iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71777

Applications of Tree Based Regressors in Predicting Factor of Safety in Slope Stability and Observing Feature Importance

Md. Nahul Rahman[1], Ejaj Ahmed Sadiq[2], Hibzur Rahman Mahim[3], Sachitra Halder[4], Fardin Rahman[5]
[1, 4, 5]*Department of Computer Science and Engineering, Military Institute of Science and Technology, Mirpur Cantonment, Dhaka-1216, Bangladesh*
[2, 3]*Department of Civil Engineering, Military Institute of Science and Technology, Mirpur Cantonment, Dhaka-1216, Bangladesh*

Abstract: In the geotechnical engineering field, it is required to anticipate the Factor of Safety (FOS) in slope stability precisely in order to assess the possibility of slope failure and guarantee infrastructure safety. This research utilizes a thorough slope stability dataset to inspect how well six tree-based regression models—Decision Tree, Random Forest, Extra Trees, AdaBoost, Gradient Boosting, and XGBoost—predict the FOS. With the target of predicting the continuous FOS value, the dataset covers 10,000 samples with eight vital geotechnical parameters and one categorical reinforcement feature. Using performance metrics like RMSE, MAE, R^2 score, and execution time, a modified study was executed. The most significant factors affecting slope stability were also resolved using feature importance analysis. The Extra Trees Regressor performs finer than other models in terms of predictive accuracy, according to the results, while cohesion, internal friction angle, slope angle, and pore water pressure ratio decrease.
Keywords: Slope Stability, Factor of Safety, Tree-Based Regressors, Feature Importance, Ensemble Learning, Geotechnical Engineering.

I. INTRODUCTION

Slope stability analysis is a fundamental feature of geotechnical engineering, necessary for preventing landslides and ensuring the structural integrity of slopes in construction and environmental projects. The Factor of Safety (FOS) serves as a quantitative measure that distinguishes stable slopes (FOS $\geq$ 1) from unstable ones (FoS < 1). Traditional analytical methods for finding FOS often depend on empirical relationships and deterministic assumptions, which may overlook the complex, nonlinear interactions among geotechnical parameters. With the emergence of machine learning, predictive modelling has become a strong equipment to mark these challenges. Among numerous machine learning techniques, tree-based regressors are comparably well-suited for interpreting and predicting continuous geotechnical outcomes due to their robustness, interpretability, and ability to capture nonlinearity. This study investigates the application of six such models—Decision Tree Regressor, Random Forest Regressor, Extra Trees Regressor, AdaBoost Regressor, Gradient Boosting Regressor, and XGBoost Regressor—in predicting the FoS using a real-world dataset. The analysis concentrated on model precision, computational capability and feature importance to provide an extensive evaluation of each regressor's applicability for slope stability prediction.

This paper points to bridge the intermission between conventional geotechnical assessments and modern data-driven approaches by highlighting how ensemble tree-based methods can enhance predictive accuracy and interpretability in slope stability studies.

II. LITERATURE STUDIES

In order to understand the work in a deep level, studies related to this work integrating with machine learning were observed and the most aligned works can be discussed. A study's main aim of is to expand a composite learning environment for anticipating the slope stability prediction. The experiment went under Random Forest (RF) and Extreme Gradient Boosting (XGBoost) for prediction and trying to apply those routes for stability prediction of 786 landslide scenarios in China. This paper extensively investigates and compares RF and XGBoost in opposition to traditional methods like SVM, LR and so on. Most importantly, this paper concentrates on detecting the 12 controlling variables to identify the key impact for slope stability predictions. Compared to our papers, this paper applies multi class classification categories and real-world field data whereas we use pure regression problem investigating factors of safety values and large synthetic data basis [1].

Another study had the objective to reduce effectively the risk of rock slope failures on highways in the Higher Himalayas and behaviour of rock masses. This paper involves a summary of 18 locations along a road cut section in the Higher Himalayan region of India using RQD, RMR, SMR, Q-slope, and GSI. In terms of comparison, this paper has no validation on independent dataset and it is dependent on static condition data analysis or traditional geomechanical classification where machine learning scopes were limited [2].

A study aimed to examine the stability of slopes based on global sensitivity analysis, aiming to provide a systematic approach to comprehend how uncertainties in input variables affect slope stability. Applying Global Sensitivity Analysis (GSA) using the Sobol method and Slope stability analysis using the bishop's simplified method, the research finds cohesion, friction angle, slope angle using the methods mentioned before. Though researchers used classical geotechnical engineering methods for prediction, the model has uncertainty in input parameters and numerical intensity of global sensitive analysis using the model [3].

A study focused on determining stability using ensemble-based hybrid machine learning approaches for improved precision and solidity by using Random Forest, XGBoost, LightGBM for ensemble learning and Feature selection techniques Recursive Feature Elimination, Principal Component Analysis and model stacking is used for hybrid modelling. Higher prediction accuracy is observed ($R^2 > 0.9$, RMSE < 10%) for both training and testing datasets in this paper [4].

Another study wanted build on the prediction of slope stability under the pair of static and dynamic conditions using Random Tree (RT) and Reduced Error Pruning Tree (REP Tree) models and provide precise models to differentiate stable or unstable slopes. Using RT and REP models with a 700 slope case datasets and Limit Equilibrium Method (LEM) using Geostudio, after implementing the model, 97.14% accuracy for random tree and 95.43% accuracy for the REP Tree model were observed. This model cannot interpret complex field conditions through modelling. In terms of our paper, we use RMSE and MAE and get a significant result with high accuracy for extra tree regression which is better than that model [5].

The research conducted by Sahoo et al. (2024) assessed the use of machine learning in analysing slope stability, emphasizing the types of models, input features, and metrics for evaluation. A comprehensive review of 53 studies from 2015 to 2023 highlighted Random Forest and Support Vector Machine as the most effective algorithms, with RMSE and AUC being commonly employed metrics. The majority of the studies relied on data-driven approaches with a restricted number of input parameters, frequently utilizing limited or historical data sets. Notable drawbacks included the absence of dynamic features, applicability specific to certain sites, and the reliance on manual tuning of models. Both studies employ tree-based regression models and analyse feature importance to predict factors of safety. Our project utilizes a large synthetic dataset that includes categorical features and concentrates exclusively on regression, contrasting with the mixed-method approaches evident in the reviewed studies [6].

Another study aimed to forecast slope stability for circular failures utilizing machine learning models (RF, SVM, XGB) enhanced through Bayesian Optimization to improve accuracy and generalizability. A total of 627 cases were used to train and optimize the models with Bayesian Optimization, which were then assessed through 5-fold cross-validation and analysed for feature significance. The RF-BO model produced the best outcomes, achieving an accuracy of 95.5%, with cohesion, friction angle, and slope height identified as crucial factors. The limitations included a small dataset, applicability restricted to homogeneous circular failures, and the exclusion of dynamic factors. Both studies utilize RF/XGB and feature analysis focusing on essential geotechnical inputs. While this research employs real-world classification with BO, our study applies regression on a large synthetic dataset without optimization [7].

A study sought to enhance slope stability classification through the use of deep learning techniques (GAN, LSTM) and an innovative feature selection method (bGGO). It utilized 627 real-world samples and incorporated data pre-processing, deep learning, and Random Forest for assessing feature importance. The bGGO-GAN model reached the highest accuracy (91.3%) and AUC (0.9285), pinpointing cohesion and unit weight as significant factors. Recognized limitations included a small dataset size, high computational demands, and restricted interpretability. Both studies leverage machine learning for analysing slope stability and feature importance. Our research employs regression on a large synthetic dataset using tree-based regressors, whereas the referenced paper focuses on classification utilizing deep learning and metaheuristic optimization [8].

Another research focused on enhancing slope stability forecasts through ensemble machine learning models by integrating bagging and boosting with foundational learners such as Decision Trees and Random Forests. Utilizing 125 real-world samples, it employed both classification and regression methods with a total of 12 machine learning models and applied dimensionality reduction using KPCA. Bagging with Decision Trees attained a classification accuracy of 96%, while LassoLarsCV, Random Forest, and MLP demonstrated strong performance in regression ($R^2 > 0.90$).

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Limitations of the study included a small dataset, high computational demands, and sensitivity to parameter tuning. Both approaches utilize ensemble tree-based regression models to assess feature significance and prediction accuracy. This research utilizes a limited real dataset and examines classification, regression, and dimensionality reduction, whereas our project employs a larger synthetic dataset and concentrates solely on regression [9].

A study focused on forecasting slope stability through the use of sophisticated machine learning models (ANN, GMDH, SGD, CN2) with non-dimensional geotechnical parameters, aiming to overcome the drawbacks of conventional techniques. Utilizing 296 actual records, the models were assessed using MAE, RMSE, and R^2, alongside conducting sensitivity analysis. The ANN model exhibited the highest performance ($R^2 = 0.946$, RMSE = 0.46), with GMDH following closely ($R^2 = 0.926$). Identified limitations included a small dataset, challenges with model interpretability, and sensitivity to input parameters. Both studies estimate the factor of safety (FOS) using regression-based machine learning techniques with geotechnical characteristics. This paper utilizes a limited real dataset and ANN/GMDH models, while our project employs a larger synthetic dataset and tree-based regressors with categorical variables [10].

III.METHODOLOGY

The experimental process starts by pre-processing the dataset including encoding the dataset to numerical representation from categorical values, as shown in Fig. 1, scaling the data values, splitting the dataset for training and testing for further model evaluation, identifying the target and dependent variables. Six tree based regressors are used in this study in order to understand the appropriate model for the problem and understand the tree dependent feature importance. The regressor models are: Decision Tree Regressor, Random Forest Regressor, Extra Trees Regressor, AdaBoost Regressor, Gradient Boosting Regressor, XGboost Regressor. The training dataset is trained using all the 6 regressors, and the testing datasets are then injected. The accuracy Metrices extracted from each regressor model are noted. The performance analysis dependent matrices are: RMSE (Root mean square error), MAE (Mean Absolute Error), R^2 Score, Execution Time (in seconds), and feature importance plot from each regressor model to understand a clear view about the most influential feature affecting the whole model while execution. All the performance matrices are jotted down and analysed for further analysis purpose.

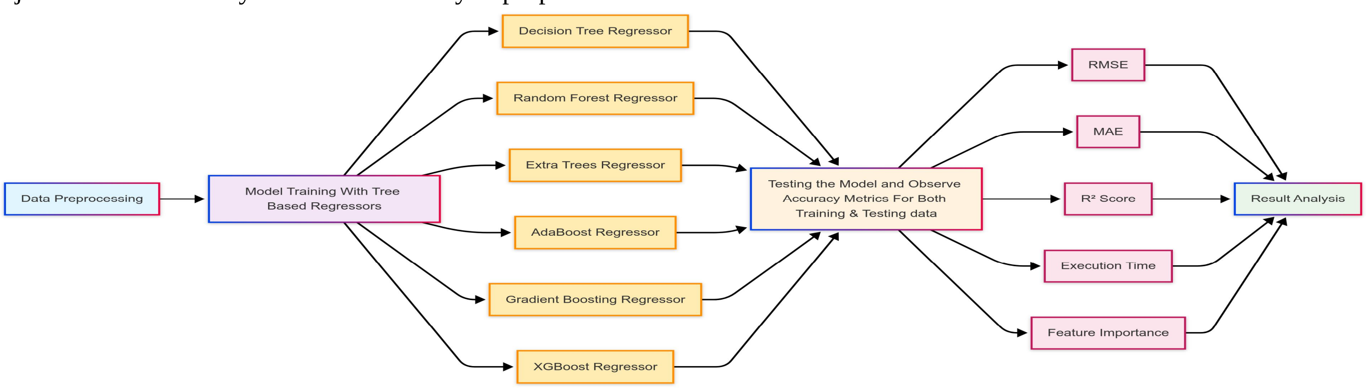

Fig. 1 Process Flow Diagram of Experimentation

This dataset designed for slope stability analysis and optimization in geotechnical engineering, is available in Kaggle containing key geotechnical parameters that influence the stability of slopes and evaluates the Factor of Safety (FS) for different reinforcement techniques. With 10,000 samples featuring 8 geotechnical and 1 reinforcement type parameter, Factor of Safety (FS) being the target variable indicating slope instability if the value is less than 1 and if the value is greater or equal than 1 indicating slope stability. Other features are:

1) Unit Weight (kN/m³) indicating density of soil/rock material
2) Cohesion (kPa): Shear strength due to soil bonding
3) Internal Friction Angle (°): soil shear resistance angle
4) Slope Angle (°): Indicating the inclination of the slope
5) Slope Height (m): Vertical height of the slope
6) Pour Water Pressure Ratio representing water pressure impact
7) Reinforcement Type of slope reinforcement
8) Reinforcement Numeric: Encoded version of reinforcement type

Including the target variable, except Reinforcement Type, all the variables are continuous in nature. Reinforcement Type is a categorical column and Reinforcement Numeric is the encoded representation of the column.

The Dataset is applicable in the fields of geotechnical engineering in predicting and getting a clear view in slope failure risks. Training predictive models for slope stability assessment can be a sector that can integrate both geotechnical engineering and artificial intelligence.

Decision trees are a simple way to make decisions based on the experimental data. Decision tree represents a tree like structure where each node represents the decision based on the feature each branch represents an outcome of that decision and each Leaf node represents the final prediction or the decision. Decision trees are built through a process of recursively splitting the data into subsets based on the most informative feature where the goal is to creating a tree which can make accurate predictions on new and unseen data. They are quick simple and interpretable and used in both classification and regression tasks making them super versatile.

Random Forest is a supervised machine learning algorithm that is used widely in classification and regression problems that leverages an ensemble of multiple decision trees to generate predictions for the data. One of the most important features of the random Forest algorithm is that it can handle the data containing both continuous variables and categorical variables that working of random Forest algorithm. It uses Ensemble learning technique that simply means combining multiple models. Ensemble uses two types of methods bagging and boosting. Bagging creates a different training subset from sample training data with replacement, the final output is based on majority voting. Boosting combines weak Learners into strong Learners by creating sequential models such that the final model has the highest accuracy steps involved in random Forest algorithm

a) In the random forest model a subset of data points and a subset of features is selected
b) Individual decision trees are constructed for each sample
c) Each decision tree will generate an output
d) Final output is considered based on majority voting or averaging

Extra Tree regressor, an extension of random forest regressor, builds multiple trees using random feature selection and random splits unlike traditional methods. They don't waste time finding the perfect split, this means they can train much faster but speed doesn't always equal accuracy. Proper tuning is still essential for the best results that can help reduce overfitting by introducing randomness. but remember it doesn't guarantee better performance.

AdaBoost, an ensemble machine learning model applicable for both classification and regression problems build a strong classifier by combining multiple poorly performing classifiers so that you will get high accuracy strong classifier popularly used in both classification and regression problems. The basic concept behind Adaboost is to set the weights of classifiers and training the data sample in each iteration such that it ensures the accurate predictions of unusual observations. Adaboost works withing 2 conditions, one the classifier or regressor should be trained interactively on various weighed training examples. Another one is in each iteration; it tries to provide an excellent fit for these examples by minimizing training error. Adaboost works initially by selecting a training subset randomly. It iteratively trains the AdaBoost machine learning model by selecting the training set based on the accurate prediction of the last training. It assigns the higher weight to wrong classified observations so that in the next iteration these observations will get the high probability for classification. Also, it assigns the weight to the trained classifier in each iteration according to the accuracy of the classifier. The more accurate classifier will get high weight. This process iterates until the complete training data fits without any error or until reached to the specified maximum number of estimators.

Gradient boosting is a powerful machine learning technique used for both regression and classification task. It is a type of Ensemble learning that combines multiple weak models, typically decision trees to create a strong predictive model. The central idea is to build model sequentially where each new model attempts to correct the errors made by the previous models. Gradient boosting Works by making small corrections to improve the overall prediction. Gradient boosting is widely used in various domains due to its high predictive accuracy and ability to handle different types of data including missing values and outliers.

XGBoost or Extreme Gradient Boosting is a powerful machine learning algorithm popularly used in both classification and regression models uses weak decision trees as its base Learners to improve predictions iteratively. The process begins with an initial prediction, typically the mean of the target variable new trees is subsequently created by focusing on minimizing the residuals or errors left from the previous steps. XGBoost employs gradient descent to update the model where the gradient indicates the direction for optimal loss reduction. Split points in the decision tree are determined based on gain which measure the improvement in the loss function. The gain is computed using gradients and Hessians which represent the second derivatives of the loss function indicating error surface curvature. Pruning is applied to prevent unnecessary splits by ceasing tree growth. When the gain Falls below a certain threshold regularization further mitigates overfitting by imposing penalties on large leaf weights and the total number of leaves in a tree shrinkage or learning rate is utilized.

IV.RESULT ANALYSIS

A. Observing Regression line based on individual Tree based regressors

The dataset for Slope Stability Analysis is a regression-oriented dataset where the target column is of continuous value, so each regression model-based regression curves based on training and testing scatter plots with perfect prediction lines were needed to be observed to get a clear observational analysis for the dataset, along with the residual plots to check for patterns in prediction errors. As shown in Fig. 2, Fig. 3, Fig. 4, Fig. 5 and Fig. 6 shows individual tree regressor based model's training, testing scatter plot along with a linear regression curve which best fits and aligns with the scatter points. For the train and test data scatter plot regression curve the axes are denoted as "Actual Factor of Safety" and "Predicted Factor of Safety" and for residual plots the both axes denoted as predicted factor of safety and residues. In all the 6 regressors, linearity was observed, pointing on the models best fit curves accuracy.

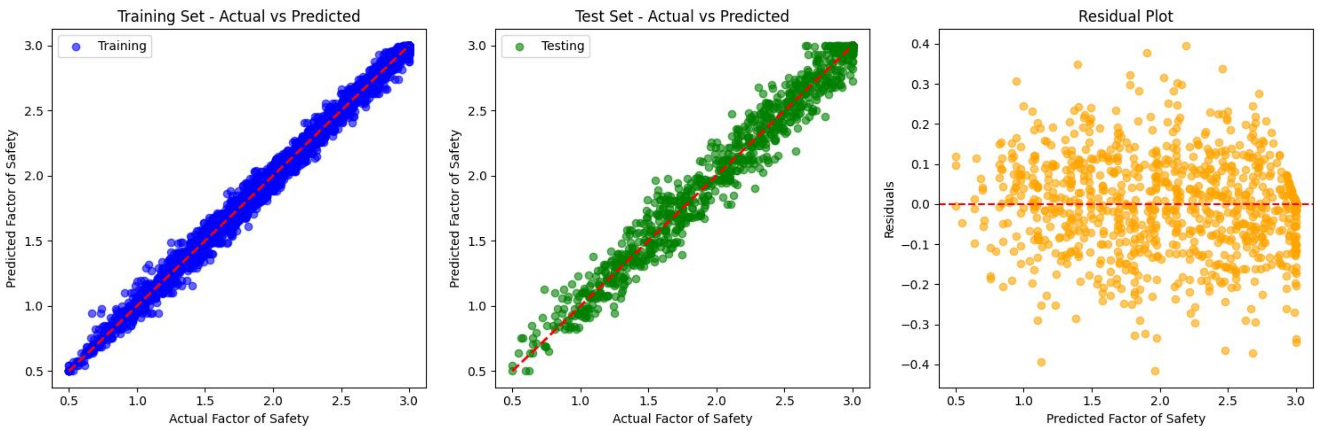

Fig. 1 Decision Tree Regressor Curve analysis

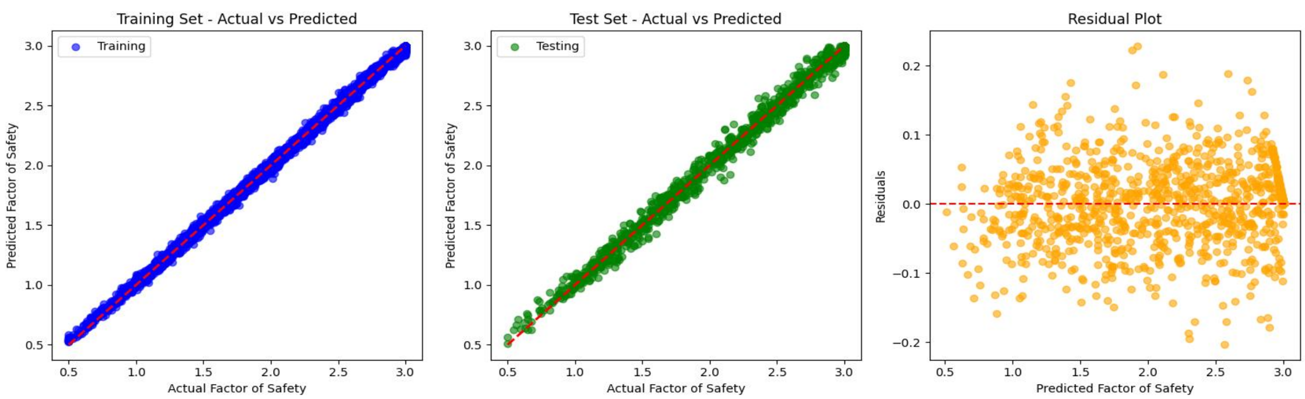

Fig. 3 Random Forest Regressor Curve analysis

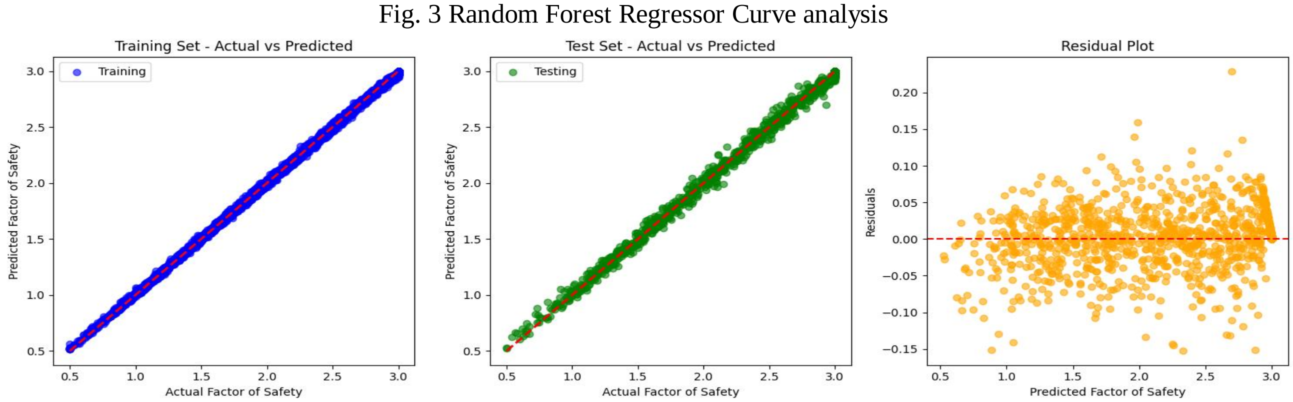

Fig. 4 Extra Tree Regression Regressor Curve analysis

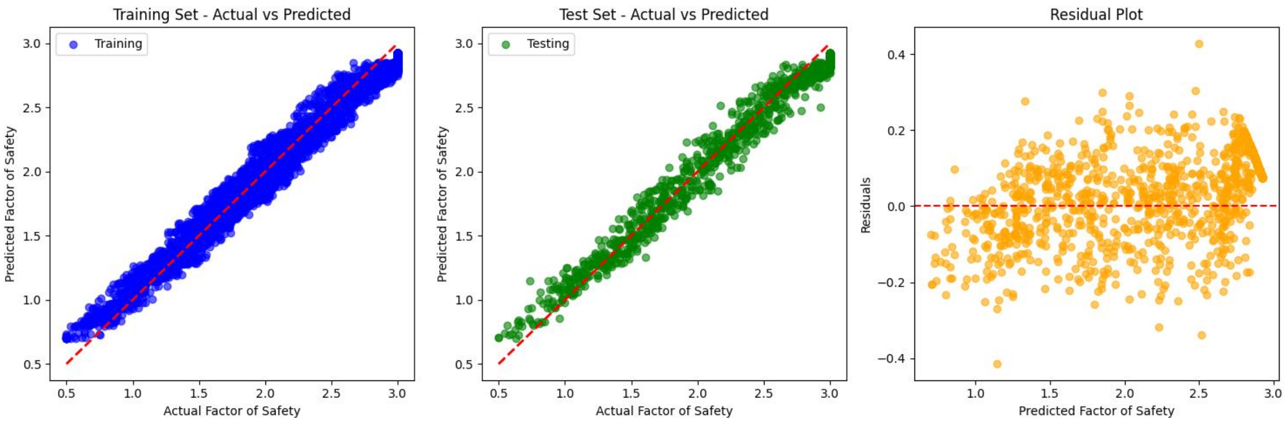

Fig. 5 AdaBoost regressor Curve analysis

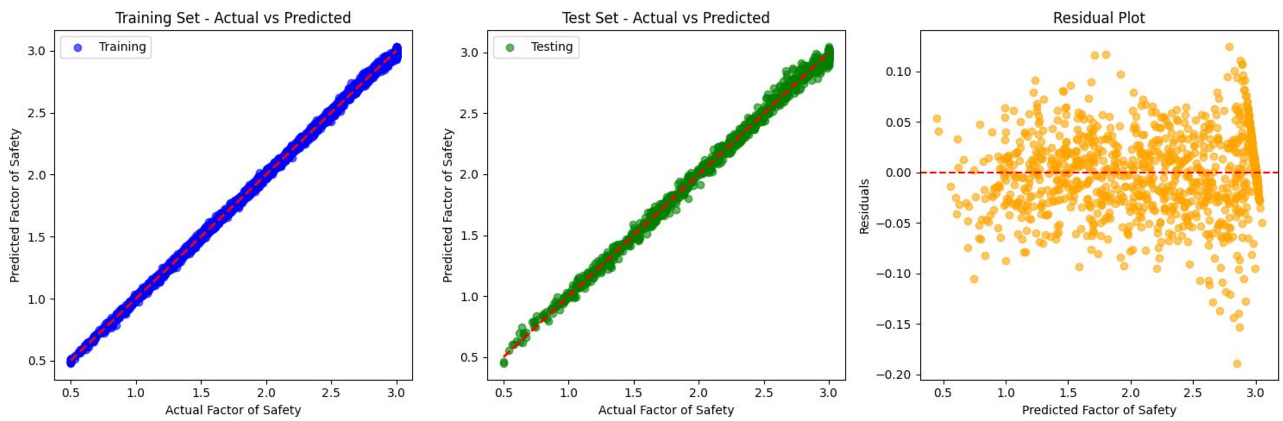

Fig. 2 Gradient Boosting Regressor Curve Analysis

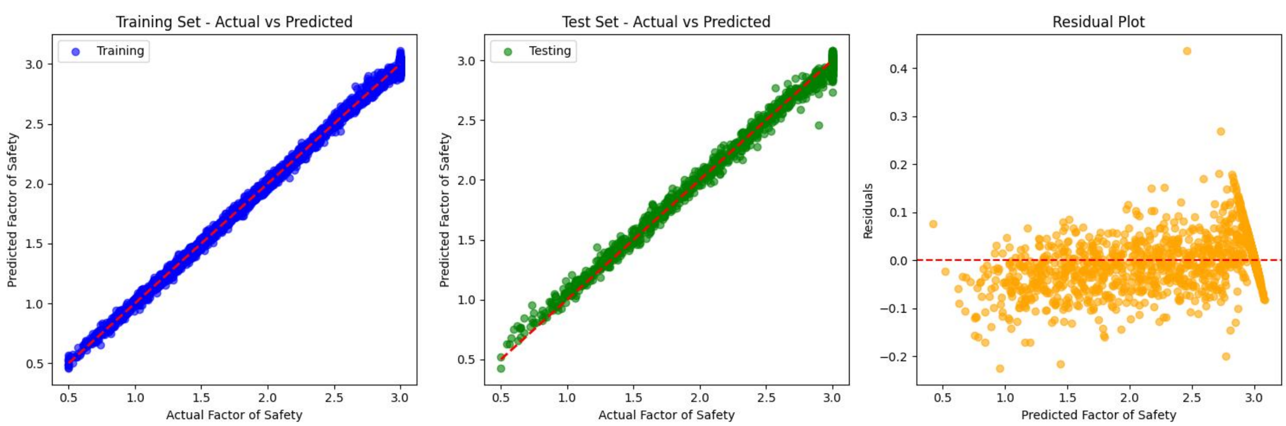

Fig. 6 XGBoost Regressor Curve Analysis

B. Model Performance Analysis

Analysing the performance of tree based regressors in order to observe which regressor performed the best in order to predict the Slope Stability status. As the dataset focused on evaluating regressor models, the accuracy metrics were focused on analysing the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R^2 Score and model execution time in seconds in order to evaluate the models.

Table 1 is a table representing all the metrics value for both training band testing data in order to observe the performance metrics where the optimal outcome is in bold print. Along with tabular representation Fig. 7 gives a corresponding visual representation for an easier comparison view.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Among multiple Root Mean Squared Error (RMSE) values, the smallest value indicates the best model performance. A lower RMSE value means the model's predictions are closer to the actual values, implying a better fit to the data and more accurate predictions. From the table and Fig. 7

Among MAE values, the best or lowest Mean Absolute Error value indicates the most accurate model. due to MAE being a negatively oriented. metric where lower value represents better performance. As MAE measures the average absolute difference between predicted and actual values, a value closer to zero signifies a more precise prediction.

Among multiple R^2 values, the higher the value, the better the model's fit, as R^2 (Coefficient of Determination) being a statistic measuring how well a model fits the data. From the pictorial representation in finding the highest R^2 value, observations were noticed in Extra Tree regressors for both training and testing data, showed the highest R^2 value, indicating better model fitting for the data training and evaluation.

From both tabular and pictorial representation, though all the regressors showed apparently very high accuracy metrics indicating all the models performing pretty well, it is observed that Extra Trees Regressor was the best performing model in terms of a lower RMSE and MAE value tending towards zero, and higher R^2 value amongst all the tree based regressors for both the train and test data.

It is observed that the basic most tree-based algorithm that is the decision tree regressor, showed the least execution time, in contrary with Gradient Boost algorithm being the most time executed for in both training and testing the whole data session.

TABLE I

PERFORMANCE ANALYSIS AMONGST INDIVIDUAL REGRESSOR ON TRAINING AND TESTING DATA

	Metrics	Tree Based Regression Models					
		Decision Tree Regressor	**Random Forest Regressor**	**Extra Trees Regressor**	**AdaBoost Regressor**	**Gradient Boosting Regressor**	**XGboost Regressor**
Training Dataset	RMSE	0.035945	0.018725	**0.011988**	0.119337	0.013871	0.025091
	MAE	0.018043	0.010117	**0.006544**	0.108413	0.008173	0.018901
	R^2 Score	0.997047	0.999199	**0.999671**	0.967447	0.99956	0.998561
	Execution Time (s)	**0.101897**	4.243867	1.075031	2.826697	5.774426	1.11519
Testing Dataset	RMSE	0.08217	0.042208	**0.030442**	0.121375	0.028866	0.046802
	MAE	0.043495	0.022878	**0.016002**	0.109863	0.016364	0.032491
	R^2 Score	0.984206	0.995833	**0.997832**	0.965539	0.998051	0.994876
	Execution Time (s)	**0.101897**	4.243867	1.075031	2.826697	5.774426	1.11519

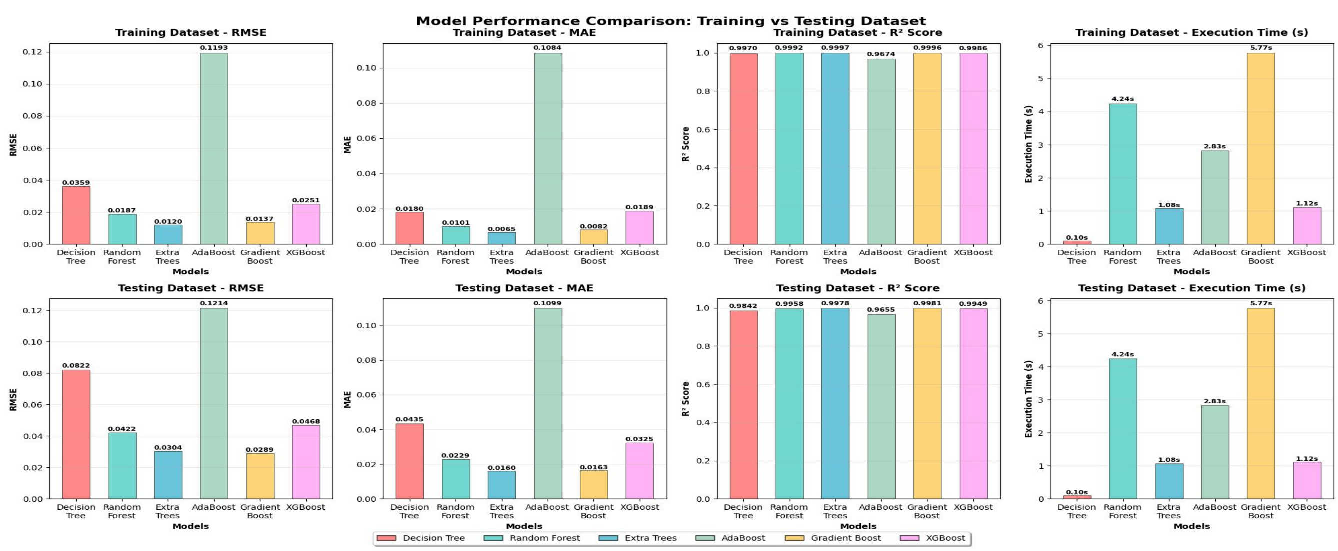

Fig. 7 Visual Representation of performance metrics on individual regressors.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

C. Feature Importance Observation

One Final Observation in this research was to observe the feature importance that will indicate which features were influencing in figuring out the slope stability. Fig. 8, Fig. 9, Fig. 10, Fig. 11, Fig. 12, Fig. 13 shows individual tree based regressor's feature importance in Descending order where he topmost features carries, he highest feature importance scores and Fig. 14 depicts XGBoost's Built-in feature importance plot. From t6he feature importance plot it was observed that the factors naming Cohesion, Internal Friction Angle, Slope Angle, Pour Water Pressure Ratio these 4 features were the most important, influential and concurrent featur4es that came after regressor testing for all the regressors including while in model training and in XGBoost's Built-in feature importance plot. The other features were minor in feature importance relevance.

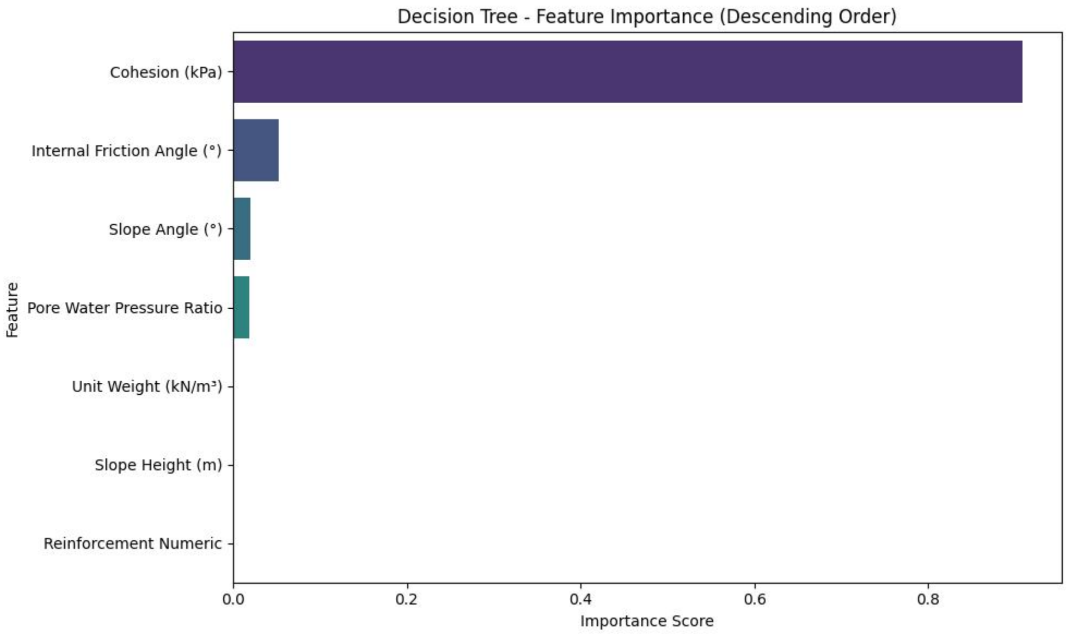

Fig. 8 Decision Tree Regressor Based Feature Importance score in Descending Order.

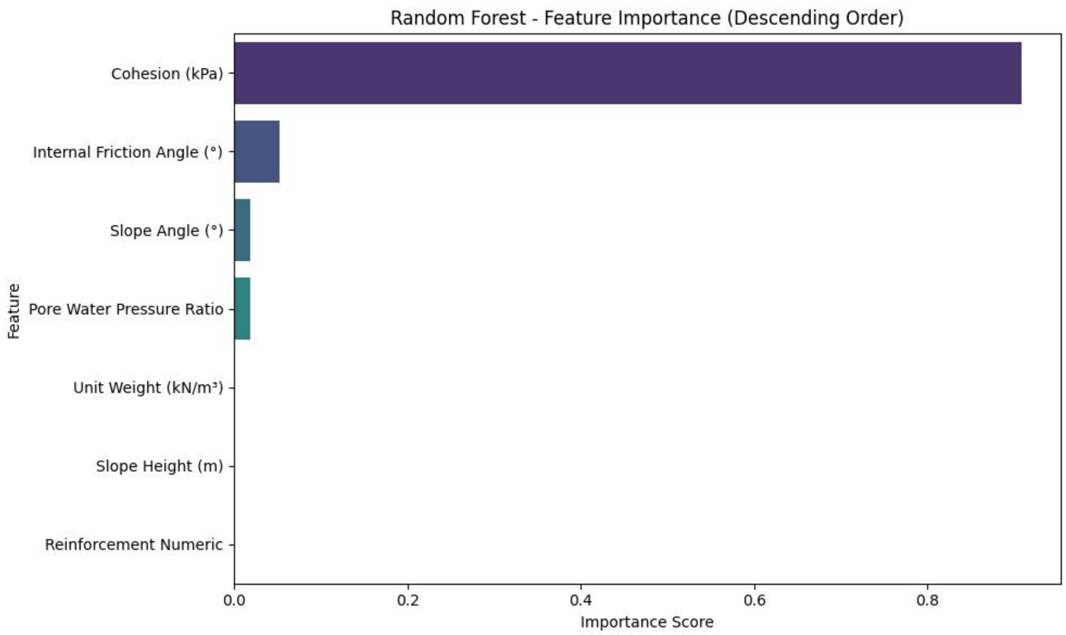

Fig. 9 Random Forest Regressor Based Feature Importance score in Descending Order.

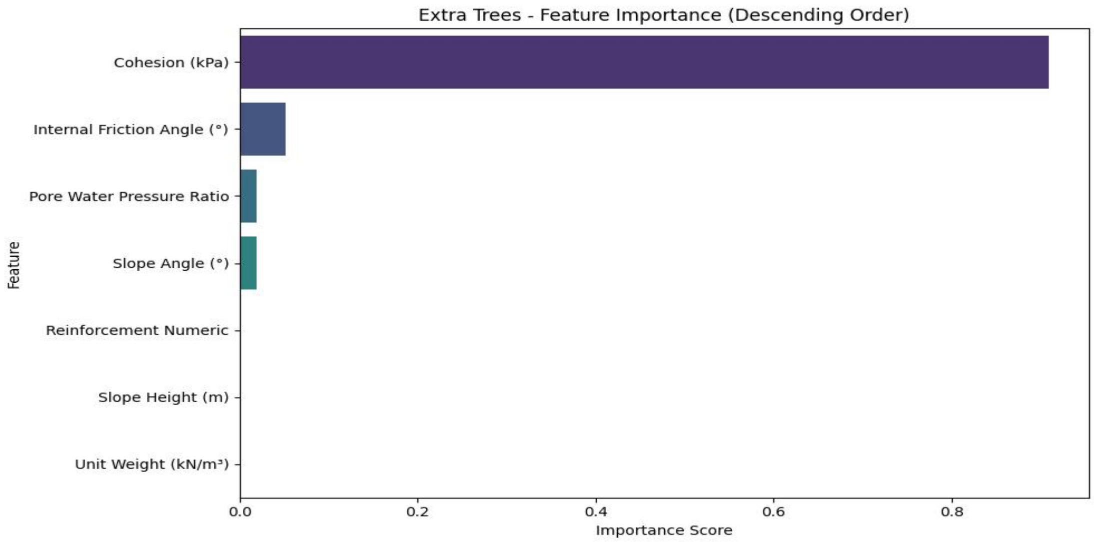

Fig. 10 Extra Trees Regressor Based Feature Importance score in Descending Order.

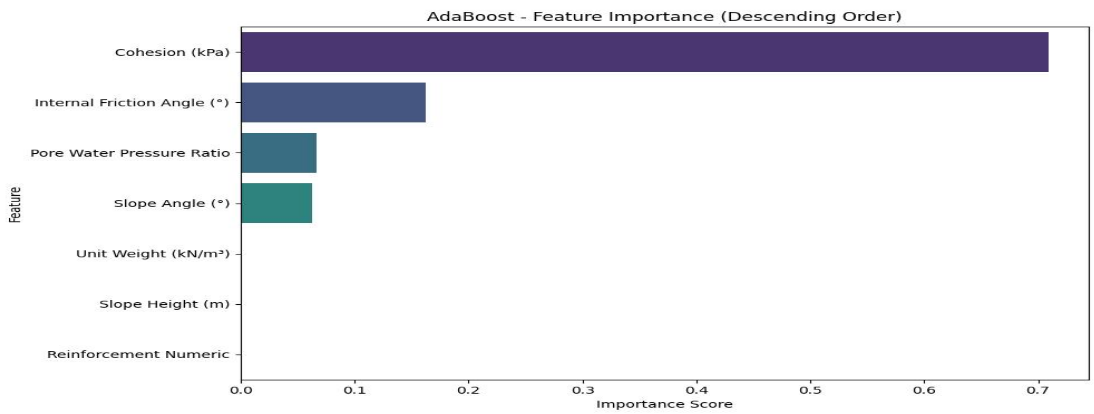

Fig. 11 AdaBoost Regressor Based Feature Importance score in Descending Order.

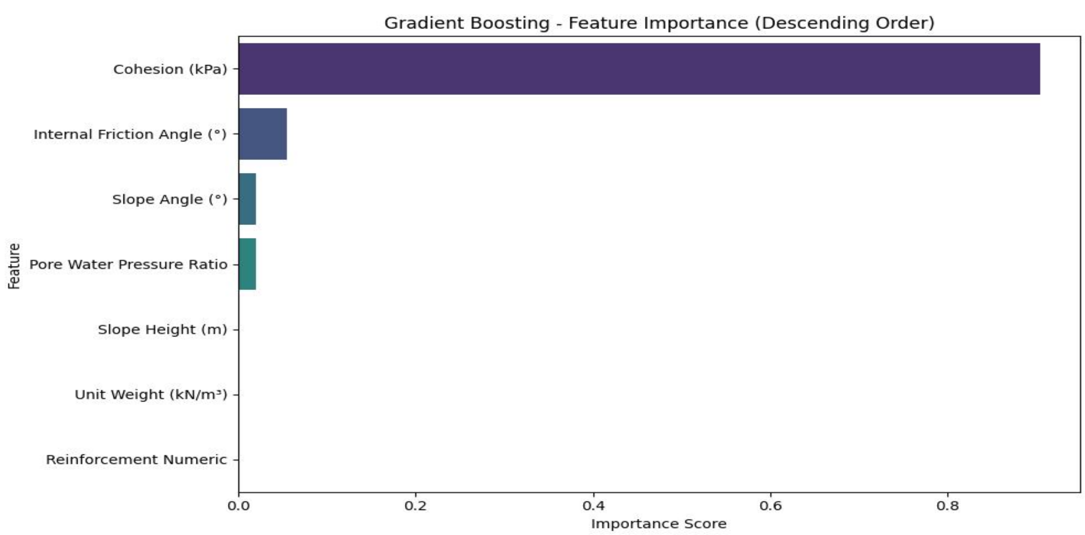

Fig. 12 Gradient Boosting Regressor Based Feature Importance score in Descending Order.

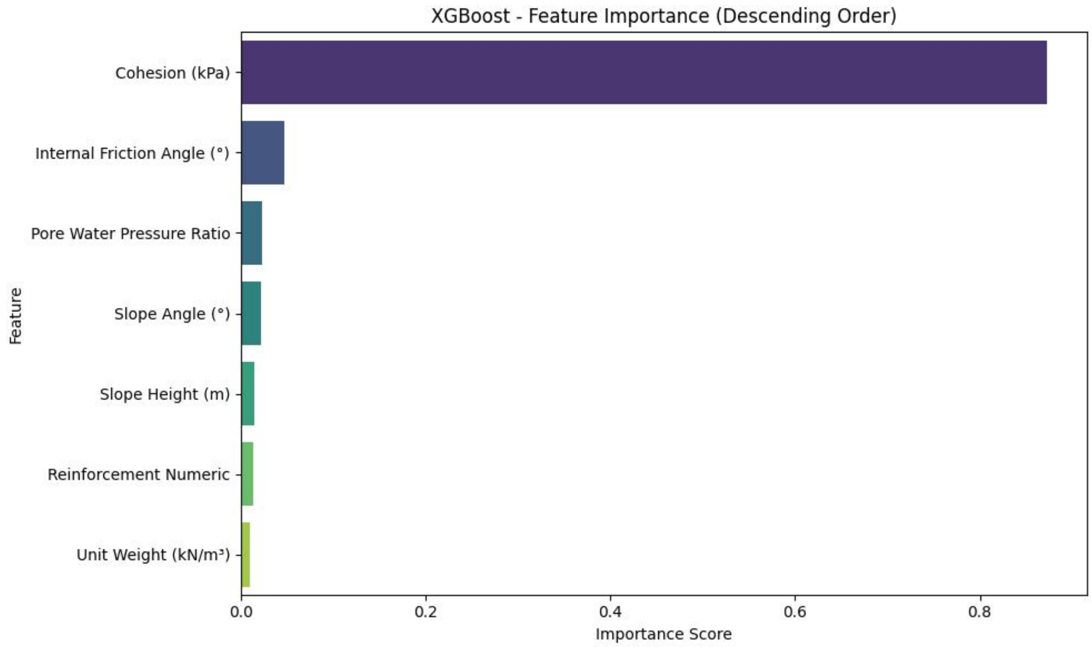

Fig. 13 XGBoost Regressor Based Feature Importance score in Descending Order..

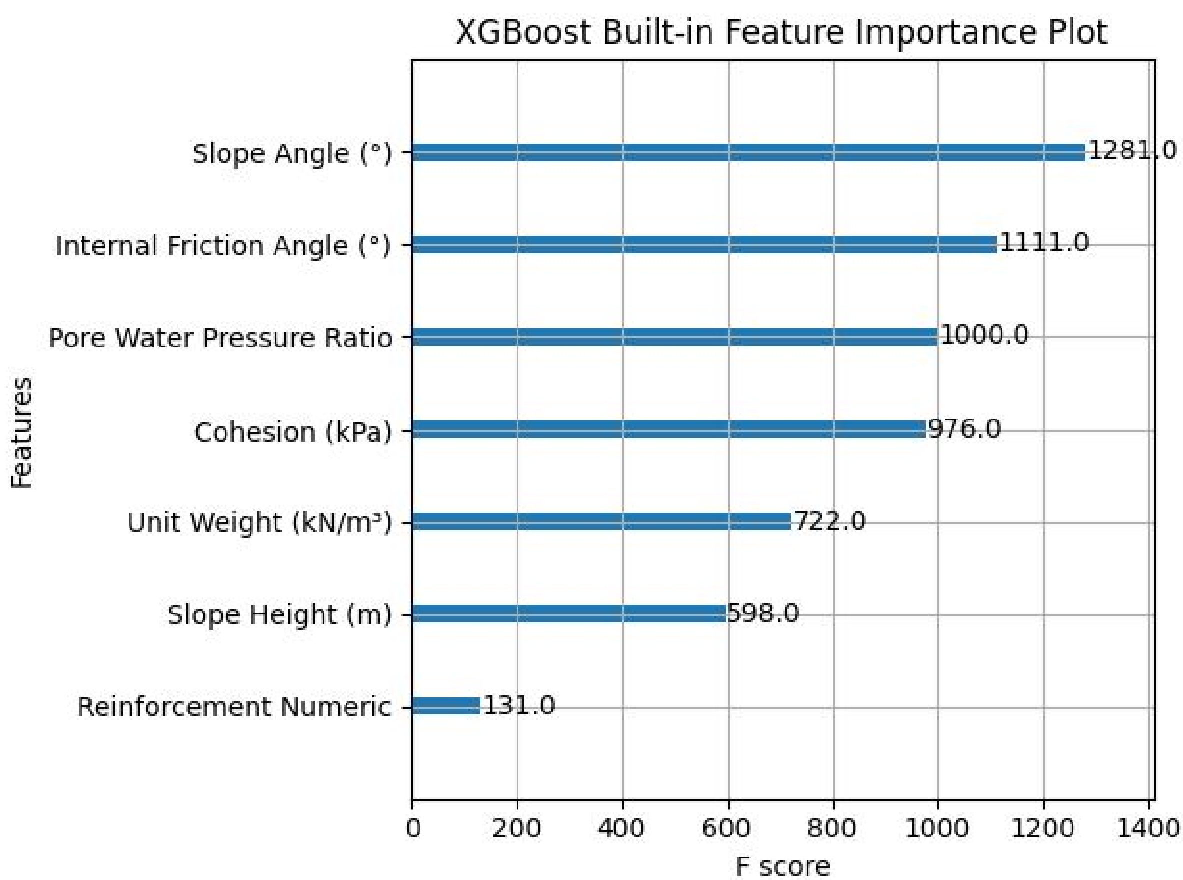

Fig. 14 XGBoost Regressor's Built-in Feature Importance Score in Descending order.

V. FUTURE WORKS

Artificial Neural Network based regression models can be analysed to observe the regression curves and feature importance. Various Meta-heuristic approaches like: Swarm Intelligence, various nature inspired algorithms or XAi (Explainable Artificial Intelligence) can be implied to understand the datasets and features in depth analysis. A small web application can be developed top easily calculate and visualize the slope stability.

VI. CONCLUSION

This study illustrates the effectiveness of tree-based regressors in precisely predicting the Factor of Safety for slope stability analysis. Among all the models, the Extra Trees Regressor consistently reached the highest implementation over RMSE, MAE, and R^2 metrics. Feature significance study announced that cohesion, internal friction angle, slope angle, and pore water pressure ratio were the most influential geotechnical factors. The outcomes emphasize the perspective of integrating machine learning with geotechnical engineering for apocalyptic modelling. Future work can expand on this by incorporating additional soil conditions and real-time monitoring data for upgraded conception.

REFERENCES

[1] W. Zhang, H. Li, L. Han, L. Chen, and L. Wang, "Slope stability prediction using ensemble learning techniques: A case study in Yunyang County, Chongqing, China," Journal of Rock Mechanics and Geotechnical Engineering, vol. 14, no. 4, pp. 1089–1099, 2022.

[2] A. Kainthola, V. H. R. Pandey, G. Kushwaha, and V. Yadav, "Role of Feature Importance in Geomechanical Classification of Rock Slopes," Deep Resources Engineering, pp. 100165, 2025.

[3] Z. Xu, X. Wang, L. Guo, and T. Yu, "Stability Evaluation of Slope Based on Global Sensitivity Analysis," Complexity, vol. 2024, no. 1, pp. 2333859, 2024.

[4] P. Ragam, N. Kushal Kumar, J. E. Ajith, G. Karthik, V. K. Himanshu, D. Sree Machupalli, and B. Ramesh Murlidhar, "Estimation of slope stability using ensemble-based hybrid machine learning approaches," Frontiers in Materials, vol. 11, pp. 1330609, 2024.

[5] F. Ahmad, X. Tang, J. Hu, M. Ahmad, and B. Gordan, "Improved prediction of slope stability under static and dynamic conditions using tree-based models," parameters, vol. 1, no. 3, 2023.

[6] A. K. Sahoo, D. P. Tripathy, and S. Jayanthu, "Application of machine learning techniques in slope stability analysis: A comprehensive overview," Journal of Mining and Environment, vol. 15, no. 3, pp. 907–921, 2024.

[7] M. H. Kadkhodaei, E. Ghasemi, and M. H. Fazel, "Slope stability prediction of circular mode failure by machine learning models based on Bayesian Optimizer," Journal of Mountain Science, vol. 22, no. 4, pp. 1482–1498, 2025.

[8] B. Zerouali, N. Bailek, A. Tariq, A. Kuriqi, M. Guermoui, A. H. Alharbi, D. S. Khafaga, and E.-S. M. El-Kenawy, "Enhancing deep learning-based slope stability classification using a novel metaheuristic optimization algorithm for feature selection," Scientific Reports, vol. 14, no. 1, pp. 21812, 2024.

[9] D. K. Yadav, S. Chattopadhyay, D. P. Tripathy, P. Mishra, and P. Singh, "Enhanced slope stability prediction using ensemble machine learning techniques," Scientific Reports, vol. 15, no. 1, pp. 7302, 2025.

[10] K. C. Onyelowe, A. M. Ebid, S. Hanandeh, and V. Kamchoom, "Evaluating the slope behavior for geophysical flow prediction with advanced machine learning combinations," Scientific Reports, vol. 15, no. 1, pp. 6531, 2025.