameters, the water table, seismic intensity, etc. Because there are certain errors or uncertainties in the measurement and estimation of these parameters, the calculation results of slope stability will also be affected by these uncertainties, which makes the stability uncertain. MCSs are widely used in stochastic analysis to estimate the propagation of uncertainty. With the development of computer science, a large number of machine learning methods (such as random forests and neural networks) are proposed and applied to build surrogate models to replace time-consuming numerical evaluations. These algorithms can directly fit the mathematical relationship between input and output, reduce the need for complex and time-consuming calculations, and directly obtain the system response. For slope stability analysis, using machine learning methods as surrogate models to evaluate the FOS can obtain highly accurate results and greatly reduce computation effort.

4.3. Prediction of the Factor of Safety

A total of 159 slope stability studies were collected from the Web of Science that were published in the last two decades, 89 of which focused on the FOS. This kind of machine learning-aided slope stability method has similarities to the so-called surrogate model methods. The key is the use of machine learning to establish high-dimensional nonlinear functions between input and output variables. These nonlinear functions can obtain a large number of predictions for reliability analysis with high efficiency and accuracy.

Taking the research of [31] as an example, a typical machine learning-aided slope stability analysis is presented in Figure 5:

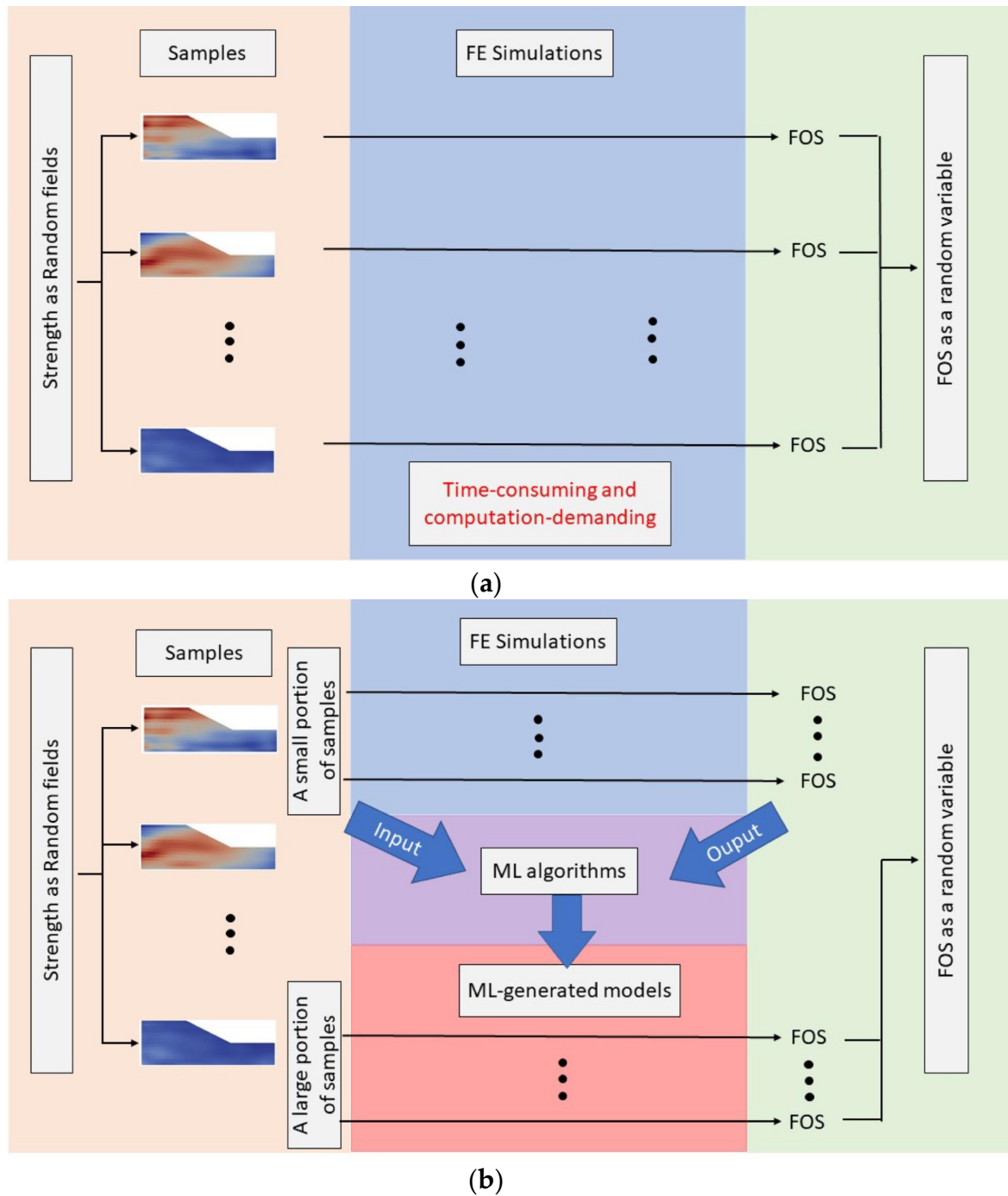


Figure 5. A process of typical machine learning-aided slope stability analysis. (Adapted from [31]) (a) Traditional statistical reliability analysis (b) Machine learning aided statistical reliability analysis.

- (1) **Data generation:** A large number of samples of random variables or random fields are generated to represent the uncertainty and spatial variability in the soil. Karhunen–Loève expansion is a common method to generate random variables or random fields [133,134].
- (2) **Data collection and preprocessing:** Random variables or random fields are mapped into numerical models, such as the LEM and FEM, and the associated *FOS* is calculated using the numerical models. All simulation results are required to be checked to ensure the correct *FOS* is obtained.
- (3) **Model selection and training:** One or more appropriate machine learning models are chosen based on the nature of the problem. A specific machine learning method is used to train and validate models with random variables or random fields as input and the *FOS* as output data. Trained models represent the nonlinear relationship between samples and the *FOS*. The conceptual function captured with machine learning is expressed as:

$$FOS = f(x_1, x_2, \dots, x_n) \quad (27)$$

where x_i represents the i th input variables, such as the cohesion and friction angle in the Mohr–Coulomb model.

- (1) **Model validation and tuning:** The trained models are evaluated using either a validation or test dataset to assess their performance. Model parameters may be adjusted, or alternative models may be explored as necessary to attain the optimal model or a model that closely approaches optimality.
- (2) **Model deployment:** Based on the well-trained models, a large number of predicted *FOS* values can be quickly obtained, and the failure probability can be counted according to the predicted results.

According to step (3) above, the implicit Equation (27) captured with machine learning is also an application of the response surface method (RSM). The RSM has been proven to be an efficient tool for slope stability analysis [127]. The classic RSM uses polynomial regression (usually quadratic regression) to approximate the actual response. The classic RSM may fail to detect the most critical slip surface if multiple failure surfaces are present. The RSM based on machine learning is more powerful when dealing with complex, nonlinear responses or large amounts of data and can perform a global search to find the most dangerous failure surface. For example, a Kriging-based RSM was proposed to conduct the system slope stability study [135]. The undrained shear strain of the soil is the input of the equations. The method performs a global approximation that allows for a more accurate assessment of the system reliability of soil slopes. Even with the same calibration samples, second-order polynomial-based RSMs are less accurate than Kriging-based RSMs.

For the problem of predicting the *FOS*, ANNs were used up to 45 times, accounting for about 50% of the reviewed studies, followed by SVM, which was used 20 times. Table 1 lists some information about the studies on the prediction of the *FOS* and gives the notations of input variables. The rest of this section describes some studies of machine learning methods for predicting the *FOS* of slopes, including high dimensional or highly nonlinear regression [136–138].

Kang et al. [138] proposed a so-called ν -support vector machine (ν -SVM) method to build a surrogate model for predicting the *FOS* and evaluating the system failure probability. Mu'azu [61] combined the teaching–learning-based optimization (TLBO) method and two machine learning algorithms (ANNs and ANFISs) to predict the *FOS* of slopes. Their results indicate that the ANN-TLBO is the best model and has the lowest error. Ji et al. [139] trained surrogate models based on the least-squares support vector machine (LS-SVM) and accurately estimated the stability for spatially variable slopes. Ahangari Nanehkaran et al. [140] compared five different machine learning models for *FOS* predictions and verified the models using a confusion matrix and errors table to confirm the accuracy evaluation indexes. Hsiao et al. [65] proposed a pre-trained model using an ANN and

CNN to directly estimate the safety factor, the trace of the slope slip surface, and finally quickly predict the probability of failure. Jiang et al. [141] trained a surrogate model with a gradient-boosting regression tree to predict the *FOS* under the effect of heavy rainfall. This model accurately evaluated the bench slope *FOS* under the rainfall intensity of the “20-year rainstorm recurrence period”, and its error was smaller than that calculated using a numerical simulation analysis. Lin et al. [142] compared the *FOS* prediction ability of eleven machine learning methods and suggested that SVM, GBR, and Bagging are considered to be the best among these eleven regression methods. Meng et al. [136] used an ANN to evaluate the slope stability of 3D homogeneous dry slopes with different shapes of slope surfaces. A graphical user interface (SlopeLab) was developed based on the pre-trained ANN models. Suman et al. [143] studied both *FOS* predictions and slope stability classifications with three different machine learning methods: functional networks (FNs), MARS, and multigene genetic programming (MGGP). Two surrogate models for predictions and classifications were trained using the same input and output dataset, and prediction model equations were provided. Their results indicate that MARS has better prediction capabilities than FN and MGGP models.

Table 1. Studies of slope stability concerning the prediction of *FOS* with ML.

Reference	Method to Obtain <i>FOS</i>	ML Method	Input Variables	Dimension	Training Sample Size	Advantages	Limitations
Kang et al. (2016) [138]	LEM	ν -SVM	Case 1: S_u, γ Case 2,3,4: c, φ, γ	2D	20	Good generalization capability; adaptable for high-dimensional problems	Single-prediction outputs
Ji et al. (2017) [139]	LEM	LS-SVM	c, φ	2D	50–500	Computationally efficient. good generalization capability; adaptable for high-dimensional problems	Single-prediction outputs
Mu’azu (2022) [61]	FEM	ANFIS, ANN	S_u, β, q, D	2D	504	High accuracy; optimization methods for tuning hyperparameter	Time-consuming
Ahangari Nanehkaran et al. (2022) [140]	LEM	ANN, DT, k-NN, RF, SVM	$\theta, G_s, \rho_d, H, \beta, c, \varphi$	2D	49	Comparison of different machine-learning methods	Unable to determine generalization ability; requires large training sample size
Pandey et al. (2022) [105]	FDM	ANN, MLR	H, β, c, φ	2D	148	High accuracy	Unreliable samples are not excluded
Hsiao et al. (2022) [65]	FEM	ANN, CNN	c, φ, γ	2D	Case1: 200 Case2: 400	High accuracy; consideration of spatial variability; prediction of failure surface	Requires large training sample size
Jiang et al. (2022) [141]	LEM	GB	H, β, R_i, R_d	2D	37	High accuracy; good generalization capability	Requires large memory; high sensitivity of hyperparameters
Lin et al. (2021) [142]	LEM, empirical	ABR, Bagging, BR, ENR, ET, DT, k-NN, MLR, RF, SVM	$H, \beta, c, \varphi, \gamma, u_w$	2D	314	Comparison of different machine-learning methods	Limited accuracy; no consideration of external or trigger factors

Table 1. Cont.

Reference	Method to Obtain FOS	ML Method	Input Variables	Dimension	Training Sample Size	Advantages	Limitations
Falae et al. (2021) [106]	FEM	MARS	c, φ, ν, E	2D	Section X-X': 128, 256 Section Y-Y': 512, 512	Consideration of spatial variability	Requires known prior information; only considers R^2 metric
Cho (2009) [144]	FDM	ANN	c, φ	2D	Case1: 15–25 Case 2: 50–200	High accuracy; parallel distributed processing; strong robustness	Requires a large number of initial parameters; consumes a lot of time during model training
Meng et al. (2021) [136]	FEM	ANN	β, D_r, W_r, D	3D	Purely cohesive soil: 168 Cohesive-frictional soil: 1254	High accuracy; consideration of 3D geometries; consideration of concave slopes surface	Only valid for homogeneous dry slopes without pore pressure
He et al. (2020) [31]	FEM	ANN, SVM	c, φ, γ	2D	300	High accuracy; consideration of spatial variability	Separate samples and training are required for each case
Ray et al. (2020) [145]	FEM, empirical	ANN	E_s, c_s, φ_s, E_r	2D	320	Consideration of weathering effects	Limited generalization capability; large number of input parameters Unable to provide sequential prediction; single-prediction outputs
Liu et al. (2019) [146]	LEM, FEM	MARS	c, φ	2D	Case1: 280 Case1: 210	No need to deal with a large amount of data and high-dimension data	Unable to provide sequential prediction; single-prediction outputs
Wang et al. (2021) [67]	FEM	CNN	c, φ	2D	5000	Adaptable for high-dimensional problems; high performance in feature classification; efficiently handles large samples	Large number of samples required
Suman et al. (2016) [143]	LEM	FN, MARS, MGGP	$c, \varphi, \gamma, H, \beta, r_u$	2D	75	Prediction model equations are provided	No consideration of spatial variability

4.4. Slope Stability Classification

Compared with the prediction of FOS, the classification of slope stability has been extensively studied, with a total of 47 publications identified in our review using the Web of Science. Slope stability classification is significant in landslide disaster prevention and mitigation [10]. Directly classifying the state of a slope can intuitively identify the slope stability. Generally, numerical simulations and analytical methods can rationally and explicitly quantify the state of a specific slope. However, when faced with many landslides or a complex situation, performing a slope stability analysis on all landslides in practice would be time-consuming, computationally expensive, or even impossible. Therefore, the classification of slope stability based on known information is necessary.

Slope stability classification can be roughly divided into two categories. The first type classifies the state of real slopes [10,32]. Researchers obtain terrain conditions, stress history, environmental and meteorological factors, and other conditions that affect slope stability using site investigation and laboratory measurements or from literature or databases [49,52,147]. These conditions and slope stability classifications are used as input and output in the ML training process. The obtained models can accurately predict slope stability under specific conditions.

The second type classifies the stability of slopes using a numerical model [27,98], which is further used to accelerate stochastic analyses. The FOS obtained using numerical calculations or analytical solutions can be used to judge slope stability. Most of the reviewed

studies performed a binary classification of slope stability [10,52]. $FOS = 1$ is often defined as the critical state of stability and failure. Therefore, $FOS > 1$ is stable and $FOS < 1$ represents failure. Therefore, some studies directly define the output as stable and unstable states. However, some studies reported that some slopes may have intermediate states or are unstable when FOS is greater than 1 and stable when FOS is less than 1 [49], so a more detailed classification of slope states is needed. In addition, a binary classification method is proposed to directly judge the relationship between FOS values and 1 without accurate calculations of the FOS [26].

A typical machine learning-aided slope stability classification is similar to the process of predicting the FOS , while the outputs are the slope stability classifications. Among the studies on slope stability classification, the most common machine learning algorithms were ANNs, which were used in twenty-two studies, followed by SVM in eighteen studies.

Table 2 presents some information about the studies on slope stability classification. Some of these studies are introduced below.

Table 2. Studies of slope stability analysis concerning classification with ML (output are slope state classifications).

Reference	Method Classification	ML Method	Input Variables	Training Sample Size	Advantages	Limitations
Zhang et al. (2022) [10]	Empirical method	RF, XGboost, SVM, LR	$E_{fb}, H, \beta, \beta_{in}$	628	High accuracy; good generalization capability	Requires large memory; large number of samples required
Zhu et al. (2022) [32]	Empirical method	ANN, BNB, CNN, DT, GNB, RF, SVM	High-resolution images	868	High accuracy; comparison of different machine-learning methods	Only valid for slopes without shade and vegetation; affected by image quality
Hanandeh (2022) [27]	LEM	GEP	$c, \varphi, \gamma, H, \beta, u$	253	High accuracy	Difficulty in choosing appropriate fitness function; limited ability to handle high-dimensional data
Zhang et al. (2022) [148]	LEM	Ensemble learning	$c, \varphi, \gamma, H, \beta, u_w$	338	High accuracy; good generalization capability	Time-consuming; relies on individual learners
Lin et al. (2022) [149]	FEM, LEM	Ensemble learning	$c, \varphi, \gamma, H, \beta, u$	400	High accuracy; good generalization capability	Time-consuming; relies on individual learners
Azmoon et al. (2021) [150]	LEM	CNN	Photos	95,400	High accuracy; efficiently handles large samples	Requires a large training sample size; affected by image quality
Pham et al. (2021) [30]	LEM	Ensemble learning	$c, \varphi, \gamma, H, \beta$	122	High accuracy; good generalization capability	Time-consuming; relies on individual learners
Yuan & Moayed. (2020) [52]	FEM	ANN	S_u, β, q, D	504	High accuracy; optimization methods for tuning hyperparameter	Unable to determine generalization ability; time-consuming
Yuan & Moayed. (2020) [151]	FEM	ANN	S_u, β, q, D	630	High accuracy; optimization methods for tuning hyperparameter	Unable to determine generalization ability; hard to determine the parameters of the optimization algorithm
Zhou et al. (2019) [29]	LEM	GB	$c, \varphi, \gamma, H, \beta, u_w$	177	High accuracy	Complexity of hyperparameter tuning; time-consuming; requires large training sample size
Lin et al. (2018) [28]	LEM	RF, SVM, GSA, NB	$c, \varphi, \gamma, H, \beta, u$	86	Comparison of different machine learning methods	Time-consuming
Huang et al. (2020) [152]	Empirical method	k-NN	$c, \varphi, \gamma, H, \beta, R_i, E_i, GWL$	50	Robust algorithm; small sample dependence	More validation required

Table 2. Cont.

Reference	Method Classification	ML Method	Input Variables	Training Sample Size	Advantages	Limitations
Ospina-Dávila et al. (2020) [153]	Empirical method	LDC, QDC, k-NN, ANN	21 parameters: soil strength; slope geometry; hydraulic and environmental parameters	165	Significant separability of the classes can be found	Large number of parameters required
Suman et al. (2016) [143]	LEM	FN, MARS, MGGP	$c, \varphi, \gamma, H, \beta, u_w$	75	Prediction model equations are provided	No consideration of spatial variability

Zhang et al. [10] considered an ensemble learning-based method that combined RF and XGBoost methods to conduct a slope stability classification and explore the importance of 12 influencing variables. Zhu et al. [32] proposed a classification framework to categorize rock blocks based on the principles of block theory using the CNN method. The surrogate model can classify three types of rocks (key blocks, trapped blocks, and stable blocks) using high-resolution images. Yuan & Moayedhi [52] used six metaheuristic algorithms to optimize the classification ability of machine learning methods. Their results show that metaheuristic algorithms could greatly improve the accuracy of classification models from 2% to 27%, among which the genetic algorithm performed best.

Ensemble learning technology is also popular in slope stability classification. Zhang et al. [148] built a margin distance minimization selective ensemble (MDMSE) method to deal with slope stability classification. This ensemble learning is established using four individual learners (k-NN, SVM, DT, and ANN). Compared with common single-machine learning models and ensemble models, MDMSE shows better generalization ability, better recognition accuracy, and faster identification speed. Lin et al. [149] developed an ensemble learning model with eight individual learners for classification problems. A parameter analysis was conducted with three types of slope parameters: material parameters, geometry parameters, and hydraulic parameters. Their results show that material parameters were the most sensitive factors to slope stability, followed by geometry and hydraulic parameters. They suggested that the cohesive force of geomaterials and the internal friction angle should be improved for treating landslides.

5. Future Research Perspective

The number of papers published in the past two decades shows that more and more machine learning methods, including several commonly used machine learning methods and some of the latest advanced machine learning methods (such as CNNs), are applied to solve slope stability problems. In addition, machine learning has greatly promoted the implementation of reliability analysis in geotechnical engineering practice. However, there are still some challenges in the current literature, including:

- (1) The spatial variation in shear strength parameters and hydraulic parameters is commonly considered in slope reliability and risk analysis. However, research on the hydro-mechanical coupling between these parameters is still limited. This means that the impact of rainfall or the groundwater table on slope stability is often not fully accounted for, as both hydraulic and shear strength parameters can have a significant influence.
- (2) Previous studies have mainly focused on simplified models, and there is a lack of realistic geotechnical engineering case applications. Moreover, most studies were limited to 2D slope simulations. Although 2D slope stability analysis has been extensively investigated, it neglects the effect of 3D spatial variation, which is a crucial factor in real-world applications. Despite being a technical challenge due to the larger computational efforts involved, several studies have demonstrated that 3D slope stability

analysis results can significantly differ from those obtained using 2D models [136]. Therefore, future research should focus on analyzing existing 3D slope stability data using ML to provide more accurate and reliable results.

- (3) A site investigation often captures only one or two properties of material spatial variability, which can make it challenging to accurately estimate fluctuation scales and autocorrelations. The generation of random fields currently relies on selecting a theoretical autocorrelation function, which may not accurately represent the true fluctuation scale and autocorrelation in soil parameters. In the future, machine learning could potentially be utilized in field surveys to capture more realistic estimates of these parameters, leading to improved reliability in slope stability analysis.
- (4) Due to the ease of establishing and connecting databases, as well as the advances in monitoring technologies, it is anticipated that more comprehensive and detailed input data can be obtained for the slope stability classification problem, including environmental factors, meteorological information, and hydrogeological conditions. This will lead to more precise predictions of slope conditions.

6. Conclusions

This paper reviews the application of ML to slope stability analysis problems in the past two decades. Slope stability has always been an important branch of geotechnical engineering. The limit equilibrium method and the finite element method are commonly used to calculate the *FOS*. Random field theory is used to simulate the spatial variability in soil.

Early use of the brute-force Monte Carlo method for reliability analysis is time-consuming and inefficient. With the development of machine learning techniques, a large number of applications have contributed to the assessment of slope stability. Some common machine learning algorithms are introduced in Section 3. ANNs and SVM occupy a mainstream position in a large number of the reviewed studies, but we can also see that various advanced methods are gradually emerging, such as ANFIS (Figure 2). Each machine learning algorithm has advantages and disadvantages. The selection of machine learning algorithms should not blindly pursue 'complexity' and an 'advanced' method but should be suitable for the research problem.

Slope stability problems can be broadly categorized into two groups: those predicting the *FOS* and those classifying slope stability. In addition, a few studies focused on searching for the most critical sliding surfaces, designing failure criteria, and other related issues. Considering that machine learning is a "black box" technology, it is critical to select evaluation metrics judiciously for studying these problems. The coefficient of determination characterizes the overall fitting level of the model, while error parameters represent the individual prediction performance of the model. Ignoring any of them may lead to a biased evaluation of the model. Therefore, it is recommended to use both the coefficient of determination and error parameters to ensure a reasonable assessment of a model's performance.

While machine learning has been applied extensively to slope stability problems, several issues require further attention in future research. The hydraulic parameters and shear strength parameters have a significant influence on slope stability, and the problem of hydro-mechanical coupling remains a challenge in slope stability analysis. The analysis of 3D slope stability with the assistance of machine learning is important for geotechnical engineers and researchers, given the impact of 3D spatial variation on the stability of slopes. To enhance the reliability of slope stability analysis, it is important to explore how to use machine learning methods to capture more accurate fluctuation scales and autocorrelations. Additionally, the proper use of monitoring data and databases can contribute to an improvement in machine learning model performance.

In general, machine learning has been widely used in slope stability analysis and shows significant potential. With the availability of various open-source machine learning libraries, it has become more accessible to researchers. In the future, machine learning is

expected to contribute to a more accurate analysis of slope stability problems and prevent disasters caused by landslides.

Author Contributions: Conceptualization, H.X.; methodology, H.X.; writing—original draft preparation, H.X., F.S. and G.N.; writing—review and editing, X.H. and D.S.; supervision, X.H. and D.S.; funding acquisition, D.S. All authors have read and agreed to the published version of this manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ANFIS	adaptive neuro-fuzzy inference system	ANN	artificial neural network
BN	Bayesian network	CNN	convolutional neural network
DNN	deep neural network	DT	decision tree
FDM	finite difference method	FEM	finite element method
FOS	factor of safety	FN	functional network
GB	gradient boosting	GIM	gravity increasing method
GPR	Gaussian process regression	HPO	hyperparameter optimization
k-NN	k-nearest neighbor	LDC	linear discriminant classifier
LEM	limit equilibrium method	LS-SVM	least squares support vector machine
MAE	mean absolute error	MAPE	mean absolute percentage error
MARS	multivariate adaptive regression splines	MCS	Monte Carlo simulation
MDMSE	margin distance minimization selective ensemble	MGGP	multigene genetic programming
ML	machine learning	MLR	multilinear regression
MSE	mean squared error	PDF	probability density function
QDC	quadratic discriminant classifier	R^2	coefficient of determination
RF	random forest	RMSE	root mean square error
RSM	response surface method	SPCE	sparse polynomial chaos expansion
SPH	smoothed-particle hydrodynamics	SRM	strength reduction method
SVM	support vector machine	TLBO	teaching–learning-based optimization
ν -SVM	ν -support vector machine		

Notations

S_u	Undrained shear strength	u_w	Pore pressure ratio
β	Slope angle	ν	Poisson's ratio
q	External force on the footing	E	Elastic modulus
D	Ratio of setback distance	W_r	Width ratio
θ	Water content	$D_g = c\gamma H \tan\phi$	Dimensionless group
G_s	Specific gravity	E_s	Young's modulus of residual soil
ρ_d	Dry density	c_s	Cohesion of residual soil
H	Slope height	ϕ_s	Friction of residual soil
c	Cohesion	E_r	Young's modulus of rock mass
ϕ	Friction angle	c_r	Cohesion of rock mass
γ	Unit weight	ϕ_r	Friction of rock mass
R_i	Rainfall intensity	c_j	Cohesion of joint interface
R_d	Rainfall duration	ϕ_j	Friction of joint interface
D_r	Depth ratio	E_{fb}	Front and back edge elevations
β_{in}	Inclination angle	D_{dip}	Dip direction
V_l	Landslide volume	C_v	Categorical variables
E_i	Earthquake intensity	GWL	Ground water level

References

1. Meyerhof, G.G. Limit Analysis and Soil Plasticity, W.F. Chen (Ed.), Elsevier, Amsterdam (1975), p. 638 Dfl. 240.00. *Eng. Geol.* **1976**, *10*, 79. [[CrossRef](#)]
2. Sloan, S.W. Lower Bound Limit Analysis Using Finite Elements and Linear Programming. *Int. J. Numer. Anal. Methods Geomech.* **1988**, *12*, 61–77. [[CrossRef](#)]
3. Dawson, E.M.; Roth, W.H.; Drescher, A. Slope Stability Analysis by Strength Reduction. *Géotechnique* **1999**, *49*, 835–840. [[CrossRef](#)]
4. Griffiths, D.V.; Lane, P.A. Slope Stability Analysis by Finite Elements. *Géotechnique* **1999**, *49*, 387–403. [[CrossRef](#)]
5. Kaur, A.; Sharma, R.K. Slope Stability Analysis Techniques: A Review. *Int. J. Eng. Appl. Sci. Technol.* **2016**, *1*, 52–57.
6. Pourkhosravani, A.; Kalantari, B. A Review of Current Methods for Slope Stability Evaluation. *Electron. J. Geotech. Eng.* **2011**, *16*, 1245–1254.
7. Wallace, M.I.; Ng, K.C. Development and Application of Underground Space Use in Hong Kong. *Tunn. Undergr. Space Technol.* **2016**, *55*, 257–279. [[CrossRef](#)]
8. Wei, X.; Zhang, L.; Yang, H.-Q.; Zhang, L.; Yao, Y.-P. Machine Learning for Pore-Water Pressure Time-Series Prediction: Application of Recurrent Neural Networks. *Geosci. Front.* **2021**, *12*, 453–467. [[CrossRef](#)]
9. Zhang, J.; Tang, W.H.; Zhang, L.M. Efficient Probabilistic Back-Analysis of Slope Stability Model Parameters. *J. Geotech. Geoenviron. Eng.* **2010**, *136*, 99–109. [[CrossRef](#)]
10. Zhang, W.; Li, H.; Han, L.; Chen, L.; Wang, L. Slope Stability Prediction Using Ensemble Learning Techniques: A Case Study in Yunyang County, Chongqing, China. *J. Rock Mech. Geotech. Eng.* **2022**, *14*, 1089–1099. [[CrossRef](#)]
11. Wesley, L.D.; Leelaratanam, V. Shear Strength Parameters from Back-Analysis of Single Slips. *Géotechnique* **2001**, *51*, 373–374. [[CrossRef](#)]
12. Tiwari, B.; Brandon, T.L.; Marui, H.; Tuladhar, G.R. Comparison of Residual Shear Strengths from Back Analysis and Ring Shear Tests on Undisturbed and Remolded Specimens. *J. Geotech. Geoenviron. Eng.* **2005**, *131*, 1071–1079. [[CrossRef](#)]
13. Griffiths, D.V.; Fenton, G.A. Bearing Capacity of Spatially Random Soil: The Undrained Clay Prandtl Problem Revisited. *Géotechnique* **2001**, *51*, 351–359. [[CrossRef](#)]
14. Sun, X.; Zeng, P.; Li, T.; Wang, S.; Jimenez, R.; Feng, X.; Xu, Q. From Probabilistic Back Analyses to Probabilistic Run-out Predictions of Landslides: A Case Study of Heifangtai Terrace, Gansu Province, China. *Eng. Geol.* **2021**, *280*, 105950. [[CrossRef](#)]
15. He, X.; Wang, F.; Li, W.; Sheng, D. Deep Learning for Efficient Stochastic Analysis with Spatial Variability. *Acta Geotech.* **2022**, *17*, 1031–1051. [[CrossRef](#)]
16. Ling, Q.; Zhang, Q.; Wei, Y.; Kong, L.; Zhu, L. Slope Reliability Evaluation Based on Multi-Objective Grey Wolf Optimization-Multi-Kernel-Based Extreme Learning Machine Agent Model. *Bull. Eng. Geol. Environ.* **2021**, *80*, 2011–2024. [[CrossRef](#)]
17. Li, J.-Z.; Zhang, S.-H.; Liu, L.-L.; Huang, L.; Cheng, Y.-M.; Dias, D. Probabilistic Analysis of Pile-Reinforced Slopes in Spatially Variable Soils with Rotated Anisotropy. *Comput. Geotech.* **2022**, *146*, 104744. [[CrossRef](#)]
18. Allaix, D.L.; Carbone, V.I. An Improvement of the Response Surface Method. *Struct. Saf.* **2011**, *33*, 165–172. [[CrossRef](#)]
19. Wong, F.S. Slope Reliability and Response Surface Method. *J. Geotech. Eng.* **1985**, *111*, 32–53. [[CrossRef](#)]
20. Kaymaz, I.; McMahon, C.A. A Response Surface Method Based on Weighted Regression for Structural Reliability Analysis. *Probabilistic Eng. Mech.* **2005**, *20*, 11–17. [[CrossRef](#)]
21. Kim, S.-H.; Na, S.-W. Response Surface Method Using Vector Projected Sampling Points. *Struct. Saf.* **1997**, *19*, 3–19. [[CrossRef](#)]
22. Desai, K.M.; Survase, S.A.; Saudagar, P.S.; Lele, S.S.; Singhal, R.S. Comparison of Artificial Neural Network (ANN) and Response Surface Methodology (RSM) in Fermentation Media Optimization: Case Study of Fermentative Production of Scleroglucan. *Biochem. Eng. J.* **2008**, *41*, 266–273. [[CrossRef](#)]
23. Abbasi, B.; Mahlooji, H. Improving Response Surface Methodology by Using Artificial Neural Network and Simulated Annealing. *Expert Syst. Appl.* **2012**, *39*, 3461–3468. [[CrossRef](#)]
24. Yang, H.-Q.; Zhang, L.; Pan, Q.; Phoon, K.-K.; Shen, Z. Bayesian Estimation of Spatially Varying Soil Parameters with Spatiotemporal Monitoring Data. *Acta Geotech.* **2021**, *16*, 263–278. [[CrossRef](#)]
25. Liu, X.; Wang, Y.; Koo, R.C.H.; Kwan, J.S.H. Development of a Slope Digital Twin for Predicting Temporal Variation of Rainfall-Induced Slope Instability Using Past Slope Performance Records and Monitoring Data. *Eng. Geol.* **2022**, *308*, 106825. [[CrossRef](#)]
26. Zeng, P.; Zhang, T.; Li, T.; Jimenez, R.; Zhang, J.; Sun, X. Binary Classification Method for Efficient and Accurate System Reliability Analyses of Layered Soil Slopes. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2022**, *16*, 435–451. [[CrossRef](#)]
27. Hanandeh, S. Evaluation Circular Failure of Soil Slopes Using Classification and Predictive Gene Expression Programming Schemes. *Front. Built Env.* **2022**, *8*, 30. [[CrossRef](#)]
28. Lin, Y.; Zhou, K.; Li, J. Prediction of Slope Stability Using Four Supervised Learning Methods. *IEEE Access* **2018**, *6*, 31169–31179. [[CrossRef](#)]
29. Zhou, J.; Li, E.; Yang, S.; Wang, M.; Shi, X.; Yao, S.; Mitri, H.S. Slope Stability Prediction for Circular Mode Failure Using Gradient Boosting Machine Approach Based on an Updated Database of Case Histories. *Saf. Sci.* **2019**, *118*, 505–518. [[CrossRef](#)]
30. Pham, K.; Kim, D.; Park, S.; Choi, H. Ensemble Learning-Based Classification Models for Slope Stability Analysis. *Catena* **2021**, *196*, 104886. [[CrossRef](#)]

31. He, X.; Xu, H.; Sabetamal, H.; Sheng, D. Machine Learning Aided Stochastic Reliability Analysis of Spatially Variable Slopes. *Comput. Geotech.* **2020**, *126*, 103711. [[CrossRef](#)]
32. Zhu, H.; Azarafza, M.; Akgün, H. Deep Learning-Based Key-Block Classification Framework for Discontinuous Rock Slopes. *J. Rock Mech. Geotech. Eng.* **2022**, *14*, 1131–1139. [[CrossRef](#)]
33. Bello, O.; Holzmann, J.; Yaqoob, T.; Teodoriu, C. Application of Artificial Intelligence Methods in Drilling System Design and Operations: A Review of the State of the Art. *J. Artif. Intell. Soft Comput. Res.* **2015**, *5*, 121–139. [[CrossRef](#)]
34. Nonoyama, H.; Moriguchi, S.; Sawada, K.; Yashima, A. Slope Stability Analysis Using Smoothed Particle Hydrodynamics (SPH) Method. *Soils Found.* **2015**, *55*, 458–470. [[CrossRef](#)]
35. Bui, H.H.; Fukagawa, R.; Sako, K.; Wells, J.C. Slope Stability Analysis and Discontinuous Slope Failure Simulation by Elasto-Plastic Smoothed Particle Hydrodynamics (SPH). *Géotechnique* **2011**, *61*, 565–574. [[CrossRef](#)]
36. He, X.; Liang, D.; Bolton, M.D. Run-out of Cut-Slope Landslides: Mesh-Free Simulations. *Géotechnique* **2018**, *68*, 50–63. [[CrossRef](#)]
37. He, X.; Liang, D.; Wu, W.; Cai, G.; Zhao, C.; Wang, S. Study of the Interaction between Dry Granular Flows and Rigid Barriers with an SPH Model. *Int. J. Numer. Anal. Methods Geomech.* **2018**, *42*, 1217–1234. [[CrossRef](#)]
38. Fredlund, D.G. Analytical Methods for Slope Stability Analysis. In Proceedings of the 4th International Symposium on Landslides, Toronto, ON, Canada, 16–21 September 1984; Volume 250.
39. Bishop, A.W. The Use of the Slip Circle in the Stability Analysis of Slopes. *Géotechnique* **1955**, *5*, 7–17. [[CrossRef](#)]
40. Kalatehjari, R.; Ali, N. A Review of Three-Dimensional Slope Stability Analyses Based on Limit Equilibrium Method. *Electron. J. Geotech. Eng.* **2013**, *18*, 119–134.
41. Abramson, L.W.; Lee, T.S.; Sharma, S.; Boyce, G.M. *Slope Stability and Stabilization Methods*; John Wiley & Sons: Hoboken, NJ, USA, 2001; ISBN 0471384933.
42. Zhu, D.Y.; Lee, C.F.; Jiang, H.D.; Lee, C.F.; Iang, H.D.J. Generalised Framework of Limit Equilibrium Methods for Slope Stability Analysis. *Géotechnique* **2003**, *53*, 377–395. [[CrossRef](#)]
43. Zienkiewicz, O.C.; Taylor, R.L.; Zhu, J.Z. *The Finite Element Method: Its Basis and Fundamentals*; Elsevier: Amsterdam, The Netherlands, 2005; ISBN 008047277X.
44. Belytschko, T.; Liu, W.K.; Moran, B.; Elkhodary, K. *Nonlinear Finite Elements for Continua and Structures*; John Wiley & Sons: Hoboken, NJ, USA, 2014; ISBN 1118632702.
45. Bathe, K.-J. *Finite Element Procedures*; Klaus-Jurgen Bathe: Cambridge, MA, USA, 2006; ISBN 097900490X.
46. Matsui, T.; San, K.-C. Finite Element Slope Stability Analysis by Shear Strength Reduction Technique. *Soils Found.* **1992**, *32*, 59–70. [[CrossRef](#)]
47. Cheng, Y.M.; Lansivaara, T.; Wei, W.B. Two-Dimensional Slope Stability Analysis by Limit Equilibrium and Strength Reduction Methods. *Comput. Geotech.* **2007**, *34*, 137–150. [[CrossRef](#)]
48. Sternik, K. Comparison of Slope Stability Predictions by Gravity Increase and Shear Strength Reduction Methods. *Czas. Tech. Sr.* **2013**, *110*, 121–130.
49. Baghbani, A.; Choudhury, T.; Costa, S.; Reiner, J. Application of Artificial Intelligence in Geotechnical Engineering: A State-of-the-Art Review. *Earth Sci. Rev.* **2022**, *228*, 103991. [[CrossRef](#)]
50. Zhang, H.; Nguyen, H.; Bui, X.-N.; Pradhan, B.; Asteris, P.G.; Costache, R.; Aryal, J. A Generalized Artificial Intelligence Model for Estimating the Friction Angle of Clays in Evaluating Slope Stability Using a Deep Neural Network and Harris Hawks Optimization Algorithm. *Eng. Comput.* **2021**, *38*, 3901–3914. [[CrossRef](#)]
51. Wang, H.; Moayedi, H.; Kok Foong, L. Genetic Algorithm Hybridized with Multilayer Perceptron to Have an Economical Slope Stability Design. *Eng. Comput.* **2021**, *37*, 3067–3078. [[CrossRef](#)]
52. Yuan, C.; Moayedi, H. Evaluation and Comparison of the Advanced Metaheuristic and Conventional Machine Learning Methods for the Prediction of Landslide Occurrence. *Eng. Comput.* **2020**, *36*, 1801–1811. [[CrossRef](#)]
53. Liu, Z.; Shao, J.; Xu, W.; Chen, H.; Zhang, Y. An Extreme Learning Machine Approach for Slope Stability Evaluation and Prediction. *Nat. Hazards* **2014**, *73*, 787–804. [[CrossRef](#)]
54. Kang, F.; Li, J.-S.; Wang, Y.; Li, J. Extreme Learning Machine-Based Surrogate Model for Analyzing System Reliability of Soil Slopes. *Eur. J. Environ. Civ. Eng.* **2017**, *21*, 1341–1362. [[CrossRef](#)]
55. Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Netw.* **1989**, *2*, 359–366. [[CrossRef](#)]
56. Palmer, A.; Montano, J.J.; Sesé, A. Designing an Artificial Neural Network for Forecasting Tourism Time Series. *Tour. Manag.* **2006**, *27*, 781–790. [[CrossRef](#)]
57. Ghiassi, M.; Saidane, H.; Zimbra, D.K. A Dynamic Artificial Neural Network Model for Forecasting Time Series Events. *Int. J. Forecast.* **2005**, *21*, 341–362. [[CrossRef](#)]
58. Billsus, D.; Pazzani, M.J. Learning Collaborative Information Filters. In Proceedings of the Fifteenth International Conference on Machine Learning, Madison, WI, USA, 24–27 July 1998; Volume 98, pp. 46–54.
59. Jang, J.-S. ANFIS: Adaptive-Neural-Based Fuzzy Inference System. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [[CrossRef](#)]
60. Jang, J.-S.R. Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm. In Proceedings of the AAI, Anaheim, CA, USA, 14–19 July 1991; Volume 91, pp. 762–767.

61. Mu'azu, M.A. Enhancing Slope Stability Prediction Using Fuzzy and Neural Frameworks Optimized by Metaheuristic Science. *Math. Geosci.* **2022**, *55*, 263–285. [[CrossRef](#)]
62. Al-Mahasneh, M.; Aljarrah, M.; Rababah, T.; Alu'datt, M. Application of Hybrid Neural Fuzzy System (ANFIS) in Food Processing and Technology. *Food Eng. Rev.* **2016**, *8*, 351–366. [[CrossRef](#)]
63. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
64. He, X.; Xu, H.; Sheng, D. Ready-to-Use Deep-Learning Surrogate Models for Problems with Spatially Variable Inputs and Outputs. *Acta Geotech.* **2022**, *18*, 1681–1698. [[CrossRef](#)]
65. Hsiao, C.-H.; Chen, A.Y.; Ge, L.; Yeh, F.-H. Performance of Artificial Neural Network and Convolutional Neural Network on Slope Failure Prediction Using Data from the Random Finite Element Method. *Acta Geotech.* **2022**, *17*, 5801–5811. [[CrossRef](#)]
66. Xu, H.; He, X.; Pradhan, B.; Sheng, D. A Pre-Trained Deep-Learning Surrogate Model for Slope Stability Analysis with Spatial Variability. *Soils Found.* **2023**, *63*, 101321. [[CrossRef](#)]
67. Wang, Z.-Z.; Goh, S.H. Novel Approach to Efficient Slope Reliability Analysis in Spatially Variable Soils. *Eng. Geol.* **2021**, *281*, 105989. [[CrossRef](#)]
68. Mikolov, T.; Karafiát, M.; Burget, L.; Cernocký, J.; Khudanpur, S. Recurrent Neural Network Based Language Model. In Proceedings of the Interspeech, Makuhari, Japan, 26–30 September 2010; Volume 2, pp. 1045–1048.
69. Graves, A.; Mohamed, A.; Hinton, G. Speech Recognition with Deep Recurrent Neural Networks. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; IEEE: New York, NY, USA, 2013; pp. 6645–6649.
70. Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A. Language Models Are Few-Shot Learners. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 1877–1901.
71. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A Training Algorithm for Optimal Margin Classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; pp. 144–152.
72. Sun, A.; Lim, E.-P.; Ng, W.-K. Web Classification Using Support Vector Machine. In Proceedings of the 4th International Workshop on Web Information and Data Management, McLean, VA, USA, 8 November 2002; pp. 96–99.
73. Byvatov, E.; Schneider, G. Support Vector Machine Applications in Bioinformatics. *Appl. Bioinform.* **2003**, *2*, 67–77.
74. Venkatesan, C.; Karthigaikumar, P.; Paul, A.; Satheeskumaran, S.; Kumar, R. ECG Signal Preprocessing and SVM Classifier-Based Abnormality Detection in Remote Healthcare Applications. *IEEE Access* **2018**, *6*, 9767–9773. [[CrossRef](#)]
75. Mahmoodzadeh, A.; Mohammadi, M.; Ibrahim, H.H.; Rashid, T.A.; Aldalwie, A.H.M.; Ali, H.F.H.; Daraei, A. Tunnel Geomechanical Parameters Prediction Using Gaussian Process Regression. *Mach. Learn. Appl.* **2021**, *3*, 100020. [[CrossRef](#)]
76. Zhu, B.; Hiraishi, T.; Pei, H.; Yang, Q. Efficient Reliability Analysis of Slopes Integrating the Random Field Method and a Gaussian Process Regression-based Surrogate Model. *Int. J. Numer. Anal. Methods Geomech.* **2021**, *45*, 478–501. [[CrossRef](#)]
77. Kang, F.; Han, S.; Salgado, R.; Li, J. System Probabilistic Stability Analysis of Soil Slopes Using Gaussian Process Regression with Latin Hypercube Sampling. *Comput. Geotech.* **2015**, *63*, 13–25. [[CrossRef](#)]
78. Wang, C.; Wu, X.; Kozłowski, T. Gaussian Process-Based Inverse Uncertainty Quantification for Trace Physical Model Parameters Using Steady-State Psbt Benchmark. *Nucl. Sci. Eng.* **2019**, *193*, 100–114. [[CrossRef](#)]
79. Kang, F.; Xu, B.; Li, J.; Zhao, S. Slope Stability Evaluation Using Gaussian Processes with Various Covariance Functions. *Appl. Soft Comput.* **2017**, *60*, 387–396. [[CrossRef](#)]
80. Snoek, J.; Larochelle, H.; Adams, R.P. Practical Bayesian Optimization of Machine Learning Algorithms. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 2951–2959.
81. Frazier, P.I. A Tutorial on Bayesian Optimization. *arXiv* **2018**, arXiv:1807.02811.
82. Kardani, N.; Zhou, A.; Nazem, M.; Shen, S.-L. Improved Prediction of Slope Stability Using a Hybrid Stacking Ensemble Method Based on Finite Element Analysis and Field Data. *J. Rock Mech. Geotech. Eng.* **2021**, *13*, 188–201. [[CrossRef](#)]
83. Pekel, E. Estimation of Soil Moisture Using Decision Tree Regression. *Theor. Appl. Clim.* **2020**, *139*, 1111–1119. [[CrossRef](#)]
84. Gareth, J.; Daniela, W.; Trevor, H.; Robert, T. *An Introduction to Statistical Learning: With Applications in R*; Springer: Berlin/Heidelberg, Germany, 2013.
85. Ben-Gal, I.; Dana, A.; Shkolnik, N.; Singer, G. Efficient Construction of Decision Trees by the Dual Information Distance Method. *Qual. Technol. Quant. Manag.* **2014**, *11*, 133–147. [[CrossRef](#)]
86. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
87. Nath, S.K.; Sengupta, A.; Srivastava, A. Remote Sensing GIS-Based Landslide Susceptibility & Risk Modeling in Darjeeling–Sikkim Himalaya Together with FEM-Based Slope Stability Analysis of the Terrain. *Nat. Hazards* **2021**, *108*, 3271–3304.
88. Ziegler, A.; König, I.R. Mining Data with Random Forests: Current Options for Real-world Applications. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2014**, *4*, 55–63. [[CrossRef](#)]
89. Shah, A.D.; Bartlett, J.W.; Carpenter, J.; Nicholas, O.; Hemingway, H. Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. *Am. J. Epidemiol.* **2014**, *179*, 764–774. [[CrossRef](#)]
90. Mendez, G.; Lohr, S. Estimating Residual Variance in Random Forest Regression. *Comput. Stat. Data Anal.* **2011**, *55*, 2937–2950. [[CrossRef](#)]

91. Pirayonesi, S.M.; El-Diraby, T.E. Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index. *J. Infrastruct. Syst.* **2020**, *26*, 04019036. [[CrossRef](#)]
92. Hastie, T.; Tibshirani, R.; Friedman, J.; Hastie, T.; Tibshirani, R.; Friedman, J. Boosting and Additive Trees. *Elem. Stat. Learn. Data Min. Inference Predict.* **2009**, *103*, 337–387.
93. Karir, D.; Ray, A.; Bharati, A.K.; Chaturvedi, U.; Rai, R.; Khandelwal, M. Stability Prediction of a Natural and Man-Made Slope Using Various Machine Learning Algorithms. *Transp. Geotech.* **2022**, *34*, 100745. [[CrossRef](#)]
94. Madeh Pirayonesi, S.; El-Diraby, T.E. Using Machine Learning to Examine Impact of Type of Performance Indicator on Flexible Pavement Deterioration Modeling. *J. Infrastruct. Syst.* **2021**, *27*, 04021005. [[CrossRef](#)]
95. Soni, J.; Ansari, U.; Sharma, D.; Soni, S. Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction. *Int. J. Comput. Appl.* **2011**, *17*, 43–48. [[CrossRef](#)]
96. Brown, I.; Mues, C. An Experimental Comparison of Classification Algorithms for Imbalanced Credit Scoring Data Sets. *Expert Syst. Appl.* **2012**, *39*, 3446–3453. [[CrossRef](#)]
97. Peterson, L.E. K-Nearest Neighbor. *Scholarpedia* **2009**, *4*, 1883. [[CrossRef](#)]
98. Mahmoodzadeh, A.; Mohammadi, M.; Farid Hama Ali, H.; Hashim Ibrahim, H.; Nariman Abdulhamid, S.; Nejati, H.R. Prediction of Safety Factors for Slope Stability: Comparison of Machine Learning Techniques. *Nat. Hazards* **2022**, *111*, 1771–1799. [[CrossRef](#)]
99. Bhatia, N. Survey of Nearest Neighbor Techniques. *arXiv* **2010**, arXiv:1007.0085.
100. Zhang, H.; Berg, A.C.; Maire, M.; Malik, J. SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; IEEE: New York, NY, USA, 2006; Volume 2, pp. 2126–2136.
101. Homaeinezhad, M.R.; Atyabi, S.A.; Tavakkoli, E.; Toosi, H.N.; Ghaffari, A.; Ebrahimpour, R. ECG Arrhythmia Recognition via a Neuro-SVM-KNN Hybrid Classifier with Virtual QRS Image-Based Geometrical Features. *Expert Syst. Appl.* **2012**, *39*, 2047–2058. [[CrossRef](#)]
102. Trstenjak, B.; Mikac, S.; Donko, D. KNN with TF-IDF Based Framework for Text Categorization. *Procedia Eng.* **2014**, *69*, 1356–1364. [[CrossRef](#)]
103. Tan, S. An Effective Refinement Strategy for KNN Text Classifier. *Expert Syst. Appl.* **2006**, *30*, 290–298. [[CrossRef](#)]
104. Adeniyi, D.A.; Wei, Z.; Yongquan, Y. Automated Web Usage Data Mining and Recommendation System Using K-Nearest Neighbor (KNN) Classification Method. *Appl. Comput. Inform.* **2016**, *12*, 90–108. [[CrossRef](#)]
105. Pandey, V.H.R.; Kainthola, A.; Sharma, V.; Srivastav, A.; Jayal, T.; Singh, T.N. Deep Learning Models for Large-Scale Slope Instability Examination in Western Uttarakhand, India. *Environ. Earth Sci.* **2022**, *81*, 487. [[CrossRef](#)]
106. Falae, P.O.; Agarwal, E.; Pain, A.; Dash, R.K.; Kanungo, D.P. A Data Driven Efficient Framework for the Probabilistic Slope Stability Analysis of Pakhi Landslide, Garhwal Himalaya. *J. Earth Syst. Sci.* **2021**, *130*, 167. [[CrossRef](#)]
107. Zhang, W.; Goh, A.T.C. Multivariate Adaptive Regression Splines and Neural Network Models for Prediction of Pile Drivability. *Geosci. Front.* **2016**, *7*, 45–52. [[CrossRef](#)]
108. Friedman, J.H.; Roosen, C.B. An Introduction to Multivariate Adaptive Regression Splines. *Stat. Methods Med. Res.* **1995**, *4*, 197–217. [[CrossRef](#)] [[PubMed](#)]
109. Lee, T.-S.; Chiu, C.-C.; Chou, Y.-C.; Lu, C.-J. Mining the Customer Credit Using Classification and Regression Tree and Multivariate Adaptive Regression Splines. *Comput. Stat. Data Anal.* **2006**, *50*, 1113–1130. [[CrossRef](#)]
110. Lee, T.-S.; Chen, I.-F. A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks and Multivariate Adaptive Regression Splines. *Expert Syst. Appl.* **2005**, *28*, 743–752. [[CrossRef](#)]
111. Leathwick, J.R.; Elith, J.; Hastie, T. Comparative Performance of Generalized Additive Models and Multivariate Adaptive Regression Splines for Statistical Modelling of Species Distributions. *Ecol. Model.* **2006**, *199*, 188–196. [[CrossRef](#)]
112. Lewis, P.A.W.; Stevens, J.G. Nonlinear Modeling of Time Series Using Multivariate Adaptive Regression Splines (MARS). *J. Am. Stat. Assoc.* **1991**, *86*, 864–877. [[CrossRef](#)]
113. Yang, H.-Q.; Yan, Y.; Wei, X.; Shen, Z.; Chen, X. Probabilistic Analysis of Highly Nonlinear Models by Adaptive Sparse Polynomial Chaos: Transient Infiltration in Unsaturated Soil. *Int. J. Comput. Methods* **2023**, *20*, 2350006. [[CrossRef](#)]
114. Guo, X.; Dias, D.; Carvajal, C.; Peyras, L.; Breul, P. Reliability Analysis of Embankment Dam Sliding Stability Using the Sparse Polynomial Chaos Expansion. *Eng. Struct.* **2018**, *174*, 295–307. [[CrossRef](#)]
115. Blatman, G.; Sudret, B. An Adaptive Algorithm to Build up Sparse Polynomial Chaos Expansions for Stochastic Finite Element Analysis. *Probabilistic Eng. Mech.* **2010**, *25*, 183–197. [[CrossRef](#)]
116. Wang, J.; Gu, X.; Huang, T. Using Bayesian Networks in Analyzing Powerful Earthquake Disaster Chains. *Nat. Hazards* **2013**, *68*, 509–527. [[CrossRef](#)]
117. Jensen, F.V.; Nielsen, T.D. *Bayesian Networks and Decision Graphs*; Springer: Berlin/Heidelberg, Germany, 2007; Volume 2.
118. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 26.
119. Montgomery, D.C. *Design and Analysis of Experiments*; John Wiley & Sons: Hoboken, NJ, USA, 2017; ISBN 1119113474.
120. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
121. Xu, H.; He, X.; Sheng, D. Rainfall-Induced Landslides from Initialization to Post-Failure Flows: Stochastic Analysis with Machine Learning. *Mathematics* **2022**, *10*, 4426. [[CrossRef](#)]

122. Phoon, K.-K.; Cao, Z.-J.; Ji, J.; Leung, Y.F.; Najjar, S.; Shuku, T.; Tang, C.; Yin, Z.-Y.; Ikumasa, Y.; Ching, J. Geotechnical Uncertainty, Modeling, and Decision Making. *Soils Found.* **2022**, *62*, 101189.
123. Griffiths, D.V.; Huang, J.; Fenton, G.A. Influence of Spatial Variability on Slope Reliability Using 2-D Random Fields. *J. Geotech. Geoenviron. Eng.* **2009**, *135*, 1367–1378. [[CrossRef](#)]
124. Li, D.-Q.; Jiang, S.-H.; Cao, Z.-J.; Zhou, W.; Zhou, C.-B.; Zhang, L.-M. A Multiple Response-Surface Method for Slope Reliability Analysis Considering Spatial Variability of Soil Properties. *Eng. Geol.* **2015**, *187*, 60–72. [[CrossRef](#)]
125. Tan, X.; Shen, M.; Hou, X.; Li, D.; Hu, N. Response Surface Method of Reliability Analysis and Its Application in Slope Stability Analysis. *Geotech. Geol. Eng.* **2013**, *31*, 1011–1025. [[CrossRef](#)]
126. Jiang, S.-H.; Huang, J.-S. Efficient Slope Reliability Analysis at Low-Probability Levels in Spatially Variable Soils. *Comput. Geotech.* **2016**, *75*, 18–27. [[CrossRef](#)]
127. Li, D.-Q.; Zheng, D.; Cao, Z.-J.; Tang, X.-S.; Phoon, K.-K. Response Surface Methods for Slope Reliability Analysis: Review and Comparison. *Eng. Geol.* **2016**, *203*, 3–14. [[CrossRef](#)]
128. Pontius, R.G.; Thontteh, O.; Chen, H. Components of Information for Multiple Resolution Comparison between Maps That Share a Real Variable. *Environ. Ecol. Stat.* **2008**, *15*, 111–142. [[CrossRef](#)]
129. Shan, F.; He, X.; Xu, H.; Armaghani, D.J.; Sheng, D. Applications of Machine Learning in Mechanised Tunnel Construction: A Systematic Review. *Eng* **2023**, *4*, 1516–1535. [[CrossRef](#)]
130. Frangopol, D.M. *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*; Wiley: Hoboken, NJ, USA, 2008.
131. Jiang, S.-H.; Huang, J.; Griffiths, D.V.; Deng, Z.-P. Advances in Reliability and Risk Analyses of Slopes in Spatially Variable Soils: A State-of-the-Art Review. *Comput. Geotech.* **2022**, *141*, 104498.
132. Fenton, G.A.; Griffiths, D.V. *Risk Assessment in Geotechnical Engineering*; John Wiley & Sons: New York, NY, USA, 2008; Volume 461.
133. Zhang, J.; Ellingwood, B. Orthogonal Series Expansions of Random Fields in Reliability Analysis. *J. Eng. Mech.* **1994**, *120*, 2660–2677. [[CrossRef](#)]
134. Betz, W.; Papaioannou, I.; Straub, D. Numerical Methods for the Discretization of Random Fields by Means of the Karhunen–Loève Expansion. *Comput. Methods Appl. Mech. Eng.* **2014**, *271*, 109–129. [[CrossRef](#)]
135. Zhang, J.; Huang, H.W.; Phoon, K.K. Application of the Kriging-Based Response Surface Method to the System Reliability of Soil Slopes. *J. Geotech. Geoenviron. Eng.* **2013**, *139*, 651–655. [[CrossRef](#)]
136. Meng, J.; Mattsson, H.; Laue, J. Three-dimensional Slope Stability Predictions Using Artificial Neural Networks. *Int. J. Numer. Anal. Methods Geomech.* **2021**, *45*, 1988–2000. [[CrossRef](#)]
137. Song, L.; Yu, X.; Xu, B.; Pang, R.; Zhang, Z. 3D Slope Reliability Analysis Based on the Intelligent Response Surface Methodology. *Bull. Eng. Geol. Environ.* **2021**, *80*, 735–749. [[CrossRef](#)]
138. Kang, F.; Xu, Q.; Li, J. Slope Reliability Analysis Using Surrogate Models via New Support Vector Machines with Swarm Intelligence. *Appl. Math. Model.* **2016**, *40*, 6105–6120. [[CrossRef](#)]
139. Ji, J.; Zhang, C.; Gui, Y.; Lü, Q.; Kodikara, J. New Observations on the Application of LS-SVM in Slope System Reliability Analysis. *J. Comput. Civ. Eng.* **2017**, *31*, 06016002. [[CrossRef](#)]
140. Ahangari Nanehkaran, Y.; Pusatli, T.; Chengyong, J.; Chen, J.; Cemiloglu, A.; Azarafza, M.; Derakhshani, R. Application of Machine Learning Techniques for the Estimation of the Safety Factor in Slope Stability Analysis. *Water* **2022**, *14*, 3743. [[CrossRef](#)]
141. Jiang, S.; Li, J.; Zhang, S.; Gu, Q.; Lu, C.; Liu, H. Landslide Risk Prediction by Using GBRT Algorithm: Application of Artificial Intelligence in Disaster Prevention of Energy Mining. *Process Saf. Environ. Prot.* **2022**, *166*, 384–392. [[CrossRef](#)]
142. Lin, S.; Zheng, H.; Han, C.; Han, B.; Li, W. Evaluation and Prediction of Slope Stability Using Machine Learning Approaches. *Front. Struct. Civ. Eng.* **2021**, *15*, 821–833. [[CrossRef](#)]
143. Suman, S.; Khan, S.Z.; Das, S.K.; Chand, S.K. Slope Stability Analysis Using Artificial Intelligence Techniques. *Nat. Hazards* **2016**, *84*, 727–748. [[CrossRef](#)]
144. Cho, S.E. Probabilistic Stability Analyses of Slopes Using the ANN-Based Response Surface. *Comput. Geotech.* **2009**, *36*, 787–797. [[CrossRef](#)]
145. Ray, A.; Kumar, V.; Kumar, A.; Rai, R.; Khandelwal, M.; Singh, T.N. Stability Prediction of Himalayan Residual Soil Slope Using Artificial Neural Network. *Nat. Hazards* **2020**, *103*, 3523–3540. [[CrossRef](#)]
146. Liu, L.; Zhang, S.; Cheng, Y.-M.; Liang, L. Advanced Reliability Analysis of Slopes in Spatially Variable Soils Using Multivariate Adaptive Regression Splines. *Geosci. Front.* **2019**, *10*, 671–682. [[CrossRef](#)]
147. He, L.; Coggan, J.; Francioni, M.; Eyre, M. Maximizing Impacts of Remote Sensing Surveys in Slope Stability—A Novel Method to Incorporate Discontinuities into Machine Learning Landslide Prediction. *ISPRS Int. J. Geoinf.* **2021**, *10*, 232. [[CrossRef](#)]
148. Zhang, H.; Wu, S.; Zhang, X.; Han, L.; Zhang, Z. Slope Stability Prediction Method Based on the Margin Distance Minimization Selective Ensemble. *Catena* **2022**, *212*, 106055. [[CrossRef](#)]
149. Lin, S.; Zheng, H.; Han, B.; Li, Y.; Han, C.; Li, W. Comparative Performance of Eight Ensemble Learning Approaches for the Development of Models of Slope Stability Prediction. *Acta Geotech.* **2022**, *17*, 1477–1502. [[CrossRef](#)]
150. Azmoon, B.; Biniyaz, A.; Liu, Z.; Sun, Y. Image-Data-Driven Slope Stability Analysis for Preventing Landslides Using Deep Learning. *IEEE Access* **2021**, *9*, 150623–150636. [[CrossRef](#)]

151. Yuan, C.; Moayedi, H. The Performance of Six Neural-Evolutionary Classification Techniques Combined with Multi-Layer Perception in Two-Layered Cohesive Slope Stability Analysis and Failure Recognition. *Eng. Comput.* **2020**, *36*, 1705–1714. [[CrossRef](#)]
152. Huang, S.; Huang, M.; Lyu, Y. An Improved KNN-Based Slope Stability Prediction Model. *Adv. Civ. Eng.* **2020**, *2020*, 8894109. [[CrossRef](#)]
153. Ospina-Dávila, Y.M.; Orozco-Alzate, M. Parsimonious Design of Pattern Recognition Systems for Slope Stability Analysis. *Earth Sci. Inform.* **2020**, *13*, 523–536. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.