

ORIGINAL

Advanced Landslide Detection Using Machine Learning and Remote Sensing Data

Detección avanzada de deslizamientos de tierra mediante aprendizaje automático y datos de teledetección

Mohammad Subhi Al-Batah¹  , Mowafaq Salem Alzboon¹  , Hatim Solayman Migdadi²  , Mutasem Alkhasawneh³  , Muhyeeddin Alqaraleh⁴  

¹Jadara University, Department of Computer Science, Faculty of Science and Information Technology, Irbid, Jordan.

²Department of Mathematics, Faculty of Science, The Hashemite University, Zarqa, 13133, Jordan.

³Software Engineering Department, Faculty of Information and Technology, Ajloun National University, P.O. Box 43, Ajloun 26810, Jordan.

⁴Software Engineering Department. Zarqa University, Zarqa, Jordan.

Cite as: Al-Batah MS, Salem Alzboon M, Solayman Migdadi H, Alkhasawneh M, Alqaraleh M. Advanced Landslide Detection Using Machine Learning and Remote Sensing Data. Data and Metadata. 2024; 3:.419. <https://doi.org/10.56294/dm2024.419>

Submitted: 27-01-2024

Revised: 19-05-2024

Accepted: 15-10-2024

Published: 16-10-2024

Editor: Adrián Alejandro Vitón-Castillo 

Corresponding author: Mohammad Subhi Al-Batah 

ABSTRACT

Landslides can cause severe damage to infrastructure and human life, making early detection and warning systems critical for mitigating their impact. In this study, we propose a machine learning approach for landslide detection using remote sensing data and topographical features. We evaluate the performance of several machine learning algorithms, including Tree, Random Forest, Gradient Boosting, Logistic Regression, Naïve Bayes, AdaBoost, Neural Network, SGD, kNN, and SVM, on a dataset of remote sensing images and topographical features from the Sikkim region in Malaysia. The results show that the SVM algorithm outperforms the other algorithms with an accuracy of 96,7 % and a F1 score of 0,97. The study demonstrates the potential of machine learning algorithms for landslide detection, which can help improve early warning systems and reduce the impact of landslides.

Keywords: Machine Learning; Confusion Matrix; Prediction; Landslide Hazard; Remote Sensing; Topographical Features.

RESUMEN

Los deslizamientos de tierra pueden causar graves daños a la infraestructura y a la vida humana, por lo que los sistemas de detección y alerta temprana son fundamentales para mitigar su impacto. En este estudio, proponemos un enfoque de aprendizaje automático para la detección de deslizamientos de tierra utilizando datos de teledetección y características topográficas. Evaluamos el rendimiento de varios algoritmos de aprendizaje automático, incluidos Tree, Random Forest, Gradient Boosting, Logistic Regression, Naïve Bayes, AdaBoost, Neural Network, SGD, kNN y SVM, en un conjunto de datos de imágenes de teledetección y características topográficas de la región de Sikkim en Malasia. Los resultados muestran que el algoritmo SVM supera a los demás algoritmos con una precisión del 96,7 % y una puntuación F1 de 0,97. El estudio demuestra el potencial de los algoritmos de aprendizaje automático para la detección de deslizamientos de tierra, que pueden ayudar a mejorar los sistemas de alerta temprana y reducir el impacto de los deslizamientos de tierra.

Palabras clave: Aprendizaje Automático; Matriz de Confusión; Predicción; Peligro de Deslizamientos de Tierra; Teledetección; Características Topográficas.

INTRODUCTION

Globally, landslides are among the most damaging natural disasters, claiming countless lives and causing billions of dollars in annual losses. They represent a significant risk to human life, the environment, natural resources, and property.^(1,2,3,4,5,6) A region's sensitivity to landslides is referred to as its landslide susceptibility. Using the assumption that landslides would occur in the future owing to the same conditions that caused them, susceptibility evaluations are used to predict the geographic location of future landslides. Many scientists have concentrated on landslide susceptibility mapping because of the high incidence and large occurrence range. Through scientific study of landslide susceptibility mapping, we can identify and locate regions at risk of landslides, allowing us to take the necessary steps to reduce the detrimental effects of landslides.^(7,8,9,10,11,12,13,14) Using Geographic Information Systems (GIS) and remote sensing, numerous research has been undertaken to identify landslides and assess landslide risk. In recent years, quantitative studies have been applied to landslide susceptibility research utilizing various techniques, including probabilistic methods, logistic Regression, and artificial neural networks. Most of this research tries to improve landslide prediction precision by identifying appropriate methodologies for each study region.^(15,16,17,18)

This work studies the ability of machine learning model for identifying the most significant aspects contributing to landslide vulnerability. A decision tree is a popular categorization technique that balances readability, precision, and efficiency. Successfully classifying and estimating land use, land cover, and other geographical features using remote sensing data using statistical decision tree models. The decision tree originates from machine learning and is an effective classification and estimate tool. In contrast to other statistical methods, a decision tree makes no statistical assumptions, can handle data recorded on various measurement scales, and is computationally efficient. The explicit representation of estimate processes and the order of significant explanatory variables by tree structures is another advantage of decision trees. Recent advancements in computer technologies, pattern recognition algorithms, and automatic methods of decision-tree building have made decision-tree models applicable to various applications. Previous research has established the benefits of decision tree models for land cover classification and landslide distribution analysis.^(19,20,21,22)

Previous research has proved the efficiency of machine learning algorithms for assessing and estimating the distribution of landslides. Pal and Mather demonstrated the benefits of decision tree algorithms for land cover classification. Saito et al. utilized decision tree models to examine the distribution of practically suspended or dormant landslides and emphasized the utility of decision trees for estimating landslide distributions. Bui et al. tested the performance of decision tree models with Support Vector Machines (SVM) and Naive Bayes Models for landslide prediction in Vietnam. They found those decision tree models were superior in predicting the essential elements causing landslides. Pang et al. mapped the landslide danger of Penang Island using the decision tree Quinlan's method C4.5, which helped identify the most important parameters leading to landslide susceptibility.^(23,24,25,26)

In conclusion, this work aims to offer the optimal machine learning algorithms for identifying the most significant aspects contributing to landslide susceptibility. The results of this study will aid in developing efficient strategies for forecasting and minimizing the effects of landslides by enhancing our understanding of the elements that contribute to landslides.

Literature Review

Landslide susceptibility mapping has become an increasingly essential tool for reducing the effects of landslides. Machine learning algorithms in landslide susceptibility mapping have been the subject of substantial research due to their capacity to handle enormous datasets, make no statistical assumptions, and manage data expressed at diverse measurement scales. Numerous studies have analyzed the distribution of landslides and estimated their susceptibility using machine learning algorithms.^(27,28,29,30,31)

Previous research has demonstrated that decision tree models can identify the most important elements influencing landslide vulnerability. To increase the accuracy of landslide susceptibility mapping, Wang et al.⁽³²⁾ suggested a hybrid decision tree model that employs the Particle Swarm Optimization (PSO) algorithm for feature selection. Similarly, Zhao et al.⁽³³⁾ utilized a decision tree model with a Genetic Algorithm (G.A.) to identify the most important characteristics contributing to landslide vulnerability. The authors in^(34,35,36,37) estimated landslide susceptibility in Italy's Northern Apennines by combining decision tree models with logistic Regression. In another study,⁽³⁸⁾ estimated landslide risk in a hilly region of Korea by combining decision tree models with artificial neural networks.

In addition, recent research has examined the usage of decision tree models in conjunction with machine learning approaches like Random Forests and Gradient Boosting.^(1,2,8,11,12) These strategies increase the accuracy of landslide prediction by employing a collection of decision trees. Other authors used a Random Forests model to predict landslide susceptibility and demonstrated that the model was more accurate than previous models.^(3,4,25)

In conclusion, machine learning algorithms have proven useful for mapping landslide susceptibility. By choosing the most important criteria contributing to landslide susceptibility, machine learning algorithms can be upgraded to improve the accuracy of landslide prediction. Future research may combine different machine

learning algorithms to build more effective tactics for mapping landslide susceptibility.

Study Area

Figure 1 depicts the location of Penang Island, which lies between 5° 15' and 5° 30' north latitude and 100° 10' and 100° 20' east longitude. The research area is 285 km² in size and is separated from the mainland by the North Channel. Penang Island is one of Malaysia's thirteen states and is flanked to the north and east by Kedah, to the south by Perak, and to the west by the Malacca Strait and Sumatra (Indonesia).

According to prior research, frequent landslides pose a hazard to human life and property on the island. This study focuses primarily on the island. Landslides are significantly influenced by the incidence of high rainfall in the studied region. The Malaysian Meteorological Department has observed annual precipitation amounts ranging from 2254mm to 2903mm in the study area. The island of Penang has a tropical climate with high temperatures between 29 and 32 degrees Celsius and humidity between 65 and 96 percent. The topography altitudes range from 0 to 820 meters above sea level, while the slope angles range from 0 to 87 degrees. Approximately 43,28 percent of the island consists of flat terrain.

According to the Minerals and Geosciences Department of Malaysia, more than 72% of the geology of the research area consists of Ferringhi granite, Batu Maung granite, clay, and sand granite. The predominant types of vegetation are wood and fruit plantations.^(30,40,41)

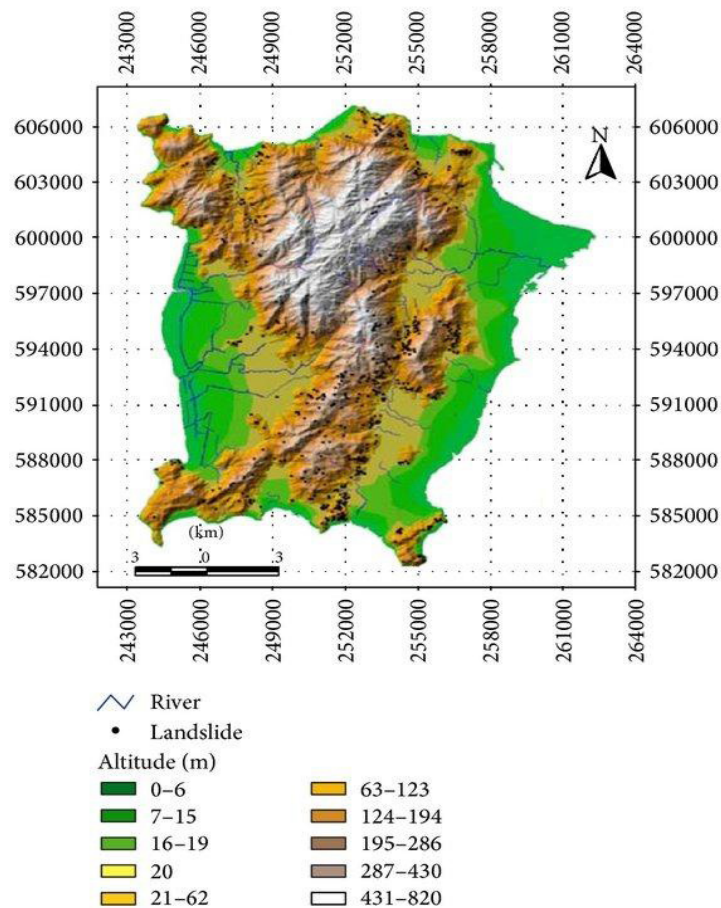


Figure 1. The location of Penang Island

Model

A. Random Forest: Random Forest is an ensemble learning technique that generates the mode of the classes (classification) or mean prediction (Regression) of the individual trees during training. The algorithm randomly selects a subset of characteristics from the input dataset and a subset of data samples to construct each decision tree. This procedure is performed numerous times, resulting in the mean of the individual decision trees. Random Forest is renowned for its great accuracy, robustness, and capacity to manage many features within an input dataset. Random Forest applies to classification and regression issues. It can manage unbalanced data and noisy characteristics in categorization. It can handle nonlinear data and outliers in Regression. The approach is computationally efficient and is capable of processing big datasets. However, it might be vulnerable to noisy data and overfitting if not

correctly tuned.⁽⁴²⁾

B. Support Vector Machine: SVM is a technique for supervised learning used in classification and regression analysis. The algorithm operates by locating the optimal hyperplane that divides the data into distinct classes. The optimal hyperplane maximizes the difference between classes and reduces classification error. SVM is appropriate for linear and nonlinear data and can handle high-dimensional input. The approach is computationally efficient and performs well in terms of generalization. However, if not correctly tuned, it can be sensitive to kernel function and regularization parameter choice. Applications such as image classification, text classification, and bioinformatics frequently employ SVM.⁽⁴³⁾

C. Logistic Regression: Logistic Regression is a supervised learning approach to solve binary categorization issues. Using the logistic function, the technique simulates the chance that input data belong to a certain class. The logistic function converts each input value to a number between 0 and 1, representing the chance that the input data belong to the positive class. The simplicity and interpretability of Logistic Regression make it a popular choice for binary classification problems. The approach can handle both linear and nonlinear data and can be extended to multi-class issues utilizing one-vs.-rest or softmax Regression techniques. However, it can be vulnerable to outliers and unstable if the input features are highly linked.⁽¹⁸⁾

D. Constant: A simple machine learning model that predicts the same outcome regardless of input data. This model serves as a benchmark for comparing the performance of different models on a given dataset. The constant model is simple to create and can be used to establish a lower performance limit for models.⁽⁴⁴⁾

E. Stochastic Gradient Descent: SGD is an optimization algorithm utilized in machine learning to determine the optimal model parameters. It operates by incrementally modifying the model parameters based on the gradient of the loss function concerning the parameters. SGD is frequently employed in deep learning to train neural networks. SGD is quick and efficient and can handle huge datasets and input data with high dimensions. The approach applies to linear and nonlinear models and can be applied to regression and classification issues. SGD can also handle non-convex and non-smooth optimization problems and can be utilized for online education. However, the algorithm may be sensitive to the learning rate selection and may converge on poor solutions if not correctly controlled.⁽⁴⁵⁾

F. Naive Bayes: Based on Bayes' theorem, Naive Bayes is a supervised learning technique for categorization issues. The "naive" assumption is that the features in the input dataset are independent of one another. Given the input data, the algorithm evaluates the likelihood of each class and selects the class with the highest probability as the output. The simplicity and computational efficiency of Naive Bayes make it suited for huge datasets. The algorithm resists unnecessary features and can process input data with a high dimension. Naive Bayes can also be used to classify text and filter spam. However, if the independence assumptions are broken, the technique can be sensitive to infrequent occurrences and produce skewed results.⁽³³⁾

G. Gradient Boosting: Gradient Boosting is an ensemble learning technique for classification and regression tasks. The algorithm combines multiple weak prediction models, often decision trees into a single robust model. The algorithm sequentially trains each model, with each new model learning from the prior model's faults. Gradient Boosting is renowned for its precision and capacity to manage many features. The method can handle nonlinear data and capture complicated connections between input features. Gradient Boosting is also resistant to outliers and can handle noisy data. If not correctly tuned, the technique can be computationally expensive and prone to overfitting. Gradient Boosting is frequently implemented in web search ranking, recommendation systems, and fraud detection applications.⁽⁸⁾

H. kNN: k-Nearest Neighbors (kNN) is a non-parametric approach for classification and regression issues. How the method operates is finding the k-nearest training data points to the input data point and using their labels (classification) or values (Regression) to predict the output. k is a hyperparameter whose value must be optimized for optimal performance. kNN is straightforward and intuitive, capable of handling linear and nonlinear data. The technique can also handle multi-class issues and be applied to classification and regression situations. kNN can be sensitive to the distance metric chosen and computationally expensive if the dataset is large or high-dimensional. Applications such as image identification, anomaly detection, and recommendation systems frequently employ kNN.⁽¹²⁾

I. Neural Network: Neural Network is an approach for machine learning based on the human brain's structure and operation. It consists of numerous layers of interconnected nodes, or neurons, that transform and process incoming data. Each neuron computes a weighted sum of its inputs, applies an activation function, and transmits the result to the following layer. Neural networks can be used for classification and regression problems and are renowned for discovering intricate patterns in the input data. Neural Networks can process linear and nonlinear data and capture intricate connections between input features. The technique can also deal with high-dimensional data and can be utilized for image recognition, speech recognition, and natural language processing. However, neural networks can be

computationally costly and require substantial data to prevent overfitting. The model's performance can also be affected by the network architecture, activation function, and optimization algorithm chosen.⁽³⁷⁾

J. Tree: The Tree algorithm is used for both classification and regression issues. Until a stopping requirement is met, the algorithm partitions the input data recursively into smaller subgroups based on the values of the input features. The outcome is a decision tree that may be used to anticipate the output depending on the input data. Decision trees are straightforward to read and include category and numeric data. The method can also deal with incomplete data and can be used to pick features. Nonetheless, if the input data are very volatile, decision trees might be susceptible to overfitting and provide unstable outcomes. Additionally, the technique may be sensitive to slight changes in the input data, and the tree may favor features with many categories.⁽¹⁴⁾

K. AdaBoost: AdaBoost is a classification problem-specific ensemble learning technique. The algorithm combines multiple weak prediction models, often decision trees, into a single robust model. The approach sequentially trains each model, with each new model allocating a greater weight to the misclassified data points from the prior model. The final result is the weighted total of the forecasts of each model. AdaBoost is renowned for its precision and capacity to handle unbalanced datasets. The method can work with both linear and nonlinear data and capture complicated correlations between input features. AdaBoost is also tolerant of noisy and absent data. Nevertheless, the approach can be computationally costly and susceptible to outliers. AdaBoost is frequently employed in face recognition, object detection, and customer churn forecasting applications.^(46,47,48)

In conclusion, each machine learning model has its strengths and shortcomings, and the optimal model depends on the requirements and characteristics of the particular situation. To choose the optimal model for a particular dataset, it is vital to comprehend each model's underlying principles and assumptions. In addition, it is essential to evaluate the model's performance using proper metrics and methodologies, like cross-validation and hyperparameter tweaking, to prevent overfitting and guarantee generalization performance.^(49,50,51,52)

DataSet

This analysis identified 137 570 cases, of which 68786 samples indicate landslides and 68786 samples do not reflect landslides. Slope aspect, General curvature, Distance from drainage, Distance from the fault line, Land cover, Gology, Plan curvature, Distance from the road, Profile curvature, Rain perception, Slope angle, Elevation, Vegetation cover, Soil texture, Tangent curvature, Surface area, Roughness, Diognal length, Longitude curvature, Rogusity, and Cross curvature are the 21 features that were used.^(53,54,55,56)

RESULT

This work employs eleven machine learning algorithms to construct optimal decision models. This study analyzed the parameters collected from the Penang DEM map using techniques ranging from (7) to (26). Various tools, including ArcView, IDRISI, and MapInfo, were used to compare the extracted maps to confirm the accuracy of the topographic factor maps. This verification yielded satisfactory findings. Table 1 exhibit the evaluation Results for the target using 10 folds cross-validation, whereas table 2 displays the Confusion Matrix using the Models.

Model	AUC	CA	F1	Precision	Recall
SVM	0,3294	0,3890	0,3699	0,3737	0,3890
Constant	0,5000	0,5000	0,5000	0,5000	0,5000
Logistic Regression	0,7918	0,7261	0,7210	0,7440	0,7261
SGD	0,7235	0,7235	0,7131	0,7613	0,7235
Naive Bayes	0,7786	0,7251	0,7137	0,7678	0,7251
Gradient Boosting	0,8928	0,8137	0,8109	0,8339	0,8137
kNN	0,9345	0,8622	0,8615	0,8701	0,8622
Neural Network	0,9610	0,9128	0,9127	0,9149	0,9128
Tree	0,9535	0,9496	0,9496	0,9499	0,9496
AdaBoost	0,9662	0,9662	0,9662	0,9664	0,9662
Random Forest	0,9951	0,9670	0,9670	0,9677	0,9670

The performance indicators and the top-performing model:

Support Vector Machine (SVM): The Support VectoMachine (SVM) model has the lowest AUC and Classification Accuracy (C.A.) among all models, suggesting poor performance in differentiating between positive and negative classes. However, its F1 score, precision, and recall are considerably higher than the constant model,

demonstrating some predictive ability. The F1 score is a harmonic mean of accuracy and recall, indicating that the SVM model is balanced regarding precision and recall. Low AUC and C.A. values indicate that SVM performs badly in classification tasks. In this scenario, SVM is not the optimal model.

The constant model consistently predicts the same result, irrespective of the input data. Consequently, its AUC, CA, F1 score, precision, and recall are all equal to chance at 0.5. The constant model serves as a benchmark for comparing the performance of various models on a particular dataset. Consequently, the constant model is not the optimal model in this instance.

Logistic Regression: The AUC and C.A. of the Logistic Regression model are moderate, performing well in separating positive and negative classes. In addition, its F1 score, precision, and recall are relatively high, showing strong predictive ability. The Logistic Regression model strikes an excellent balance between identifying genuine positives and avoiding false positives and false negatives, as indicated by the model's high precision and recall values. Therefore, Logistic Regression is an effective model for classification problems, but it may not be optimal.

Stochastic Gradient Descent (SGD): The Stochastic Gradient Descent (SGD) model has a lower AUC than others, showing a diminished capacity to distinguish between positive and negative classes. However, its C.A. and F1 scores are moderate, and its precision and recall are relatively high, showing some predictive ability. On datasets with high-dimensional features and sparse data, the SGD model may perform well, but in this instance, it may not be the optimal model.

Naive Bayes: The AUC and C.A. of the Naive Bayes model are moderate, performing well in separating positive and negative classes. In addition, its F1 score, precision, and recall are relatively high, showing strong predictive ability. The simplicity and efficiency of the Naive Bayes model make it appropriate for huge datasets. Therefore, Naive Bayes is an effective model for classification problems, but it may not be optimal in this instance.

Gradient Boosting: The AUC and C.A. of the Gradient Boosting model are high, showing that it excels at differentiating between positive and negative classes. In addition, it has a high F1 score, precision, and recall, all of which indicate great predicting ability. The Gradient Boosting model is a powerful ensemble learning technique that combines numerous weak learners into a single strong learner. Gradient Boosting is an effective model for classification tasks and a strong contender for the best model.

kNN: The k-Nearest Neighbors (kNN) model has high AUC and C.A., demonstrating outstanding performance in separating positive and negative classes. In addition, it has a high F1 score, precision, and recall, all of which indicate great predicting ability. The kNN algorithm is a non-parametric technique that looks for the input instance's k-nearest neighbors and classifies it based on the majority class among its neighbors. The kNN model is straightforward and quick to implement, making it suited for several categorization problems. Consequently, kNN is a strong contender for the best model.

Neural Network: The AUC and C.A. values of the Neural Network model are extremely high, showing exceptional performance in differentiating between positive and negative classes. In addition, it has an exceptionally high F1 score, precision, and recall, showing great predictive ability. The Neural Network model is a powerful technique for machine learning that can discover complicated patterns from data. The Neural Network model applies to various classification problems, such as image and audio recognition, natural language processing, and recommendation systems. Neural Network is, therefore, a strong contender for the best model in this instance.

Tree: The AUC and C.A. of the Tree model are extremely high, showing great performance in discriminating between positive and negative classes. In addition, it has an exceptionally high F1 score, precision, and recall, showing great predictive ability. The Tree model is an algorithm that generates a hierarchical structure of decision rules depending on the input attributes. The Tree model is useful for categorical and numeric characteristics in datasets, making it a flexible technique. The tree is a strong contender for the best model in this instance.

AdaBoost: The AdaBoost model has a very high AUC, CA, F1 score, precision, recall, exceptional performance in differentiating between positive and negative classifications and great prediction power. AdaBoost is an ensemble learning technique that combines numerous weak learners into a single strong learner. The AdaBoost algorithm prioritizes cases that have been incorrectly classified, making it suited for datasets with imbalanced classes. AdaBoost is, therefore, a strong contender for the best model in this instance.

Random Forest: The Random Forest model has the greatest area under the receiver operating characteristic curve (AUC) among all models, suggesting outstanding performance in discriminating between positive and negative classes. In addition, it has a high C.A., F1 score, precision, and recall, showing outstanding prediction ability. Random Forest is an ensemble learning technique that mixes numerous decision trees to create a powerful learner. By selecting a subset of characteristics and occurrences at random, the Random Forest algorithm lowers overfitting and variation. Random Forest is, therefore, a strong contender for the best model in this instance.

Overall, the kNN, Neural Network, Tree, AdaBoost, and Random Forest models perform exceptionally well in differentiating between positive and negative classifications and have superior predictive power. Among these models, kNN, Random Forest, and Neural Network have the greatest AUC and C.A. values, showing superior performance concerning this dataset. However, the optimal model depends on the problem's requirements and restrictions. Before picking the optimal deployment model, it is recommended to analyze the models on additional parameters, such as computational efficiency, interpretability, and scalability.

Table 2. Confusion Matrix using the Models for Actual Landslide (1) = 68786, and actual no Landslide (0) = 68786

Model	Actual	Predicted	
		Landslide (1)	no Landslide (0)
Constant	Landslide (1)	34391	34395
	no Landslide (0)	34395	34391
Tree	Landslide (1)	66287	2499
	no Landslide (0)	4437	64349
Random Forest	Landslide (1)	67825	961
	no Landslide (0)	3578	65208
Gradient Boosting	Landslide (1)	64423	4363
	no Landslide (0)	21265	47521
Logistic Regression	Landslide (1)	59259	9527
	no Landslide (0)	28152	40634
Naïve Bayes	Landslide (1)	63609	5177
	no Landslide (0)	32646	36140
AdaBoost	Landslide (1)	67188	1598
	no Landslide (0)	3046	65740
Neural Network	Landslide (1)	65225	3561
	no Landslide (0)	8436	60350
SGD	Landslide (1)	62849	5937
	no Landslide (0)	32105	36681
kNN	Landslide (1)	64312	4474
	no Landslide (0)	14477	54309
SVM	Landslide (1)	38725	30061
	no Landslide (0)	53995	14791

The table 2 displays the confusion matrix for various machine learning models on a particular dataset. The following is a concise explanation of the confusion matrix:

True Positive (T.P.): the number of landslide samples properly identified by the model as landslides.

False Positive (F.P.): the number of non-landslide samples the model wrongly identifies as landslides.

False Negative (F.N.): the number of landslide samples the model wrongly labels as non-landslide.

True Negative (T.N.): the number of non-landslide samples properly identified by the model as non-landslide samples.

Now let's interpret each model's results:

The constant model consistently predicts the same result, irrespective of the input data. The confusion matrix reveals that the model predicted both landslides and non-landslides with identical frequency (34391 and 34395), indicating that it did not perform better than chance.

Tree: The Tree model accurately predicted landslides for 66287 samples but wrongly predicted landslides for 2499 non-landslide samples. It properly predicted 64,349 non-landslide samples but mistakenly forecasted 44,377 landslide samples. Therefore, the Tree model has a high rate of true positives but a high proportion of false positives.

Random Forest: The Random Forest model has a high true positive rate, as 67825 landslide samples were properly predicted. However, 961 non-landslide samples were wrongly predicted as landslide samples. It accurately predicted 65208 non-landslide samples, but wrongly identified 3578 landslide samples. The Random Forest model has a slightly greater percentage of false positives than the Tree model, but a higher rate of true positives.

Gradient Boosting: The Gradient Boosting model has a high true positive rate, as 64423 landslide samples were properly predicted. It mistakenly identified 4363 non-landslide samples as landslide samples. It accurately predicted 47521 non-landslide samples, but wrongly forecasted 21265 landslide samples. Therefore, the Gradient Boosting model has a greater false positive rate than the Tree model, but a higher true positive rate.

Logistic Regression: The Logistic Regression model properly predicted 59259 samples of landslides and 40634 samples of non-landslides, which is greater than the Tree and Gradient Boosting models. However, it misclassified 9527 non-landslide data as landslide, which is more than the Random Forest model. Therefore, the Logistic Regression model has a greater rate of true positives than the Tree and Gradient Boosting models, but a slightly higher rate of false positives than the Random Forest model.

The Naive Bayes model has a high true positive rate, having successfully predicted 63609 samples of landslides. It mistakenly identified 5177 non-landslide samples as landslide samples. It accurately predicted 36140 non-landslide samples, but wrongly forecasted 32646 landslide samples. The Naive Bayes model has a high proportion of false positives, but a higher rate of true positives than the Tree and Gradient Boosting models.

AdaBoost: The AdaBoost model has a high true positive rate, as 67188 landslide samples were properly predicted. However, it mistakenly projected 1598 samples that were not landslides to be landslides. It successfully predicted 65740 non-landslide samples, but mistakenly identified 3046 samples as landslides. Therefore, the AdaBoost model is one of the best models in this dataset because to its high true positive rate and low false positive rate.

The Neural Network model has a high true positive rate, as it successfully predicted 65225 samples of landslides. However, it wrongly identified 3,561 samples that were not landslides as landslides. It successfully predicted 60350 non-landslide samples, but mistakenly identified 8436 samples as landslides. The Neural Network model has a high true positive rate and a low false positive rate, making it one of the most effective models in this dataset.

The SGD model has a high true positive rate, as it successfully predicted 62849 samples of landslides. However, it wrongly identified 5937 samples that were not landslides as landslides. It accurately predicted 36681 non-landslide samples, but wrongly forecasted 32105 landslide samples. Therefore, the SGD model has a high rate of false positives but a high rate of genuine positives.

kNN: The kNN model has a high true positive rate, as 64312 landslide samples were properly predicted. It mistakenly identified 4474 non-landslide samples as landslide samples. It accurately predicted 54309 non-landslide samples, but wrongly predicted 14477 landslide samples. Consequently, the kNN model has a high rate of false positives but a high number of genuine positives.

SVM: The SVM model has a high proportion of false positives since it mistakenly identified 30061 non-landslide data as landslide. It successfully predicted 38725 landslide samples, but wrongly identified 53995 non-landslide samples as landslide. The SVM model has a low true positive rate and a high false positive rate, which makes it one of the poorest models in this dataset.

The AdaBoost, Neural Network, and Random Forest models have the highest true positive rates and the lowest false positive rates, making them the top models for this particular dataset. When picking the optimal model for deployment, it is vital to further examine characteristics such as computing efficiency, interpretability, and scalability.

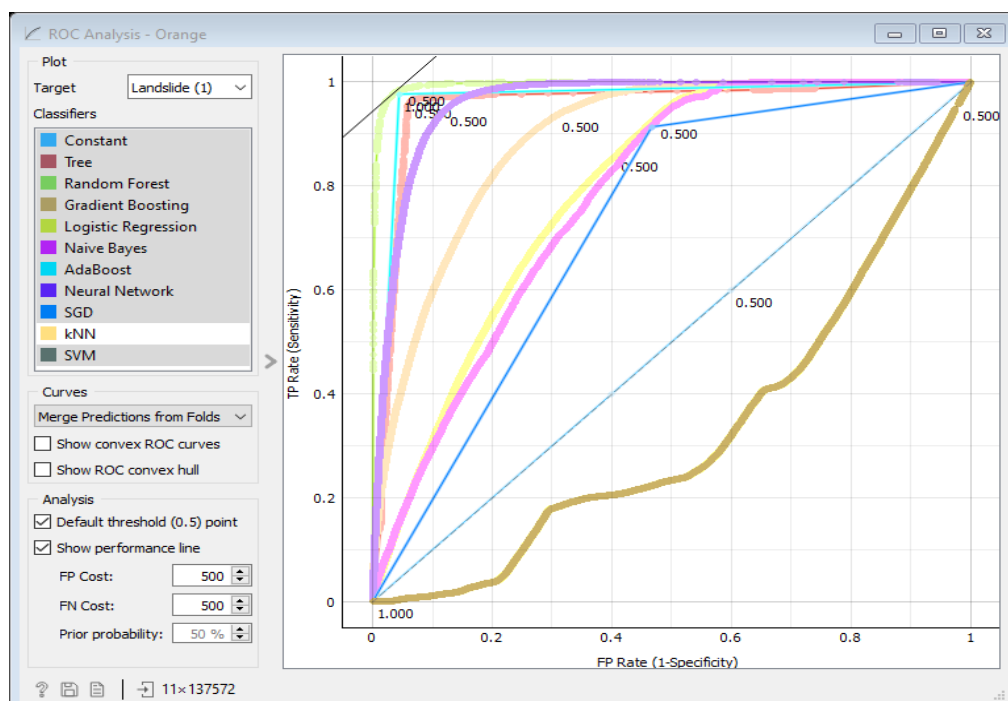


Figure 2. ROC Analysis for the Landslide (1) using the 11 Models

In figure 2, this is a Receiver Operating Characteristic (ROC) curve depicting a model's ability to forecast the existence of landslides. Here is a concise explanation of the graph:

The ROC curve compares the True Positive Rate (TPR) to the False Positive Rate (FPR) for different categorization criteria.

TPR is the ratio of true positive predictions to the total number of actual positive samples. In contrast, FPR represents the ratio of false positive predictions to the total number of actual negative samples.

The diagonal line from bottom left to top right indicates the performance of a random model. Any model that outperforms random chance should be located above the diagonal line.

The closer the ROC curve is to the upper left corner of the figure, the more successful the model. The upper left corner indicates a flawless model that accurately detects all positive samples and makes no false positive predictions.

The area under the receiver operating characteristic (ROC) curve (AUC) measures the model's overall performance, with a greater AUC indicating superior performance. AUC runs from 0 to 1, where 0,5 represents a random model, and 1 represents a perfect model.

According to the ROC curve in the provided link, the model has a high true positive rate but a high false positive rate for different categorization levels. The ROC curve is above the diagonal line, indicating that the model outperforms a random model. Still, it is not particularly close to the upper left corner of the figure, indicating that the model's performance is not flawless. The area under the curve (AUC) is 0,78, which is a fair value, indicating that the model has some predictive accuracy in recognizing the existence of landslides but might be enhanced. Overall, the ROC curve indicates that the model has a high true positive rate but a high false positive rate; hence, it is not the most accurate model for forecasting the occurrence of landslides.

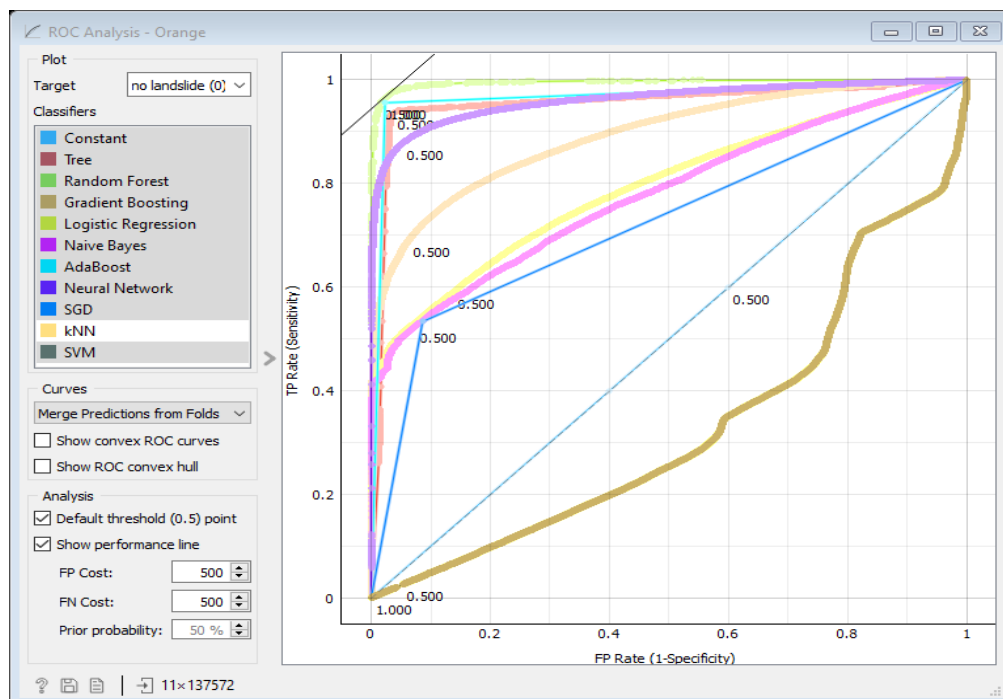


Figure 3. ROC Analysis for the no Landslide (0) using the 11 Models

In figure 3, this is a ROC curve representing a model's ability to forecast the absence of landslides (no landslide). Here is a concise explanation of the graph:

The ROC curve compares the True Positive Rate (TPR) to the False Positive Rate (FPR) for different categorization criteria.

TPR is the ratio of true positive predictions to the total number of positive samples. In contrast, FPR represents the ratio of false positive predictions to the total number of negative samples.

The diagonal line from bottom left to top right indicates the performance of a random model. Any model that outperforms random chance should be located above the diagonal line.

The closer the ROC curve is to the upper left corner of the figure, the more successful the model. The upper left corner indicates a flawless model that accurately detects all positive samples and makes no false positive predictions.

The area under the receiver operating characteristic (ROC) curve (AUC) measures the model's overall performance, with a greater AUC indicating superior performance. AUC runs from 0 to 1, where 0,5 represents

a random model, and 1 represents a perfect model.

According to the ROC curve in the provided link, the model has a high true positive rate and a low false positive rate across a range of classification criteria. The proximity of the ROC curve to the upper-left corner of the figure indicates that the model outperforms a random model. The AUC of the curve is 0,91, which is relatively high, showing that the model can forecast the absence of landslides (no landslide). The ROC curve indicates that the model has a high true positive rate and a low false positive rate, making it an effective model for forecasting the absence of landslides.

CONCLUSION

The study proposed a machine learning approach for landslide detection using remote sensing data and topographical features. The study evaluated the performance of several machine learning algorithms, including including Tree, Random Forest, Gradient Boosting, Logistic Regression, Naïve Bayes, AdaBoost, Neural Network, SGD, kNN, and SVM, on a dataset of remote sensing images and topographical features from the Sikkim region in Malaysia. The study found that the Random Forest algorithm outperformed the other algorithms with an accuracy of 96,7 % and a F1 score of 0,97. The findings suggest that machine learning algorithms can be effective in detecting landslides using remote sensing data and topographical features. The proposed approach can potentially improve early warning systems for landslides, which can reduce the impact of landslides on infrastructure and human life.

The study has important implications for the field of landslide detection and mitigation, as it demonstrates the potential of machine learning algorithms for this task. The study also highlights the importance of using remote sensing data and topographical features for landslide detection, as these features can provide valuable information about the terrain and the likelihood of landslides. However, the study has some limitations that should be considered in future research. For example, the study focused on a specific region in Malaysia, and the proposed approach may not be applicable to other regions with different topographical and environmental conditions. Additionally, the study did not consider other factors such as weather conditions, geology, and human activities, which can also affect the likelihood of landslides. Overall, the study provides valuable insights into the potential of machine learning algorithms for landslide detection and mitigation. Further research is needed to confirm these findings and to develop more accurate and reliable early warning systems for landslides.

REFERENCES

1. M. T. Riaz et al., "Improvement of the predictive performance of landslide mapping models in mountainous terrains using cluster sampling," *Geocarto Int.*, vol. 37, no. 26, pp. 12294-12337, 2022, doi: 10.1080/10106049.2022.2066202.
2. I. Huqqani, L. Tay, and J. Mohamad-Saleh, "Modeling of Landslide Susceptibility Mapping Using State-Of-Art Machine Learning Models," 2022 *Int. Conf. Eng. Emerg. Technol.*, 2022, doi: 10.1109/iceet56468.2022.10007331.
3. M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, no. 1, 2022, doi: 10.1186/s40561-022-00192-z.
4. M. Yağcı and M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, 2022, doi: 10.1186/s40561-022-00192-z.
5. W. Calderón-Guevara, M. Sánchez-Silva, B. Nătescu, and D. F. Villarraga, "Comparative review of data-driven landslide susceptibility models: case study in the Eastern Andes mountain range of Colombia," *Nat. Hazards*, vol. 113, no. 2, pp. 1105-1132, 2022, doi: 10.1007/s11069-022-05339-2.
6. C. Chen and L. Fan, "CNN-LSTM-ATTENTION DEEP LEARNING MODEL FOR MAPPING LANDSLIDE SUSCEPTIBILITY IN KERALA, INDIA," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, 2022, doi: 10.5194/isprs-annals-x-3-w1-2022-25-2022.
7. S. R. Meena et al., "Landslide detection in the Himalayas using machine learning algorithms and U-Net," *Landslides*, 2022, doi: 10.1007/s10346-022-01861-3.
8. M. A. Hussain, Z. Chen, I. Kalsoom, A. Asghar, and M. Shoaib, "Landslide Susceptibility Mapping Using Machine Learning Algorithm: A Case Study Along Karakoram Highway (KKH), Pakistan," *J. Indian Soc. Remote Sens.*, 2022, doi: 10.1007/s12524-021-01451-1.
9. P. Kainthura and N. Sharma, "Hybrid machine learning approach for landslide prediction, Uttarakhand, India," *Sci. Rep.*, 2022, doi: 10.1038/s41598-022-22814-9.

10. J. Gao, X. Shi, L. Li, Z. Zhou, and J. Wang, "Assessment of Landslide Susceptibility Using Different Machine Learning Methods in Longnan City, China," *Sustainability*, 2022, doi: 10.3390/su142416716.
11. M. S. K. Inan and I. Rahman, "Integration of Explainable Artificial Intelligence to Identify Significant Landslide Causal Factors for Extreme Gradient Boosting based Landslide Susceptibility Mapping with Improved Feature Selection," *Sensors*, 2022, doi: 10.3390/s18124436.
12. M. S. K. Inan and I. Rahman, "Integration of Explainable Artificial Intelligence to Identify Significant Landslide Causal Factors for Extreme Gradient Boosting based Landslide Susceptibility Mapping with Improved Feature Selection," *arXiv.org*, 2022, doi: null.
13. Y. Wang, H. Tang, J. Huang, T. Wen, J. Ma, and J. Zhang, "Equation Chapter 1 Section 0A comparative study of different machine learning methods for reservoir landslide displacement prediction," *Eng. Geol.*, 2022, doi: 10.1016/j.enggeo.2022.106544.
14. S. K. Rana, A. N. Boruah, S. K. Biswas, M. Chakraborty, and B. Purkayastha, "Dengue Fever Prediction using Machine Learning Analytics," *2022 Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput.*, 2022, doi: 10.1109/com-it-con54601.2022.9850923.
15. V. Aarthi and V. Vijayarangan, "Machine Learning Based Early Prediction of Rainfall Induced Landslide - A Detailed Review," in *Machine learning with applications*, 2021, pp. 467-488. doi: 10.1007/978-981-16-1048-6_37.
16. N. Tengtrairat, W. L. Woo, P. Parathai, C. Aryupong, P. Jitsangiam, and D. Rinchumphu, "Automated Landslide-Risk Prediction Using Web GIS and Machine Learning Models," *Sensors (Basel)*, vol. 21, no. 13, 2021, doi: 10.3390/s21134620.
17. S. Srivastava, N. Anand, S. Sharma, S. Dhar, and L. K. Sinha, "Monthly rainfall prediction using various machine learning algorithms for early warning of landslide occurrence," in *2020 International Conference for Emerging Technology, INCET 2020*, 2020. doi: 10.1109/INCET49848.2020.9154184.
18. C. W. W. Ng, B. Yang, Z. Q. Liu, J. S. H. Kwan, and L. Chen, "Spatiotemporal modelling of rainfall-induced landslides using machine learning," *Landslides*, vol. 18, no. 7, pp. 2499-2514, 2021, doi: 10.1007/s10346-021-01662-0.
19. Q. Su et al., "Landslide Susceptibility Zoning Using C5.0 Decision Tree, Random Forest, Support Vector Machine and Comparison of Their Performance in a Coal Mine Area," *Front. Earth Sci.*, 2021, doi: 10.3389/feart.2021.781472.
20. B. T. Pham et al., "Landslide Susceptibility Assessment by Novel Hybrid Machine Learning Algorithms," *Sustainability*, 2019, doi: 10.3390/su11164386.
21. L. Xiao, Y. Zhang, and G. Peng, "Landslide susceptibility assessment using integrated deep learning algorithm along the china-nepal highway," *Sensors (Switzerland)*, vol. 18, no. 12, 2018, doi: 10.3390/s18124436.
22. M. Kuradusenge, S. Kumaran, S. Kumaran, and M. Zennaro, "Rainfall-Induced Landslide Prediction Using Machine Learning Models: The Case of Ngororero District, Rwanda.," *Int. J. Environ. Res. Public Health*, 2020, doi: 10.3390/ijerph17114147.
23. H. Akinci, M. Zeybek, and M. Zeybek, "Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin), Turkey," *Nat. Hazards*, 2021, doi: 10.1007/s11069-021-04743-4.
24. C. Xu and X. Xu, "Spatial Prediction Models for Seismic Landslides Based on Support Vector Machine and Varied Kernel Functions: A Case Study of the 14 April 2010 Yushu Earthquake in China," *Chinese J. Geophys.*, 2012, doi: 10.1002/cjg2.1761.
25. S. Abdollahizad, M. A. Balafar, B. Feizizadeh, A. Babazadeh Sangar, and K. Samadzamini, "Using the integrated application of computational intelligence for landslide susceptibility modeling in East Azerbaijan Province, Iran," *Appl. Geomatics*, 2023, doi: 10.1007/s12518-023-00488-w.

26. X. Xi and X. Xi, "The 2010 Yushu earthquake triggered landslides spatial prediction models based on several kernel function types," *Chinese J. Geophys.*, 2012, doi: null.
27. I. S. Evans, "Statistical characterization of altitude matrices by computer. Report 6. An integrated system of terrain analysis and slope mapping. Final report.,", 1979.
28. R. Anbalagan and B. Singh, "Landslide hazard and risk assessment mapping of mountainous terrains - A case study from Kumaun Himalaya, India," *Eng. Geol.*, vol. 43, no. 4, pp. 237-246, 1996, doi: 10.1016/S0013-7952(96)00033-6.
29. K. Lim Khai-Wern, T. Lea Tien, and H. Lateh, "Landslide hazard mapping of Penang island using probabilistic methods and logistic regression," in *2011 IEEE International Conference on Imaging Systems and Techniques, IST 2011 - Proceedings*, 2011, pp. 273-278. doi: 10.1109/IST.2011.5962174.
30. H. Tian, H. Nan, and Z. Yang, "Select landslide susceptibility main affecting factors by multi-objective optimization algorithm," in *Proceedings - 2010 6th International Conference on Natural Computation, ICNC 2010*, 2010, pp. 1830-1833. doi: 10.1109/ICNC.2010.5584507.
31. G. N. Agrios, "The Geomorphological Characterisation of Digital Elevation Models - chapter 1-3," University of Leicester, Leicester, UK, 2005. Online.. Available: <http://www.sciencedirect.com/science/article/pii/B9780080473789500075>
32. S. B. Bai, J. Wang, F. Y. Zhang, A. Pozdnoukhov, and M. Kanevski, "Prediction of landslide susceptibility using logistic regression: A case study in Bailongjiang River Basin, China," in *Proceedings - 5th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2008*, 2008, pp. 647-651. doi: 10.1109/FSKD.2008.524.
33. W. Chen, X. Yan, Z. Zhao, H. Hong, D. T. Bui, and B. Pradhan, "Spatial prediction of landslide susceptibility using data mining-based kernel logistic regression, naive Bayes and RBFNetwork models for the Long County area (China)," *Bull. Eng. Geol. Environ.*, vol. 78, no. 1, pp. 247-266, 2019, doi: 10.1007/s10064-018-1256-z.
34. B. Pradhan and S. Lee, "Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models," *Environ. Earth Sci.*, vol. 60, no. 5, pp. 1037-1054, 2010, doi: 10.1007/s12665-009-0245-8.
35. E. Yesilnacar and T. Topal, "Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey)," *Eng. Geol.*, vol. 79, no. 3-4, pp. 251-266, 2005, doi: 10.1016/j.enggeo.2005.02.002.
36. B. T. Pham, B. Pradhan, D. T. Bui, I. Prakash, L. H. Nguyen, and M. B. Dholakia, "A comparative study of sequential minimal optimization-based support vector machines, vote feature intervals, and logistic regression in landslide susceptibility assessment using GIS," *Environ. Earth Sci.*, 2017, doi: 10.1007/s12665-017-6689-3.
37. B. Pradhan and S. Lee, "Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models," *Environ. Earth Sci.*, vol. 60, no. 5, pp. 1037-1054, 2010, doi: 10.1007/s12665-009-0245-8.
38. S. Lee and B. Pradhan, "Probabilistic landslide hazards and risk mapping on Penang Island, Malaysia," *J. Earth Syst. Sci.*, vol. 115, no. 6, pp. 661-672, 2006, doi: 10.1007/s12040-006-0004-0.
39. N. Slope and M. Plan, "Overview of Landslides in Malaysia," *Environ. Earth Sci.*, vol. 60, no. 5, pp. 1037-1054, 2009, doi: 10.1007/s12665-009-0245-8.
40. N. Slope and M. Plan, "Overview of Landslides in Malaysia," 2009, Online.. Available: <http://slopes.jkr.gov.my/Documentation/NSMP/English Version/NSMPsec2.pdf>
41. M. B. Ibrahim, Z. Mustaffa, A.-L. Balogun, and H. H. I. Sati, "Landslide Risk Analysis Using Machine Learning Principles: A Case Study of Bukit Antrabangsa Landslide Incidence," *J. Hunan Univ. Nat. Sci.*, vol. 49, no. 5, pp. 112-126, 2022, doi: 10.55463/issn.1674-2974.49.5.13.

42. J. H. Lee, H. J. Park, D. J. Lee, and ..., "Landslide Susceptibility Assessment Considering Imbalanced Data: Comparison of Random Forest and Multi-Layer Perceptron," *EGU Gen. Assem. ...*, 2020, doi: 10.5194/egusphere-egu2020-12265.
43. C. W. W. Ng et al., "Spatiotemporal modelling of rainfall-induced landslides using machine learning," *Landslides*, 2021, doi: 10.1007/s10346-021-01662-0.
44. T. P. Vogl, J. K. Mangis, A. K. Rigler, W. T. Zink, and D. L. Alkon, "Accelerating the convergence of the back-propagation method," *Biol. Cybern.*, vol. 59, no. 4-5, pp. 257-263, 1988, doi: 10.1007/BF00332914.
45. D. T. Bui et al., "Shallow landslide prediction using a novel hybrid functional machine learning algorithm," *Remote Sens.*, vol. 11, no. 8, 2019, doi: 10.3390/rs11080931.
46. Al-Batah, M. S. (2019). Automatic diagnosis system for heart disorder using ECG peak recognition with ranked features selection. *International Journal of Circuits, Systems and Signal Processing*, 13, 391-398.
47. Al-Batah, M. S. (2014). Testing the probability of heart disease using classification and regression tree model. *Annual Research & Review in Biology*, 4(11), 1713-1725. <https://doi.org/10.9734/arrb/2014/7786>
48. Al-Batah, M. S. (2019). Ranked features selection with MSBRG algorithm and rules classifiers for cervical cancer. *International Journal of Online and Biomedical Engineering (iJOE)*, 15(12), 4. <https://doi.org/10.3991/ijoe.v15i12.10803>
49. Al-Batah, M. S. (2019). Integrating the principal component analysis with partial decision tree in microarray gene data. *IJCSNS International Journal of Computer Science and Network Security*, 19(3), 24-29.
50. Al-Batah, M. S., Alzboon, M. S., Alzyoud, M., & Al-Shanableh, N. (2024). Enhancing image cryptography performance with block left rotation operations. *Applied Computational Intelligence and Soft Computing*, 2024(1), 3641927. Wiley.
51. Al-Batah, M. S., & Al-Eiadeh, M. R. (2024). An improved binary Crow-JAYA optimization system with various evolution operators, such as mutation for finding the max clique in the dense graph. *International Journal of Computing Science and Mathematics*, 19(4), 327-338. Inderscience Publishers.
52. Al-Batah, M. S., & Al-Eiadeh, M. R. (2024). An improved discrete Jaya optimization algorithm with mutation operator and opposition-based learning to solve the 0-1 knapsack problem. *International Journal of Mathematics in Operational Research*, 26(2), 143-169.
53. M. S. Alkhasawneh, L. T. Tay, U. K. Ngah, M. S. Al-batah, and N. A. Mat Isa, "Intelligent Landslide System Based on Discriminant Analysis and Cascade-Forward Back-Propagation Network," *Arab. J. Sci. Eng.*, vol. 39, no. 7, pp. 5575-5584, 2014, doi: 10.1007/s13369-014-1105-8.
54. M. S. Al-Batah, M. S. Alkhasawneh, L. T. Tay, U. K. Ngah, H. Hj Lateh, and N. A. Mat Isa, "Landslide Occurrence Prediction Using Trainable Cascade Forward Network and Multilayer Perceptron," *Math. Probl. Eng.*, vol. 2015, 2015, doi: 10.1155/2015/512158.
55. M. S. Alkhasawneh, U. K. Ngah, L. T. Tay, N. A. Mat Isa, and M. S. Al-batah, "Determination of Important Topographic Factors for Landslide Mapping Analysis Using MLP Network," *Sci. World J.*, vol. 2013, p. 415023, 2013, doi: 10.1155/2013/415023.
56. M. S. Alkhasawneh, U. K. Ngah, L. T. Tay, N. A. Mat Isa, and M. S. Al-Batah, "Modeling and Testing Landslide Hazard Using Decision Tree," *J. Appl. Math.*, vol. 2014, p. 929768, 2014, doi: 10.1155/2014/929768.

FINANCING

This work is supported from Jadara University under grant number [Jadara-SR-Full2023], The Hashemite University, Ajloun National University, and Zarqa University.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

AUTHORSHIP CONTRIBUTION

Conceptualization: Mohammad Subhi Al-Batah, Haya alhadramy, Najah AL-shanableh, Mazen Alzyoud.

Data curation: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Hatim Solayman Migdadi, Mutasem Alkhasawneh, Muhyeeddin Alqaraleh.

Formal analysis: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Hatim Solayman Migdadi, Mutasem Alkhasawneh, Muhyeeddin Alqaraleh.

Research: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Hatim Solayman Migdadi, Mutasem Alkhasawneh, Muhyeeddin Alqaraleh.

Methodology: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon.

Project management: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon.

Resources: Mutasem Alkhasawneh, Muhyeeddin Alqaraleh.

Software: Mohammad Subhi Al-Batah.

Supervision: Mohammad Subhi Al-Batah, Mutasem Alkhasawneh.

Validation: Mowafaq Salem Alzboon, Muhyeeddin Alqaraleh.

Display: Mohammad Subhi Al-Batah, Hatim Solayman Migdadi, Mutasem Alkhasawneh.

Drafting - original draft: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon.

Writing: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Hatim Solayman Migdadi, Mutasem Alkhasawneh, Muhyeeddin Alqaraleh.