Revista
GEOCIÊNCIAS

unesp

ISSN: 1980-900X (online)

MACHINE LEARNING ALGORITHMS APPLIED TO THE PREDICTION OF SLOPE FAILURES IN EARTH DAMS TRIGGERED BY RAINFALL

ALGORITMOS DE APRENDIZADO DE MÁQUINA APLICADOS À PREVISÃO DE FALHAS EM TALUDES DE BARRAGENS DE TERRA PROVOCADAS POR CHUVAS

Isaida Flores Berenguer[1], Adriana Ortega Rubio[2], Alejandro Rosete[2], Yoermes González Haramboure[2], Jenny García Tristá[2]

[1]Ecoturismo Mundial y Proyectos Ambientales (EMPACA S.R.L). Calle Wenceslao Álvarez 60, Distrito Nacional, santo Domingo, Republica Dominicana. E-mail: isiflores92@gmail.com

[2]Universidad Tecnológica de La Habana "José Antonio Echeverría" (CUJAE). Facultad de Ingeniería Civil. Calle # 50 entre Avenida 25 y Avenida 27. Playa, La Habana. Cuba. E-mail: isiflores92@gmail.com; yoermes@civil.cujae.edu.cu; alejandro.rosete.suarez@gmail.com; jenny@civil.cujae.edu.cu

RESUMO - A estabilidade dos taludes em barragens de terra é uma questão de importância global. A precipitação é um fator crucial nessa análise, pois a chuva é um dos principais fatores que provocam deslizamentos de terra. Os recentes avanços no sistema computacional que integram modelos numéricos com algoritmos de aprendizado de máquina oferecem uma abordagem robusta para previsões precisas. Esta pesquisa emprega sete algoritmos amplamente utilizados para prever a estabilidade de taludes: redes neurais artificiais, máquinas de vetores de suporte, árvores de decisão, florestas aleatórias, vizinhos mais próximos, regressão logística e Naive Bayes. Essa pesquisa utiliza um conjunto de dados com 5.637 casos de barragens de terra gerados por meio de métodos numéricos acoplados, que consideram a infiltração transiente e o impacto da precipitação na estabilidade do talude. O objetivo é desenvolver modelos de previsão com alta precisão. Os resultados indicam que o melhor caso de resposta geral é o de k-vizinhos mais próximos, com uma precisão de 0,983 e um erro de 0,017, seguido pelas redes neurais artificiais, com uma precisão de 0,966 e um erro de 0,034. Os resultados mais desfavoráveis foram obtidos para o Naive Bayes, com precisão de 0,852 e erro de 0,148, seguido pelas máquinas de vetor de suporte, com precisão de 0,863 e erro de 0,137. Além disso, três inclinações do mundo real com características semelhantes às propostas no conjunto de dados foram selecionadas para validar os resultados individuais obtidos com cada um dos setes algoritmos. A partir dessa análise, foi detectado que a rede neural artificial, as máquinas de vetor de suporte e as florestas aleatórias foram as únicas que previram com precisão a resposta esperada na validação.
Palavras-chave: Falha de talude. Precipitação. Aprendizado de máquina. Precisão. Erro. Previsão de deslizamento de terra.

ABSTRACT - The stability of slopes in earth dams is a matter of global significance. Precipitation is a crucial factor in this analysis, as rainfall is a major trigger for landslides. Recent advancements in computational systems that integrate numerical models with machine learning algorithms offer a robust approach for accurate predictions. This investigation employs seven widely used algorithms for predicting slope stability: artificial neural networks, support vector machines, decision trees, random forests, k-nearest neighbors, logistic regression, and Naive Bayes. This research utilizes a dataset with 5637 cases of earth dams generated through coupled numerical methods, which consider transient seepage and the impact of precipitation on slope stability. The goal is to develop prediction models with high accuracy. The results indicate that the best overall response case is k-nearest neighbors with an accuracy of 0.983 and an error of 0.017; followed by artificial neural networks with an accuracy of 0.966 and an error of 0.034. The most unfavorable results were obtained for Naive Bayes with accuracy of 0.852 and an error of 0.148; followed by support vector machines with an accuracy of 0.863 and an error of 0.137. In addition, three real-worlds slopes with similar characteristics to those proposed in the dataset were selected to validate the individual results obtained with each of the seven algorithms. From this analysis it was detected that artificial neural network, support vector machines, and random forests were the only ones that accurately predicted the expected response in the validation.
Keywords: Slope failure. Precipitation. Machine learning. Accuracy. Error. Landslide prediction.

INTRODUCTION

In recent years, the use of Artificial Intelligence algorithms in Civil Engineering has been increasing, particularly in geotechnical problems (Xu et al., 2023). One of the most studied structures in this field is earth dams, which have numerous applications for meeting the needs of the population and industry (Flores et al., 2022).

An essential aspect that must be considered regarding slopes is the analysis of their stability, which refers to the safety of a soil mass against failure or movement (Huang et al., 2023; Nanehkaran et al., 2023; Wang et al., 2023). Slope failure is the movement of instability involving a significant amount of material due to a lack of support at the base of the slope, primarily caused by shear failure along one or more failure surfaces (Flores et al., 2022).

According to Nanehkaran et al. (2023) and Xu et al. (2023), landslides are among the most destructive geological processes affecting humans, representing sudden and severe disasters that cause thousands of deaths and substantial economic losses (He et al., 2020). They can lead to changes in terrain morphology, various environmental damages, impact infrastructure, destroy homes, bridges, and block rivers. Predicting landslides constitutes one of the most active areas in geotechnical studies (Xu et al., 2023).

It is common to define the stability of a slope in terms of a Factor of Safety (FoS), obtained from a deterministic mathematical analysis, where models must consider most factors affecting stability (Flores et al., 2024). The deterministic method involves minimizing the FoS over a range of potential failure surfaces, determining the surface with the minimum FoS, known as the critical slip surface (Fernández et al., 2018). The FoS is also defined as the factor by which the soil's shear strength must be reduced to bring a potentially unstable mass to a limit equilibrium along a preselected slip surface (Xu et al., 2023). The FoS value is obtained by comparing the shear stress at failure with the soil's shear strength and must always be greater than or equal to the value set by standards or project requirements (Flores et al., 2022).

Precipitation is one of the primary causes that can trigger slope failures (Tsang et al., 2024) and is the main factor leading to landslides (Wengang et al., 2023). During rainfall, water infiltrates the soil, increasing soil moisture content, reducing matric suction, and, upon reaching saturation, raising pore water pressure, thereby decreasing soil shear strength. This results in increased vulnerability of slopes to failures, collapses, and flows (Flores et al., 2022; Kim et al., 2017; Sidle & Bogaard, 2016).

When heavy rain affects slopes, soil hydraulic properties and rainfall patterns control water infiltration behavior. The soil's saturated hydraulic conductivity limits the range of precipitation infiltration, while rainfall intensity and duration control the amount that can infiltrate the soil (Flores et al., 2022). The individual effect of precipitation and fluctuating reservoir water levels depends on various factors, such as precipitation intensity, soil properties, geological structure, and the groundwater level (Kafle et al., 2022). In areas with heavy rainfall and under certain soil conditions, precipitation can have very severe effects.

Flores et al. (2022) and Huat et al. (2006) suggested that seasonal precipitation is the cause of most slope failures, primarily due to the loss of soil matric suction from rainwater infiltration. They also demonstrated that higher rainfall intensity leads to a higher infiltration rate in the soil, thus reducing the FoS.

In recent decades, machine learning methods have gained increasing attention from researchers for their utility as surrogate models to improve computational efficiency in slope stability studies. The fundamental approaches of these methods in slope stability focus on predicting the FoS and its classification (Xu et al., 2023). Compared to traditional probabilistic or analytical models, machine learning methods offer advantages such as high accuracy and not requiring any initial assumptions between inputs and outputs. They can also handle large datasets and incomplete data. By using these methods, geotechnical engineers can avoid limitations of traditional techniques, such as the need for prior assumptions or high computational requirements, and accurately predict slope stability (Xu et al., 2023; Flores et al., 2024). By linking the strategic importance of earth dams and their susceptibility to landslides due to precipitation effects, predicting these phenomena becomes crucial.

This research employs various machine learning algorithms to predict the effect of precipitation on the downstream slope stability of earth dams. A dataset constructed using hybrid modelling, combining results from Flores et al. (2023; 2024), including various rainfall intensities, is used. The algorithms employed include artificial neural networks, support vector machines, decision trees, random forests, k-nearest neighbors, and Naive Bayes. The results obtained allow for a comparison between all algorithms and a contrast with literature reports regarding accuracy, error, and Co-hen's Kappa coefficient.

METHODOLOGY

Here is a succinct overview of the AI techniques that are being used in this study.

Artificial Neural Networks (ANN) are commonly used machine learning systems that mimic the functioning of the human brain (Aminpour et al., 2023; Tablada & Torres, 2006). This algorithm aims to adjust the model parameters so that the training results for each input closely match the output. The optimal number of layers and hidden neurons in an ANN depends on the complexity of the problem (Flores et al., 2024). While its effectiveness is high, the mathematical analyses developed within the network make the system complex to interpret, inaccessible, and non-modifiable, which has led to a field of active research known as Explainable Artificial Intelligence (Barredo et al., 2020).

Support Vector Machines (SVM) originate from statistical studies (Nanehkaran et al., 2023). They are supervised machine learning algorithms applicable to classification problems (binary or multiclass) (Kwag et al., 2020; Salamanca, 2021). SVM are highly effective, but their interpretation is highly complex, making them difficult to understand and modify. For implementation, observations of a sample are separated into broad spaces by a hyperplane or a set of hyperplanes, also called linear separators, according to their class. In most real cases, perfect linear separation of data is not possible. When examples are not linearly separable, non-linear base function sets are efficiently used to define high-dimensional transformed spaces and find optimal separation hyperplanes in feature spaces. SVM belongs to a class of machine learning algorithms called kernel methods and are also known as kernel machines (Salamanca, 2021).

Decision Trees (DT) are a non-parametric supervised learning algorithm known for their ease of interpretation and computational efficiency (Alaminos, 2023). They can be used for both classification and regression tasks (Wengang et al., 2023). They are represented as a tree structure similar to a flowchart, consisting of different parts: a root node, intermediate nodes, leaf nodes, and branches connecting the nodes (Arana, 2021). Each path from the root node to the output node represents a decision rule showing the relationship between input and output variables (Jong et al., 2021). Pruning methods can be incorporated to remove lower branches that do not significantly contribute to accuracy, simplifying the model and improving generalization and understanding (Flores et al., 2024). While DT offer an easy-to-interpret structure governed by understandable rules, they are not as precise as other data mining techniques and are susceptible to data with excessive variation or missing information (Jong et al., 2021).

A Random Forest (RF) consists of a defined number of DT (Medina & Ñique, 2017). It is an intelligent classification algorithm based on the statistical theory proposed by Breiman (2001). Random forests are a widely used technique due to their high accuracy (Huang et al., 2023), simplicity, and flexibility, making them suitable for both classification and regression (Aminpour et al., 2023; Wang et al., 2023; Xu et al., 2023). It is a very fast decision method that works with large datasets. Essentially, RFs are a collection of DTs and thus more accurate than individual DT. As the number of DT in an RF increases, the generalization error decreases and stabilizes at a limiting value. However, a high number of trees makes the model more complex and harder to understand.

Logistic Regression (LR) is a classification and regression algorithm used to predict the probability of a categorical dependent variable given one or multiple independent variables (Ranjitha et al., 2020). It is a multivariate analysis method most useful when there is a dichotomous dependent variable (taking only one of two mutually exclusive values) and a set of predictor variables, allowing estimation of the probability of an event occurring considering other variables (Martínez et al., 2020).

The k-Nearest Neighbours (KNN) method is a basic supervised machine learning algorithm primarily used for classification and prediction of both categorical and numerical variables with a non-parametric nature (Narváez et al., 2022). KNN assumes that similar data points are located near each other. Therefore, it calculates the distance between data points, typically using Euclidean distance, and then assigns a category based on the most frequent category or the average among the closest points (Toribio et al., 2024). The method does not require a training process, significantly reducing computational costs and time.

Naive Bayes (NB) is a simple probabilistic classifier based on conditional probability. The algorithm uses Bayes' Theorem and assumes that all variables are independent given the class variable. Its widespread use is due not only to its simplicity but also to the effectiveness and robustness of the algorithm. It is used for real-time prediction because it is very fast in the training process (López et al., 2023; Wickramasinghe & Kalutarage, 2021).

To apply the machine learning algorithms, a dataset with 5637 examples established by Flores (2024) and used by Flores et al. (2024) is employed. Developed models for earth dams of 15, 30, and 40 meters-height, with embankments consisting of 40 clayey soils were obtained using the Monte Carlo method derived from actual data on clayey soils that are components of constructed earth dams (Flores et al., 2023). All analyses used coupled numerical models (Flores et al., 2023), considering the soil in a partially saturated state and including three rainfall intensities: 3 mm/day, 50 mm/day, and 150 mm/day.

Additionally, the effect of soil infiltration in the earth dam slopes under the given conditions was taken into account. For each of these 5637 examples of conditions, the FoS were obtained by using a numerical method. This dataset is used to train the AI models in order to predict the FoS given the set of different values of the variables in each situation.

In terms of FoS, this research considered it as a categorical, not numerical, variable. According to Armas & Horta (1987); Flores et al (2024); and Mucuta et al. (2020), the relationship between FoS value and slope stability is defined as follows: if FoS < 1, the slope is in failure; if $1 \leq$ FoS < 1.5, the slope is considered stable; and if FoS ≥ 1.5, the slope is considered stable and safe.

Various computational programs can be used to apply these algorithms to stability problems, e.g. the KNIME (Konstanz Information Miner) program, which allows model development in a visual environment (Flores et al., 2024). KNIME is designed as a graphical tool and features nodes (encapsulating different types of algorithms) and arrows (representing data flow). In this research, the KNIME program is used to apply the selected algorithms due to the positive results obtained in similar studies (Flores et al., 2023; Flores et al., 2024).

RESULTS AND DISCUSSION

The results obtained in the numerical modeling are grouped considering all variables associated with it: geometric parameters (embankment high, crown width), soil data (specific weight, cohesion, internal friction angle, saturated permeability, volumetric water content, volumetric compressibility index, effective diameter for 60 %, effective diameter for 10 %, liquid limit), external loading conditions (precipitation intensities: 3 mm/day; 50 mm/day and 150 mm/day), and 24 hours duration of the phenomenon. In all these experiments, the dataset is randomly divided in a training set with 70% of the dataset and a test set with the other 30%, in order to train and to evaluate the quality of AI models obtained.

Artificial Neural Networks (ANN)

In KNIME, the default configuration for the ANN node includes one hidden layer (1HL), ten neurons (10N), and 100 iterations (100It). However, calibration is necessary. Finally, a configuration with one hidden layer consisting of 10 neurons and 10,000 iterations is selected. This configuration yields an accuracy of 0.966 and an error of 0.034. Table 1 presents the final results of the analysis.

Table 1 - Results of ANN Prediction vs. Expected Value

ANN Prediction FoS \ Expected FoS	Stable and safe	Stable	Failure
Stable and safe	1527	47	11
Stable	11	710	19
Failure	52	51	3209

Table 1 shows all the results obtained by applying the ANN model. As for the case classified as Stable and safe, ANN submitted 63 incorrect proposals, of which 17.5 % were predicted as Stable and the remaining 82.5 % were predicted as Failure. For the case classified as Stable, ANN submitted 98 incorrect proposals, of which 48.0 % were predicted as Stable and Safe and the remaining 52.0 % were predicted as Failure. For the case classified as Failure, ANN submitted 30 incorrect proposals, of which 36.7 % were predicted as Stable and safe and the remaining 63.3 % were predicted as Stable.

The results show that the predictions related to Failure have the lowest number of incorrect values proposed. Moreover, for both Stable and safe and Stable results, most incorrect predictions propose Failure. These results allow us to place the predictions on the safe side.

Support Vector Machines (SVM)

Table 2 presents the final results of the analysis using SVM.

Table 2 - Results of SVM Prediction vs. Expected Value.

SVM Prediction FoS \ Expected FoS	Stable and safe	Stable	Failure
Stable and safe	1190	72	195
Stable	180	630	190
Failure	108	27	3045

Table 2 shows all the results obtained by applying the SVM model. As for the case classified as Stable and safe, SVM submitted 288 incorrect proposals, of which 62.5 % were predicted as Stable and the remaining 37.5 % were predicted as Failure.

For the case classified as Stable, SVM submitted 99 incorrect proposals, of which 72.7 % were predicted as Stable and Safe and the remaining 27.3 % were predicted as Failure. For the case classified as Failure, SVM submitted 385 incorrect proposals, of which 50.6 % were predicted as Stable and safe and the remaining 49.4 % were pre-dicted as Stable.

The results show a remarkable difference with respect to those obtained with ANN. In this case there are many more wrong predictions and, contrary to the previous case, Failure is not proposed as the predominant case. This reduces the confidence in the predictions and increase the uncertainty with the SVM model.

Decision Tree (DT)

Table 3 presents the final results of the analysis using DT.

Table 3 - Results of DT Prediction vs. Expected Value.

DT Prediction FoS \ Expected FoS	Stable and safe	Stable	Failure
Stable and safe	1405	41	136
Stable	57	720	97
Failure	5	3	3173

Table 3 shows all the results obtained by applying the DT model. As for the case classified as Stable and safe, DT submitted 62 incorrect proposals, of which 91.9 % were predicted as Stable and the remaining 8.1 % were predicted as Failure. For the case classified as Stable, DT submitted 44 incorrect proposals, of which 93.2 % were predicted as Stable and Safe and the remaining 6.8 % were predicted as Failure. For the case classified as Failure, DT submitted 233 incorrect proposals, of which 58.4 % were predicted as Stable and safe and the remaining 41.6 % were predicted as Stable.

The results show that most of the errors in the prediction are obtained for the Failure cases. Furthermore, in the Stable and safe and Stable cases, the prediction moved away from security by proposing fewer Failures and prioritizing slope safety in all cases.

In this case, the initial node of the tree is the rainfall intensity, which branches into internal friction angle and duration. This means that these are the most relevant variables defining the subsequent analyses. The DT analysis identified fourteen conditions leading to failure. According to these rules, the probability of failure increases for cases with intensities greater than 100 mm/day. Additionally, in all these Failure rules, the duration of rainfall exceeds 5.5 hours. It is noteworthy that this intensity is considered heavy rain for conditions in Cuba (Durán, 2016). The analysis of these rules shows one of the fundamental advantages of DTs: they are easily interpretable.

For cases with rainfall intensity less than or equal to 100 mm/day, the minimum continuous rainfall duration is 15.5 hours. For rainfall intensities below 100 mm/day that trigger landslides, the duration time is well over 15.5 hours.

Random Forest (RF)

Table 4 shows the final results of the analysis using RF. Table 4 shows all the results obtained by applying the DT model. As for the case classified

Table 4 - Results of RF Prediction vs. Expected Value.

RF Prediction FoS \ Expected FoS	Stable and safe	Stable	Failure
Stable and safe	1509	9	46
Stable	23	798	62
Failure	0	2	3194

as Stable and safe, RF submitted 23 incorrect proposals, of which total were predicted as Stable and none is proposed as Failure. For the case classified as Stable, RF submitted 11 incorrect proposals, of which 81.8 % were predicted as Stable and Safe and the remaining 18.2 % were predicted as Failure. For the case classified as Failure, RF submitted 108 incorrect proposals, of which 42.6 % were predicted as Stable and safe and the remaining 57.4 % were predicted as Stable. The results show the lowest number of errors in the proposals for FoS Stable and safe and Stable, but not for the Failure cases.

Logistic Regression (LR)

Table 5 presents the final results of the analysis using LR.

Table 5 - Results of LR Prediction vs. Expected Value.

LR Prediction FoS \ Expected FoS	Stable and safe	Stable	Failure
Stable and safe	1319	87	197
Stable	60	720	119
Failure	77	21	3037

Table 5 shows all the results obtained by applying the LR model. As for the case classified as Stable and safe, LR submitted 137 incorrect proposals, of which 43.8 % were predicted as Stable and the remaining 56.2 % were predicted as Failure. For the case classified as Stable, LR submitted 108 incorrect proposals, of which 80.6 % were predicted as Stable and Safe and the remaining 19.4 % were predicted as Failure. For the case classified as Failure, LR submitted 316 incorrect proposals, of which 62.3 % were predicted as Stable and safe and the remaining 37.7 % were predicted as Stable.

In this case there are many more wrong predictions and Failure is not proposed as the predominant case. This reduces the confidence in the predictions and increase the uncertainty with the LR model.

K-Nearest Neighbours (KNN)

For employing the KNN algorithm, it is necessary to conduct a calibration process. In this case, the configuration corresponding to 3 nearest neighbors is selected, which yields an accuracy of 0.983 and an error of 0.017. In addition, when the number of KNN increases, the error also rises while accuracy decreases. Table 6 presents the final results of the analysis using KNN.

Table 6. Results between prediction and expected value with KNN.

KNN Prediction FoS \ Expected FoS	Stable and safe	Stable	Failure
Stable and safe	1482	19	0
Stable	6	836	0
Failure	19	53	3222

Table 6 shows all the results obtained by applying the KNN model. As for the case classified as Stable and safe, KNN submitted 25 incorrect proposals, of which 24.0 % were predicted as Stable and the remaining 76.0 % were predicted as Failure. For the case classified as Stable, KNN submitted 72 incorrect proposals, of which 26.4 % were predicted as Stable and Safe and the remaining 76.3 % were predicted as Failure.

For the case classified as Failure, KNN does not make mistakes in prediction.

In this case, the most favorable results of those obtained so far is shown. In the training, KNN models does not make errors in the prediction of the Failure cases. In the Stable and safe and Stable cases, most of the erroneously predicted cases are proposed as Failure as well, placing the results on the security side.

Naive Bayes (NB)

Table 7 shows the final results of the analysis using NB.

Table 7 - Results between prediction and expected value with NB.

NB Prediction FoS \ Expected FoS	Stable and safe	Stable	Failure
Stable and safe	1196	72	253
Stable	228	555	117
Failure	131	31	3054

Table 7 shows all the results obtained by applying the NB model. As for the case classified as Stable and safe, NB submitted 359 incorrect proposals, of which 63.5 % were predicted as Stable and the remaining 36.5 % were predicted as Failure. For the case classified as Stable, NB submitted 108 incorrect proposals, of which 66.7 % were predicted as Stable and Safe and the remaining 33.3 % were predicted as Failure. For the case classified as Failure, NB submitted 364 incorrect proposals, of which 69.5 % were predicted as Stable and safe and the remaining 30.5 % were predicted as Stable.

In this case, the most unfavorable results of those obtained so far is shown. The total number of erroneously predicted cases exceeds that of all previous models. Moreover, in this respect, the predictions proposed as Failure are lower than those proposed as Stable and safe and Stable, which reduces the overall security of the model.

General Comparison

Figure 1 presents the general results regarding the number of correctly and incorrectly classified cases.

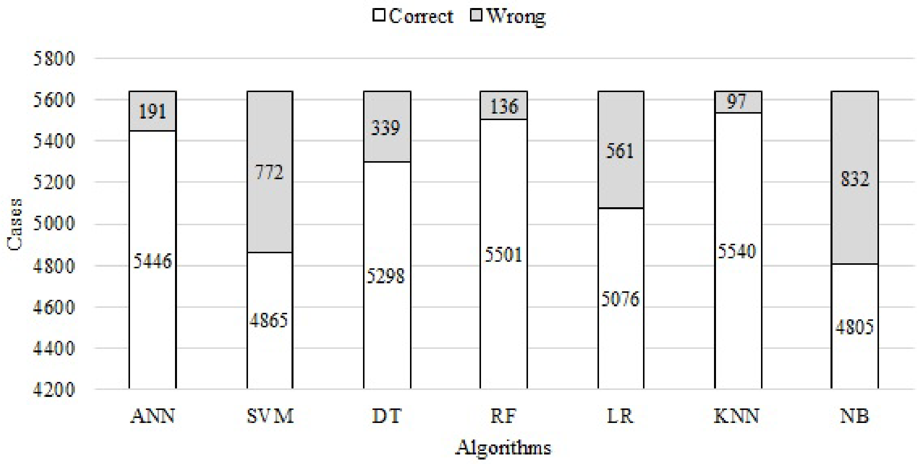

Figure 1 - General results for correctly and incorrectly classified cases for each algorithm.

Figure 1 shows that the NB algorithm has the fewest correctly classified cases (4805) and the highest number of incorrect cases (832), followed by SVM and LR. Conversely, the KNN algorithm has the highest number of correct cases (5540) and the fewest incorrect cases (97), followed by RF and ANN.

The results for accuracy, error, and Cohen's kappa coefficient are displayed in Figure 2.

Figure 2 reveals that all employed algorithms present accuracies above 0.8 for the established configurations.

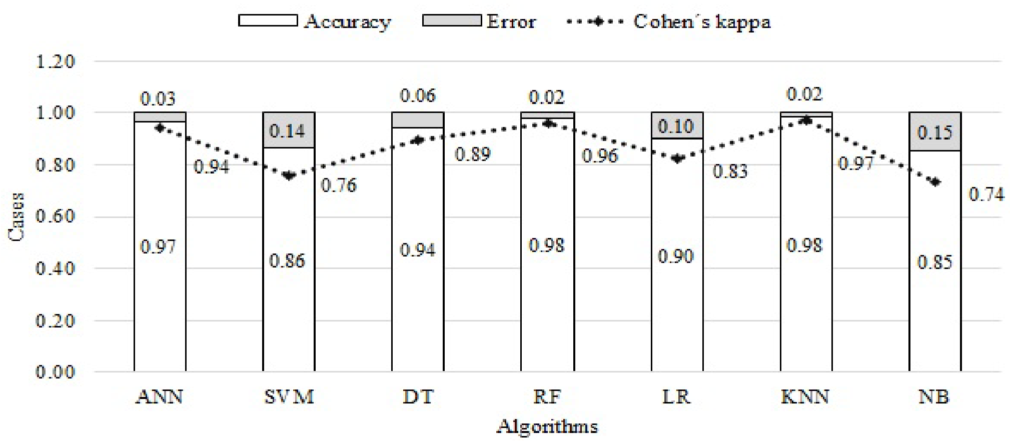

Figure 2 - General results for accuracy, error, and Cohen's kappa coefficient for each algorithm.

However, only ANN, RF, and KNN show accuracies exceeding 0.95.
The algorithm with the highest accuracy is KNN, while the one with the lowest accuracy is NB. Regarding errors, all cases are below 0.15, with ANN, RF, and KNN achieving errors below 0.05. The algorithm providing the best results for error is KNN, and the one with the highest error is NB. For Cohen's kappa coefficient, ANN, RF, and KNN display kappa values above 0.9, indicating a perfect fit between predicted and expected results (Lin et al., 2018). The algorithm with the best kappa value is KNN, while the lowest fit is shown by NB

Validation Of Prediction Models

Based on the obtained results, the trained models were employed for model validation. This involves three specific real-world cases: the slope of the La Cidra dam, which experienced a landslide in May 2014, referred to as Slope 1 for these purposes; and two cases from research by Liu & Wang (2023), called Slope 2 and Slope 3. The three cases proposed respond to slopes built with a single material, in partially saturated conditions and including rainfall as an external load or varying intensity and duration. In all three cases, slope failure occurs, which is considered the correct condition. The prediction results are shown in Figure 3. Figure 3 demonstrates that while the obtained prediction models exhibit high accuracy (greater than 0.85) and errors below 0.15 in all cases, they are not infallible in their evaluation. This implies that it is crucial to consider the inherent uncertainty in each predicted case. Although these models are effective for rapid analysis, they do not eliminate the need for numerical modelling, in certain cases that warrant it, or when there is time available to do so. Nevertheless, ANN, SVM, and RF models correctly predicted the outcomes, aligning with the criteria set by Xu et al. (2023), which reflects that these are the most internationally used algorithms for solving slope stability issues.

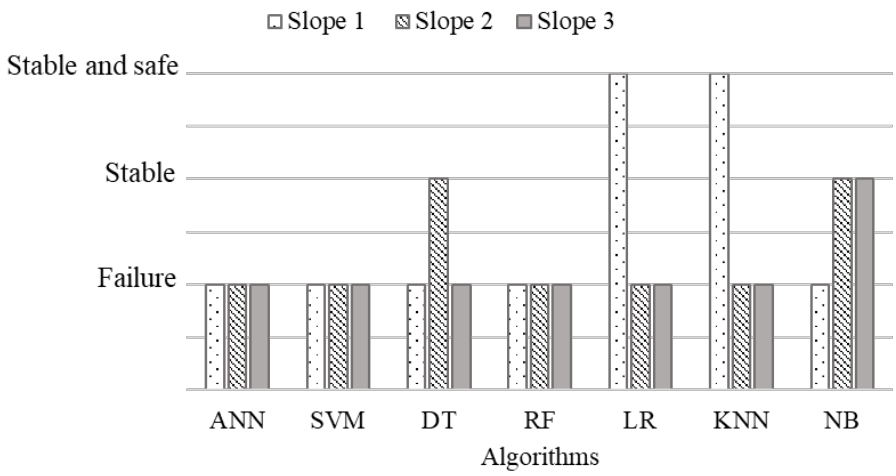

Figure 3 - Results of the validation process of the employed algorithms.

CONCLUSIONS

In recent years, there has been a growing trend in combining AI models with traditional numerical modelling for predicting slope stability. Although there is no international consensus on the most appropriate prediction methods, commonly used approaches include ANN, SVM, DT, RF, LR, KNN, and NB. This research trained these seven algorithms to predict the stability of earth dams subjected to precipitation. In all cases, the accuracy achieved was above 0.80 and the error was below 0.15.

In this research the algorithm with the highest accuracy and lowest error is KNN, while the one with the lowest accuracy and major error is NB. For Cohen's kappa coefficient, ANN, RF, and KNN display kappa values above 0.9, indicating a perfect fit between predicted and expected results.

Validation of the prediction models was carried out using three real-world cases re-ported in the literature. The results indicated that DT, LR, KNN, and NB models made incorrect predictions in at least one case, whereas ANN, SVM, and RF correctly predicted all cases. This demonstrates that, regardless of the accuracy achieved during training, algorithms are subject to errors that, even if small, can influence the prediction outcome in particular situations.

REFERENCES

ALAMINOS, A. **Árboles de decisión en R con Random Forest.** Editorial: Limencop, S.L, 2023. ISBN 978-84-09-49460-6

AMINPOUR, M.; ALAIE, R.; KHOSRAVI, S.; KARDANI, N.; MORIDPOUR, S. Slope stability machine learning predictions on spatially variable random fields with and without factor of safety calculations. **Computers and Geotechnics**, v. 153, p. 2–16, 2023. https://doi.org/10.1016/j.compgeo.2022.105094

ARANA, C. Modelos de aprendizaje automático mediante árboles de decisión. No. 778. Serie Documentos de Trabajo. 2021. https://doi.org/https://hdl.handle.net/10419/238403

Armas, R. & Horta, E. **Presas de Tierra**. Ingeniería Hidráulica. Instituto Superior Politécnico José Antonio Echeverría (Cujae), 1987. La Habana, Cuba.

BARREDO, A.; DÍAZ, N.; DEL, J.; BENNETOT, A.; TABIK, S.; BARBADO, A.; GARCÍA, S.; GIL-LÓPEZ, S.; MOLINA, D.; BENJAMINS, R.; CHATILA, R.; HERRERA, F. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. **Information Fusion**, v. 58 (October 2019), p. 82–115, 2020. https://doi.org/10.1016/j.inffus.2019.12.012

BREIMAN, L. Random Forests. **Machine Learning**, v. 45, p. 5–32, 2001. http://dx.doi.org/10.1023/A:1010933404324

DURÁN, I. Cantidad de días con lluvia y su distribución por intervalos en condiciones normales y de sequía severa en el occidente de Cuba. **Revista Cubana de Meteorología**, v. 22, n. 1, p. 49–65, 2016.

FERNÁNDEZ, W.; VILLALOBOS, S.; KING, R. Probabilistic analysis of slope stability in completely decomposed granite residual soils. Revista Ingeniería de Construcción, v. 33, n. 1, p. 5–14, 2018. ISSN 0718-5073. http://dx.doi.org/10.4067/S0718-50732018000100005.

FLORES, I.; GONZÁLEZ, Y.; GARCÍA, J. **Consideraciones para el análisis de la estabilidad de taludes en presas de tierra con suelos parcialmente saturados**. 2022. ISBN: 978-959-281-616-5.

FLORES, I.; GONZÁLEZ, Y.; GARCÍA, J.; ROSETE, A. Artificial Neural Network Model to Predict the Factor of Safety in Earth Dams Subjected to Rapid Drawdown. **Revista Politécnica,** v. 54, n. 1, 2024. https://doi.org/10.33333/rp.vol54n1.04.

FLORES, I.; HERNÁNDEZ, F.; GONZÁLEZ, Y.; GARCÍA, J. Relationship between soil physical-mechanical parameters and earth dam slope stability using stochastic and numerical modeling. **Revista Geociencias**, v. 42, n. 2, p. 387–402, 2023. https://doi.org/https://doi.org/10.5016/geociencias.v42i3.17683

HE, X.; XU, H.; SABETAMAL, H.; SHENG, D.. Machine learning aided stochastic relia-bility analysis of spatially variable slopes. **Computers and Geotechnics**, 103711, 2020 https://doi.org/10.1016/j.compgeo.2020.103711

HUANG, F.; XIONG, H.; CHEN, S.; LV, Z.; HUANG, J.; CHANG, Z.; CATANI, F. Slope stability prediction based on a long short - term memory neural network: comparisons with convolutional neural networks, support vector machines and random forest models. **International Journal of Coal Science & Technology**, v. 4, p. 4–10, 2023. https://doi.org/ 10.1007/s40789-023-00579-4

HUAT, B.; ALI, F.; LOW, T. Water infiltration characteristics of unsaturated soil slope and its effect on suction and stability. **Geotechnical and Geological Engineerin*g***, v. 24, p. 1293–1306, 2006. https://doi.org/10.1007/s10706-005-1881-8

JONG, S.; ONG, D.; OH, E. (). State-of-the-art review of geotechnical-driven artificial intelligence techniques in underground soil-structure interaction. **Tunnelling and Underground Space Technology,** v. 113 (April), p. 3–16, 2021. https://doi.org/10.1016/j.tust.2021.103946

KAFLE, L.; XU, W.; ZENG, S.; NAGEL, T. A numerical investigation of slope stability influenced by the combined effects of reservoir water level fluctuations and precipitation: A case study of the Bianjiazhai landslide in China. **Engineering Geology**, v. 297 (December 2021), p. 1–10. 2022, https://doi.org/https://doi.org/10.1016/j.enggeo.2021.106508

KIM, J.; KIM, Y.; SANGSEOM, J.; HONG, M. Rainfall - induced landslides by deficit field matric suction in unsaturated soil slopes. **Environmental Earth Science*s***, v. 76, n. 23, p. 3–28, 2017. https://doi.org/10.1007/s12665-017-7127-2

KWAG, S.; HAHM, D.; KIM, M.; EEM, S. Development of a Probabilistic Seismic Performance Assessment Model of Slope Using Machine Learning Methods. **Sustainability**, v. 12, n. 3269, p. 2–19, 2020. https://doi.org/10.3390/su12083269

LIN, Y.; ZHOOU, K.; LI, J. Prediction of Slope Stability Using Four Supervised Learning Methods. **IEEE Access**, v. 6, p. 31169–31179, 2018. https://doi.org/10.1109/ACCESS.2018.2843787

LIU, X. & WANG, Y. Probabilistic simulation of entire process of rainfall-induced landslide using random finite element and material point methods with hydromechanical coupling. **Computers and Geotechnics**, v. 12, 2021. https://doi.org/https://doi.org/10.1016/j.compgeo

LÓPEZ, J. Herramientas digitales y uso de inteligencia artificial en geotecnia: Un enfoque en la evaluación de taludes con Matlab. **Revista Latinoamericana de Ciencias Sociales y Humanidades**, v. 5, n. 1, p. 880–893. 2024. https://doi.org/https://doi.org/10.56712/latam.v5i1.1640

MARTÍNEZ, J.; FERRÁS, Y.; BERMÚDEZ, L.; ORTÍZ, Y. Regresión logística y predicción del bajo rendimiento académico de estudiantes en la carrera Medicina. **Revista Electrónica**, v. 45, n. 4, 2020. https://revzoilomarinello.sld.cu/index.php/zmv/article/view/2230

MEDINA, R. & ÑIQUE, C. Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python. **Interfases**, p. 165–190, 2017. https://doi.org/10.26439/interfases 2017. n10.1775

MUCUTA, H.; CARTAYA, M.; WATSON, R. L. Evaluación de estabilidad en taludes del yacimiento Castellano mediante el cálculo del factor de seguridad. *Minería y Geología*, v. 36, n. 4, p. 441–450, 2020. https://revista.ismm.edu.cu/index.php/revistamg/article/view/art6_No4_2020

NANEHKARAN, Y.; LICAI, Z.; CHENGYONG, J.; CHEN, J.; ANWAR, S.; AZARAFZA, M.; DE-RAKHSHANI, R. (). Comparative Analysis for Slope Stability by Using Machine Learning Methods. **Applied Sciences,** v. 13, n. 1555, p. 1–14, 2023 https://doi.org/https://doi.org/10.3390/app13031555

NARVÁEZ, M.; MOYANO, R.; PALACIOS, D.; IZURIETA, G. Predicción de clientes potenciales utilizando K vecino más cercano en el área de negocios de la Cooperativa Riobamba. **Revista a perspectivas**, v. 4, n. 1, p. 22–25, 2022. https://doi.org/https://doi.org/10.47187/perspectivas.4.1.150

RANJITHA, S.; MERUGUMALA, G.; PADIGALA, S.; VANGA, R.; KANTAPALLI, B. Employee Attrition Prediction using Logistic Regression. **International Journal for Research in Applied Science and Engineering Technology**, v. 8, n. V, p. 2872–2874. 2020. https://doi.org/http://doi.org/10.22214/ijraset.2020.5481

SALAMANCA, Y. **Análisis metalográfico en MatLab a través de la herramienta SVM con enfoque estadístico.** Bogotá. 2021. Tesis de Doctorado. Universidad Distrital Francisco José de Caldas, D.C. http://hdl.handle.net/11349/28745

SIDLE, R. C. & BOGAARD, T. A. Dynamic earth system and ecological controls of rainfall-initiated landslides. **Earth-Science Reviews**, v. 159, p. 275–291, 2016. https://doi.org/http://dx.doi.org/10.1016/j.earscirev.2016.05.013

TABLADA, C. & TORRES, G. Redes Neuronales Artificiales. *Inteligencia Artificial*, v. 1, n. 1, p. 22–30. 2006. https://doi.org/10.4114/ia.v1i1.630

TORIBIO, M.; GONZÁLEZ, G.; MAGADAN, A. Método para identificar estados cognitivos a partir de datos artificial y algoritmos de aprendizaje automático. **XIKUA Boletín Científico De La Escuela Superior De Tlahuelilpan**, v. 12, n. 23, p. 48-55. 2024. https://doi.org/10.29057/xikua.v12i23.11795

TSANG, L.; GHORBANI, A.; MOHAMMAD, S.; KHATAMI, H. Intelligent Classification of Stable and Unstable Slope Conditions Based on Landslide Movement. **Journal of Rehabilitation in Civil Engineering**, v. 12, n. 3, p. 17–31, 2024. https://doi.org/https://doi.org/10.22075/jrce.2023.30293.1833

WANG, G.; ZHAO, B.; WU, B.; ZHANG, C.; LIU, W. Intelligent prediction of slope stability based on visual exploratory data analysis of 77 in situ cases. **International Journal of Mining Science and Technology**, v. 33, n. 1, p. 47–59, 2023. https://doi.org/10.1016/j.ijmst.2022.07.002

WENGANG, Z.; HANLONG, L.; LIN, W.; XING, Z.; YANMEI, Z. **Application of Machine Learning in Slope Stability Assessment**. Science Press. 2023. https://doi.org/https://doi.org/ 10.1007/978-981-99-2756-2

WICKRAMASINGHE, I. & KALUTARAGE, H. Naive Bayes: applications, variations and vul-nerabilities: a review of literature with code snippets for implementation. **Soft Computing**, v. 25, n. 3, p. 2277–2293, 2021. https://doi.org/https://doi.org/10.1007/s00500-020-05297-6 Naive

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,** v. 4, p. 426–453, 2023. https://doi.org/https://doi.org/10.3390/ modelling4040025

Submetido em 14 de novembro de 2024
Aceito para publicação em 5 de agosto de 2025