JURNAL EMACS
[Engineering, MAthematics and Computer Science] Vol.8 No.2 May 2026: 121-132
e-ISSN: 2686-2573
DOI: 10.21512/emacsjournal.v8i2.16161

Integrating Finite Element Analysis and Machine Learning to Predict the Bearing Capacity of Strip Footings on Slopes

Aditya Dwi Kurniawan[1*], Lindung Zalbuin Mase[2], Muharram Nur Fikri[3], Rena Misliniyati[4], Aidil Fitriansyah[5]

[1-5]Department of Civil Engineering, Faculty of Engineering,
University of Bengkulu,
Bengkulu 38371, Indonesia
g1b022029.adityadwik@mhs.unib.ac.id; lmase@unib.ac.id; m.nurfikri@unib.ac.id; rena_misliniyati@unib.ac.id; aidil.fitriansyah@unib.ac.id

*Correspondence: g1b022029.adityadwik@mhs.ac.id

Abstract – *Predicting the ultimate bearing capacity (qult) of strip footings on slopes remains a major challenge in geotechnical engineering, as classical methods were developed for level ground and lose reliability for complex slope-foundation geometries. Although finite element analysis (FEA) provides better accuracy, its high computational cost limits large-scale parametric studies. This study proposes a hybrid FEA–machine learning (ML) framework to estimate qult of strip footings on slopes, overcoming the accuracy limitations of existing analytical solutions for different slope-foundation configurations. The dataset of 600 finite element simulations was developed under the Mohr-Coulomb plane strain constitutive framework. Six variables were examined: unit weight (γ), cohesion (c), friction angle (φ), applied load (P), foundation width (B), and embedment depth (Df). Seven predictive models were developed: multiple linear regression, polynomial regression, support vector regression, decision trees, random forests, k-nearest neighbors, and extreme gradient boosting (XGBoost). Model performance was assessed using R², RMSE, MAPE, and the a20 index, with R² and RMSE as the primary ranking criteria, while Shapley Additive Explanations (SHAP) were applied to interpret feature contributions. XGBoost has the highest prediction accuracy on both the training and test datasets. It is followed by Support Vector Regression (SVR). The most influencing parameter in all seven models was the foundation depth (Df), followed by the friction angle (φ) and the foundation width (B), while the slope angle consistently decreased the predicted bearing capacity. The results confirm the accuracy, interpretability, and computational efficiency of the integrated FEA-ML approach as an alternative to traditional bearing capacity analysis.*

Keywords: *Bearing capacity; Strip footing; Slope; Finite element analysis; Machine learning*

I. INTRODUCTION

A major difficulty in geotechnical engineering is the bearing capability of weak foundations on or near sloping ground. Presence of slope decreases the confining resistance of the soil and changes the failure mechanism as compared to footing on the level ground (Moayedi & Hayati, 2018; Acharyya, 2019). Field investigations of riverbank slope failure during flooding events show that slope instability causes significant financial losses and damage to nearby infrastructures (Mase et al., 2023). Mistaken calculation of the ultimate bearing capacity (q_{ult}) of strip footings on slopes may lead to uneconomical over-design or more gravely foundation failure and overall slope instability. Determining q_{ult} under such conditions therefore requires methods that can adequately capture the combined influence of footing geometry, soil strength parameters, and slope inclination.

Traditionally, the ultimate bearing capacity, q_{ult}, has been approximated utilizing analytical theories of bearing capacity similar to that of Terzaghi, Meyerhof and Hansen that are still commonly utilized in the early stages of design. (Acharyya, 2019). However, these closed-form solutions were originally developed for level ground, offering simplifying assumptions that tend to decrease the accuracy as the slope geometry becomes more complex.

Therefore, numerical methods, in particular finite element analysis (FEA), are becoming more and more popular, since they can simulate the nonlinear soil behavior and the interaction between footing and slope more realistically and can distinguish between an overall slope failure and a localized bearing capacity failure (Moayedi & Hayati, 2018; Acharyya, 2019). This capability has also been demonstrated for slope-stability applications where finite element simulation based on the shear strength reduction method has been shown to reproduce observed failure mechanisms, factor of safety distributions, and displacement patterns on riverbank slopes (Mase et al., 2023). In the last few years, three-dimensional finite element limit analysis (FELA) has been applied to the specific problem of rectangular footings resting on cohesive-frictional slopes, where slope angle, footing geometry and soil shear strength parameters were combined in a single stability framework (Le et al., 2025). However, FEA is computationally expensive when a large number of parametric scenarios are to be explored, thus limiting its applicability to fast design iterations. The problem is aggravated by the uncertainty in soil parameters in geotechnical practice.

Recently, ML has emerged as a viable substitute for numerical analysis, capable of approximating the intricate nonlinear interactions governing bearing capacity from input variables. Based on the data from FEA, ML can replace numerical analysis with computational efficiency once the model is trained. (Moayedi & Hayati, 2018; Acharyya, 2019). Different algorithms such as simple regressions, ensemble, and boosting have been applied to solve the geotechnical prediction problem with promising accuracy (Yaghoubi et al., 2024). For instance, a hybrid approach of 3D data obtained from the Finite Element Limit Analysis (FELA) method and an XGBoost model fine-tuned by Ant Colony Optimization has shown a high accuracy in estimating the bearing capacity factor of rectangular foundations on slopes, allowing the development of parametric design charts without the need for repeated numerical simulations (Le et al., 2025). The interpretability of the predictive approaches is becoming an important aspect when comparing different predictive approaches in geotechnical engineering, with a view to moving away from "black box" predictions (Yaghoubi et al., 2024). In the broader machine learning literature, methods such as Shapley Additive exPlanations (SHAP) have been used to interpret the contribution of individual input parameters to the model output, and to validate the physical consistency of the learned relationships. Based on the SHAP, feature importance analysis has shown that the shear strength characteristics of the soil were identified as the dominant variables governing footing capacity on sloped terrain, followed by the footing geometry and slope angle, and gave a physically interpretable basis for the predictions (Le et al., 2025).

But there are no studies combining a large, validated FEA-generated dataset with systematic benchmarking of multiple ML algorithms and SHAP-based interpretability for strip footings on slopes. Most of the previous works either are limited to a small set of algorithms or do not check the physical plausibility of the obtained predictions (Moayedi & Hayati, 2018; Acharyya, 2019; Liu & Liang, 2024). Even the recent hybrid FELA-machine learning studies on footings near the slopes have been generally based on a single optimized algorithm applied to rectangular geometries (Le et al., 2025) or benchmarked across multiple models but restricted to seismic loading conditions on purely cohesive soils (Kumar et al., 2023), leaving open the question of how multiple predictive models compare when applied specifically to strip footing configurations. To bridge the gap, the present study provides a hybrid FEA-ML framework built around 600 finite element simulation runs to train and assess seven prediction models to estimate the final bearing capacity of strip footings on slopes. The predictive performance of each model is evaluated using a variety of statistical measures and a SHAP evaluation is performed to quantify the contribution of individual input parameters. This work provides a dependable forecasting tool and physically important information for supporting foundation design on sloped ground conditions.

II. METHODS

2.1 Research Framework

The whole research workflow used in this study is shown in Figure 1. The study began with an extensive assessment of previous numerical research on the ultimate bearing capacity (q_{ult}) of strip footings on or adjacent to sloped terrain. The assessment of existing literature was undertaken to provide a sound theoretical framework, to identify gaps in the present understanding and to guide the selection of acceptable geotechnical parameters for the next step of numerical modelling. Furthermore, a series of parametric variations were defined to account for the critical physical and geometrical conditions that affect the bearing capacity behaviour, which are: (i) the engineering properties of stiff clay, (ii) the material properties of the foundation, (iii) the slope gradient, and (iv) the magnitude of the applied load. The soil parameters chosen for these variations are derived from well-established empirical correlations as cited by Bowles (1997) and Das & Sivakugan (2018), ensuring that the input parameters are kept within realistic and physically consistent bounds. Prior to the full-scale parametric analysis, the numerical model undergoes validation, with the predicted q_{ult} values compared to previously published numerical analysis studies. This framework then proceeds to the next stage only after satisfactory agreement with the reference data has been achieved, thereby ensuring that the finite element model is sufficiently predictive before generating training data. Numerical runs were constructed by independently varying six key geotechnical parameters, including unit weight (γ), shear strength components (cohesion c and friction angle φ), external applied vertical load (*P*), footing width (*B*) and embedment depth (D_f). This parametric matrix was designed to comprehensively cover the range of field conditions under which strip footings on slopes are constructed, and to produce a robust and physically grounded dataset for the subsequent stage of predictive modelling.

The dataset obtained from the finite element simulation was then used to train a machine learning-based predictive model capable of predicting bearing capacity as a function of the above parameters. Statistical performance indicators were employed to evaluate the predictive accuracy of the models, which provide an objective comparison of the accuracy and the generalization capability of the models.

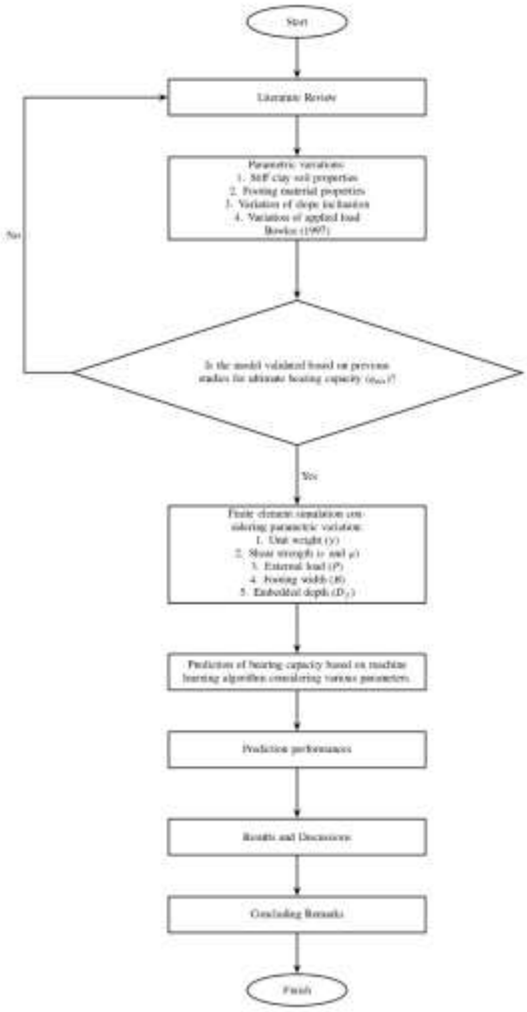


Figure 1. Research framework flowchart illustrating the sequential steps of the hybrid FEA–ML methodology: from literature review and parametric variation definition, through finite element simulation and model validation, to machine learning model training, performance evaluation

These results were integrated and critically discussed in the Results and Discussion section, followed by the Conclusions. In summary, this framework offers a hybrid numerical-data-driven approach where finite element simulations calibrated and validated using existing literature serve as a high-precision data generator; and machine learning models provide a computationally efficient alternative for rapid and reliable bearing capacity estimates for diverse slope and foundation configurations; thus filling the void between rigorous geotechnical numerical analysis and practical, scalable predictive tools for foundation design on sloping terrains.

2.2 Geometrical Model

The Geometrical Finite Element Analysis (Figure 2) has an 8 m height with slope angles (*β*) of 10°, 20°, 30°, and 40°. A strip foundation is placed at the slope edge (conservative condition), with depth (D_f) of 0, 1.0, and 2.0 m; width (*B*) of 1.0, 1.5, and 2.0 m; and loads of 50, 75, and 100 kN. The soil is modeled with Mohr-Coulomb, a typical approach to slope stability, to calculate the load resistance with shear strength. The concrete foundation has a linear-elastic model. In total 600 models were analyzed.

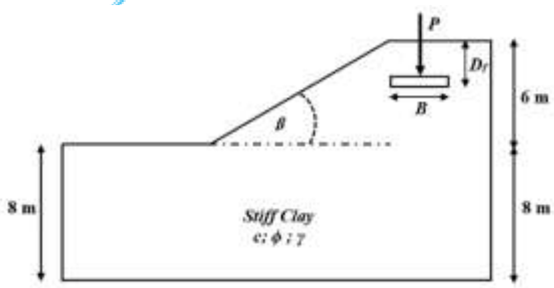


Figure 2. Geometrical configuration of the finite element analysis

2.3 Bearing capacity

The determination of the q_{ult} is of highest relevance in the construction of foundations in sloping terrain. The q_{ult} value and slope stability are particularly essential in geotechnical engineering (Mase et al., 2025). Accurate data and the use of the correct analytical procedures are required to accurately determine the q_{ult} value of the slope in issue (Mase et al., 2025). Strip footing bearing capacity on slopes is generally determined through either closed-form analytical solutions or computational numerical approaches (El-Emam et al., 2023; Mase et al., 2025). Numerical methods have greatly improved in recent years and often give better results than other methods (Mase et al., 2025). These methods can also identify various ways in which a foundation on a slope may fail. It can also help distinguish between overall slope failure and failure of load-bearing capacity (El-Emam et al., 2023). The analytical models of Meyerhof, Terzaghi and Hansen are still in common use especially in the early phases of design (El-Emam et al., 2023; Mase et al., 2025). However, numerical techniques such as finite element analysis (FEA) usually produce better outcomes. These techniques also aid in understanding the interaction between slopes and foundations (El-Emam et al., 2023; Mase et al., 2023).

Most study today favor numerical analysis notably Finite Element Analysis (FEA) (Mase et al., 2025). FEA numerical analysis in geotechnical engineering provides many advantages, for instance saving time and money, and offering accurate answers for solving slope stability problems (Mase et al., 2025). q_{ult} also plays a very essential role in the solution of problem of stability of foundation. The q_{ult} represents the peak load intensity a soil can sustain prior to shear failure (Mase et al., 2025). But geotechnical engineering is intrinsically uncertain (Otake & Honjo, 2022; Mase et al., 2025).

2.4 Finite Element Analysis

The model employed in this investigation was plane strain. Soil behaviour was simulated using Mohr–Coulomb model. This model uses highly refined triangular elements to produce accurate results. The initial conditions were defined without groundwater table and then loads were applied. The maximum load-bearing capacity of a soil or foundation (q_{ult}) can be determined using the equation below:

$$q_{ult} = (SumM - loadA \cdot P) + (\gamma_{conc} \cdot t_{cp}) \qquad (1)$$

q_{ult} (kN/m²) is denoted as the ultimate bearing capacity and $SumM$ denotes the total sum of mobilized vertical nodal reaction forces extracted from the finite element output (kN/m), representing the total vertical resistance mobilized by the soil at the point of failure; $loadA$ is a dimensionless load amplification factor applied within the FEA to scale the external loading during the analysis; and P is the externally applied vertical load per unit length (kN/m). The product $loadA \cdot P$ therefore represents the net additional vertical load imposed on the foundation, so that the term $(SumM - loadA \cdot P)$ reflects the net bearing reaction attributed to the soil's shear resistance. Furthermore, γ_{conc} is the unit weight of concrete (24 kN/m³ is assumed) and t_{cp} is the thickness of the concrete foundation slab, taken as 0.30 m in all simulations. The homogeneous stiff clay slope properties are listed in Table 1.

Table 1. Parameters utilized for FEA

Parameters	Unit	Stiff Clay			Concrete
Unit Weight (γ)	kN/m³	17 to 21	17 to 21	17 to 21	24
Vertical permeability (k_y)	m/day	0.009	0.009	0.009	-
Horizontal permeability (k_x)	m/day	0.009	0.009	0.009	-
Modulus of elasticity (E)	kN/m²	20,000	20,000	20,000	2.35×10^7
Poisson ratio (v)	-	0.2	0.2	0.2	0.15
Soil cohesion (c)	kN/m²	7.5	7.5	7.5	-
Internal friction angle (φ)	°	10	15	25	-
Dilatancy angle (ψ)	°	0	0	0	-
Material model	-	Mohr Coulomb	Mohr Coulomb	Mohr Coulomb	Linear Elastic

2.5 Model Consistency

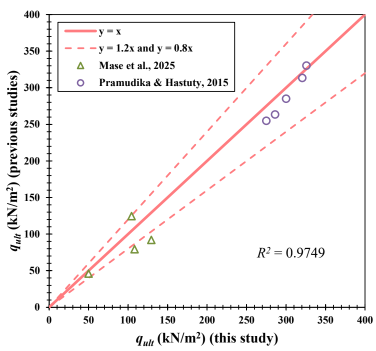


Figure 3. Comparison of predicted ultimate bearing capacity (q_{ult}) values from this FEA model, compared to the reference values reported by Pramudika & Hastuty (2015) and Mase et al. (2025). The solid line is the perfect fit (y = x) and the dashed lines are the ±20% deviation limit. The numerical model is consistent (R^2 = 0.9749).

The performance of the finite element model to describe the identical problem presented in earlier studies. Pramudika & Hastuty (2015) and Mase et al. (2025) found some studies. The variables considered in this study encompass embedment depth, slope angle, as well as the physical and engineering characteristics of the soil.

The consistency analysis indicates an R^2 value of 0.9749. It indicates that the analysis using this new approach is consistent with previous findings. Therefore, this model is suitable to use in this study. Figure 3 shows the consistency of the q_{ult} values. The figure shows that the q_{ult} values are in general consistent with the values from previous studies. This study re-analyzes FEA results to verify the model consistency. Hence, this analysis is performed using FEA.

2.6 Model Performance Evaluation

Each model's predictive capability was assessed through four widely adopted statistical indicators frequently used in geotechnical soft-computing studies, comprising four evaluation metrics: R², RMSE, MAPE, and the a20 index (Mase et al., 2025; Khatti & Grover, 2025). A model with R^2 close to one suggests that most of the variance in the measured bearing capacity is explained, while reduced RMSE and MAPE suggest smaller average prediction errors. If $_{a20}$ index is close to one, it suggests that the model rarely predicts a result that is substantially different from the finite element outcome.

The four performance indicators were R2 and RMSE as the main criteria for model ranking as they both reflect the proportion of variance explained and the absolute magnitude of prediction error, while MAPE and $_{a20}$ index were used as complementary metrics to assess relative prediction accuracy and to quantify the proportion of predictions within an acceptable ±20% deviation from FEA reference values.

Each metric was independently evaluated on both training and testing partitions to assess model generalization and detect potential overfitting (Mase et al., 2025), shown in Table 2.

Table 2. Performance metrics for seven machine learning models.
The approach to evaluating performance metrics follows the protocol described in Mase et al. (2025)

Alghorithms	**Training**				**Testing**			
	RMSE	**R²**	**MAPE**	***a20* Index**	**RMSE**	**R²**	**MAPE**	***a20* Index**
Linear Regression	31.804	0.836	17.833	68.685	34.352	0.802	21.490	68.333
Polynomial Regression	15.620	0.960	8.093	93.111	18.099	0.945	9.290	86.667
SVR	14.171	0.967	5.034	95.825	16.997	0.952	7.840	95.000
Decision Tree	14.948	0.964	7.641	93.319	22.657	0.914	11.342	87.500
Random Forest	15.294	0.962	7.375	93.111	21.082	0.925	10.703	85.833
KNN	21.107	0.928	11.439	84.969	28.702	0.862	16.758	77.500
XGBoost	12.638	0.974	5.981	97.077	16.819	0.953	8.325	92.500

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (3)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (4) \qquad (4)$$

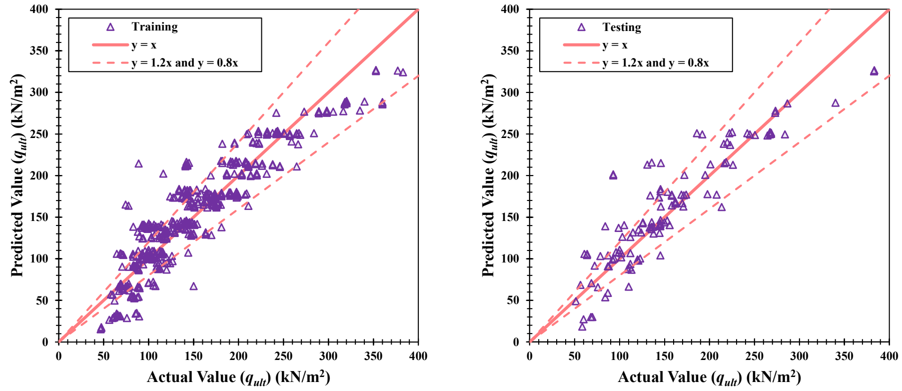
$$a_{20} = \frac{1}{n}\sum_{i=1}^{n} 1\left(0.8 \leq \frac{\hat{y}_i}{y_i} \leq 1.2\right) \times 100\% \qquad (5)$$

III. RESULTS AND DISCUSSION

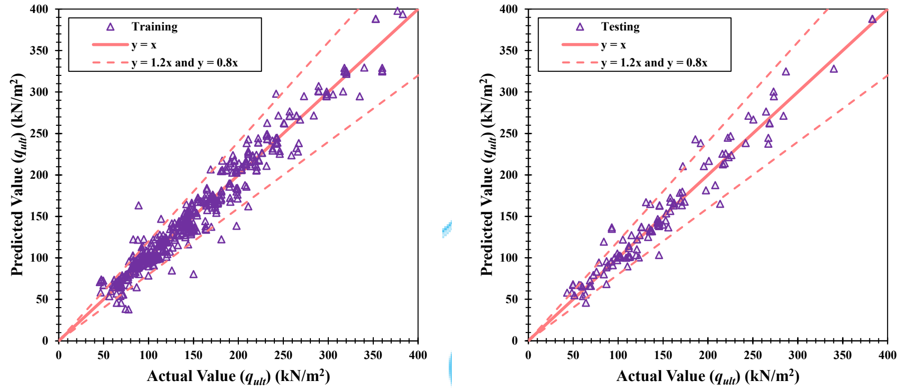
3.1 Finite Bearing Capacity Prediction Based on Machine Learning Algorithm

Figure 4 shows the actual and predicted q_{ult} values for each machine learning model, with a 1:1 line and an error range of 1.2 (upper) and 0.8 (lower).

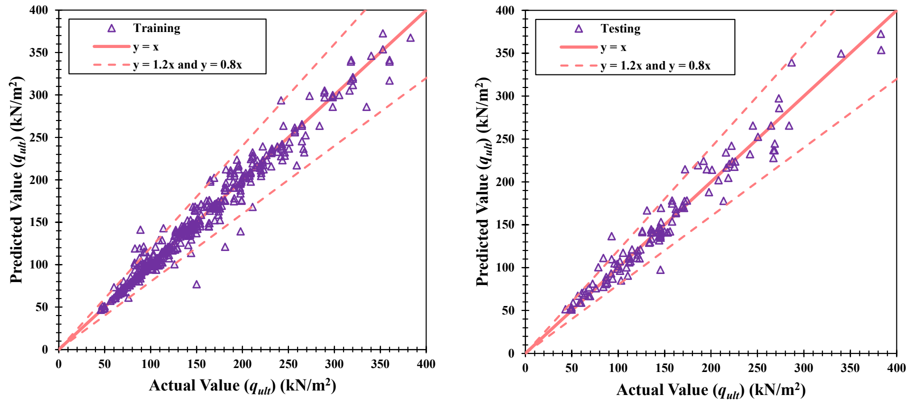
The predictive performance of the seven models was evaluated by four complementary metrics, RMSE, MAPE, R^2 and $_{a20}$ index. The metrics were calculated separately on training and testing datasets, in order to evaluate both fitting accuracy and generalization capability. Tabel 2. Performance Among all seven models, XGBoost achieved the best overall performance on the training


Training
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m²)
Actual Value (q_{ult}) (kN/m²)
Testing
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m²)
Actual Value (q_{ult}) (kN/m²)

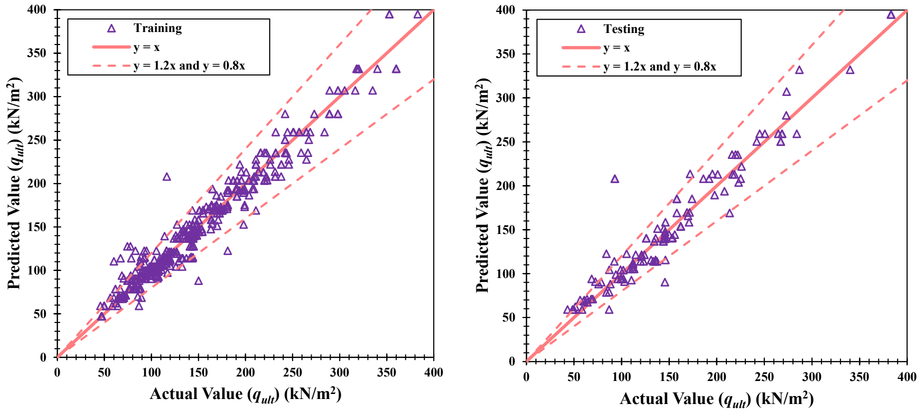
(a)


Training
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m²)
Actual Value (q_{ult}) (kN/m²)
Testing
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m²)
Actual Value (q_{ult}) (kN/m²)

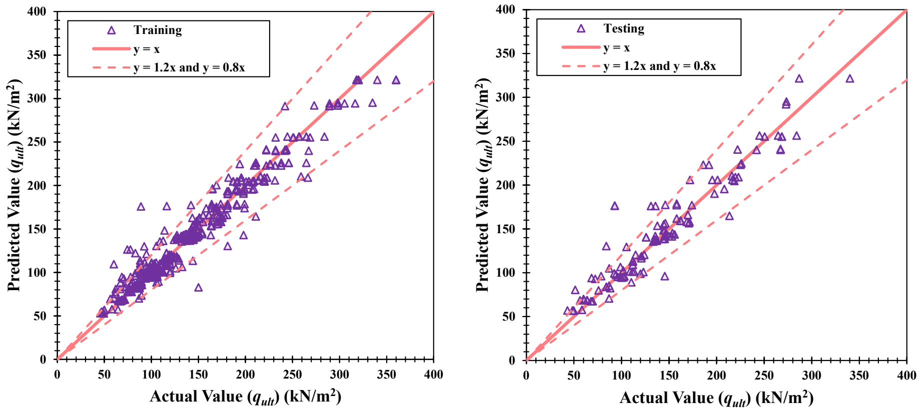
(b)


Training
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m²)
Actual Value (q_{ult}) (kN/m²)
Testing
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m²)
Actual Value (q_{ult}) (kN/m²)

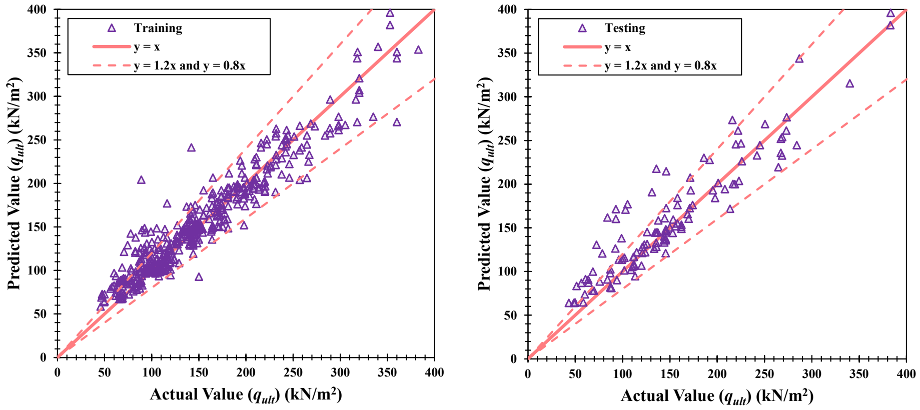
(c)


Training
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m^2)
Actual Value (q_{ult}) (kN/m^2)
Testing
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m^2)
Actual Value (q_{ult}) (kN/m^2)

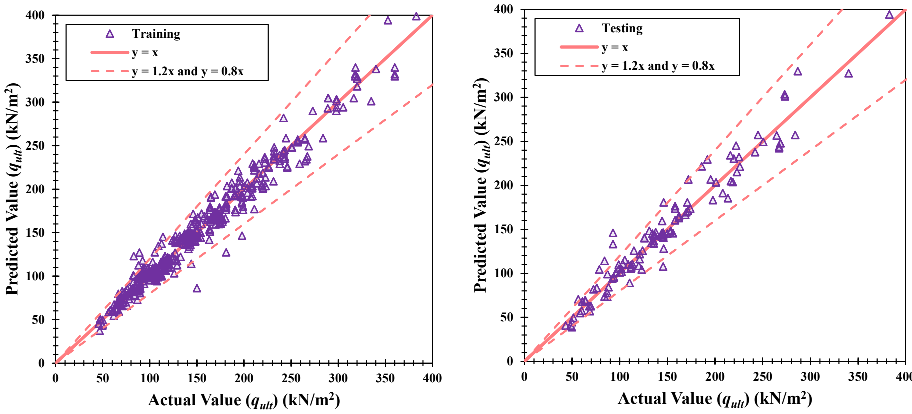
(d)


Training
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m^2)
Actual Value (q_{ult}) (kN/m^2)
Testing
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m^2)
Actual Value (q_{ult}) (kN/m^2)

(e)


Training
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m^2)
Actual Value (q_{ult}) (kN/m^2)
Testing
y = x
y = 1.2x and y = 0.8x
Predicted Value (q_{ult}) (kN/m^2)
Actual Value (q_{ult}) (kN/m^2)

(f)



(g)

Figure 4. Scatter plots comparing predicted q_{ult} (kN/m²) against FEA-computed q_{ult} (kN/m²) for training (left) and testing (right) subsets for: (a) Multiple Linear Regression, (b) Polynomial Regression, (c) Support Vector Regression, (d) Decision Tree, (e) Random Forest, (f) K-Nearest Neighbors, and (g) XGBoost. The solid red line represents the 1:1 perfect-prediction reference; dashed lines indicate the ±20% error bounds (y = 1.2x and y = 0.8x).

set with the lowest RMSE (12.64) and highest R^2 (0.974) and a_{20} index (97.08%), followed closely by SVR (RMSE = 14.17, R^2 = 0.967, a_{20} index = 95.82%) and Decision Tree (RMSE = 14.95, R^2 = 0.964, a_{20} = 93.32%). Random Forest and Polynomial Regression had similar performance with R^2 values of 0.962 and 0.960, respectively. Among the non-linear models, KNN performed quite weak (R^2 = 0.928, MAPE = 11.44%). Compared to KNN, Linear Regression (MLR) had the lowest performance overall with the highest RMSE (31.80) and MAPE (17.83%) and the lowest R^2 (0.836) and a_{20} index (68.68%) indicating its limited ability of capturing the non-linear relationship between slope geometry, footing dimensions and the bearing capacity.

The general ranking pattern is mostly preserved on the testing set, but as expected, accuracy of all models degrades. XGBoost achieved the highest prediction performance again, obtaining the lowest RMSE (16.82) and highest R^2 (0.953), followed closely by SVR (RMSE = 17.00, R^2 = 0.952) which slightly outperformed XGBoost on these two metrics (MAPE = 8.33%, a_{20} index = 92.5%) by obtaining the lowest MAPE (7.84%) and highest a_{20} index (95.0%) among all seven models. This means that while XGBoost provides the most accurate predictions in terms of absolute error magnitude, SVR offers slightly better consistency in relative prediction accuracy on unseen data. Polynomial Regression, Random Forest and Decision Tree performed moderately on testing, with R^2 values of 0.945, 0.925 and 0.914 respectively, while KNN once again showed relatively weaker generalization (R^2 = 0.862, MAPE = 16.76%). Linear Regression remained the worst performing model on the testing set (R^2 = 0.802, RMSE = 34.35), consistent with its performance during the training phase.

The outcomes have real-world implications for geotechnical foundation design on sloping terrain. The better accuracy and generalization stability of XGBoost and SVR indicate that either of these models can be confidently used as a surrogate for quick estimation of the bearing capacity during the preliminary design stage, without the need for time-consuming FEA simulations when a representative dataset has been used for training. For example, a practitioner can give site specific values of γ, c, φ, B, D_f and the slope angle β as inputs to the trained XGBoost model and get a reliable estimate of q_{ult} within milliseconds relative to the hours it takes to set up and run a full FEA model.

The comparison between training and testing metrics further exemplifies the models' different generalization behaviors. The KNN model experienced the greatest R^2 reduction between the training and testing phases (from 0.928 to 0.862, a 0.066 decrease), accompanied by an increase in the RMSE and the MAPE, indicative of a higher sensitivity to unseen data points, which can be related to the model's dependence on the local neighborhood structure. The Decision Tree model also presented a relatively large performance gap (drop in R^2 of 0.050; the RMSE increased from 14.95 to 22.66, about 52%) consistent with the well-known tendency of single decision trees to overfit without ensemble averaging. The XGBoost and the SVR models exhibited the least relative degradation in R^2 (0.021 and 0.015, respectively) despite the highest training accuracy, both models successfully generalized to unseen data without excessive overfitting. Random Forest, as an ensemble method, had an intermediate gap (0.037 R-squared reduction), better than the single Decision Tree but not matching the generalization stability of XGBoost or SVR. In conclusion, the results proved the best accuracy and generalization of the XGBoost model among the seven models used for the prediction of bearing capacity. The SVR model is a good and robust alternative.

Although XGBoost showed the best overall performance in terms of the absolute error magnitude, the choice between XGBoost and SVR in practice should be driven by the particular application context. XGBoost is recommended when the maximum predictive accuracy is the primary objective and the training dataset is sufficiently large and well-distributed, as its gradient boosting architecture allows it to capture complex nonlinear parameter interactions. SVR, by contrast, is preferable when dataset size is limited or when consistency of predictions on unseen data is critical its superior $_{a20}$ index on the test set (95.0% vs. 92.5% for XGBoost) and lower MAPE (7.84% vs. 8.33%) indicate that it produces fewer large relative errors, which is an important consideration in safety-critical geotechnical applications. SVR is also more straightforward to deploy in lightweight design tools or real-time assessment platforms.

3.2 Shapley Additive exPlanation

To analyze the contribution of each input parameter to the predicted carrying capacity and to analyze the physical validity of the trained models, a SHAP (SHapley Additive exPlanations) analysis (Figure 5) was performed for the following seven models: XGBoost, K-Nearest Neighbors (KNN), Random Forest (RF), Decision Tree (DT), Support Vector Regression (SVR), Polynomial Regression, and Multiple Linear Regression (MLR). The SHAP summary chart illustrates how SHAP values are distributed across each individual feature. The horizontal position reflects the amount and direction of the feature's effect on the model's output, while the color denotes the feature's value (red = high, blue = low).

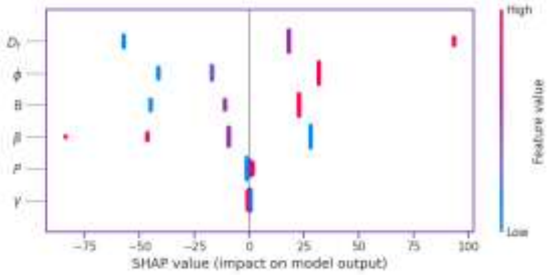
In all seven models, embedment depth (D_f) is consistently at the top of the SHAP plot as the most important parameter. This means that a greater embedment depth always increases the predicted bearing capacity and shallow embedding always decreases it for all model architectures tested, consistent with prior studies that also identified foundation depth as the dominant parameter using SHAP analysis on ensemble and boosting models such as XGBoost, GBM, and Random Forest (El Gendy, 2025; Shah et al., 2025).

Friction angle (φ) and footing width (B) are ranked second and third in importance for most models, with KNN being a noticeable exception where B is slightly more important than φ. Both parameters have a consistent positive relationship with model performance: high friction angles and larger footing widths correspond to positive SHAP values, whereas low values correspond to worse predictions. This ordering broadly agrees with Shah et al. (2025), who reported a similar hierarchy across XGBoost, Random Forest, and kNN models, reinforcing the physical consistency of φ and B as the next most influential parameters after depth. The directional relationship is similar across all seven models, but the range of SHAP values seems wider for tree-based and ensemble models (XGBoost, RF, DT) compared to MLR.

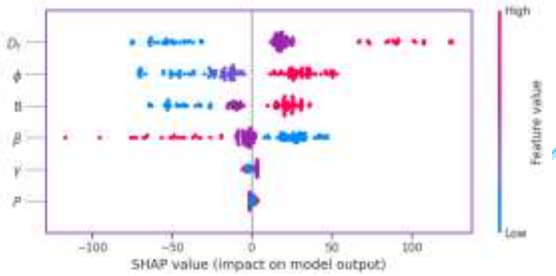
The slope angle (β) displays a distinctly opposite pattern compared to D_f, φ and B across all models. This inverse relationship is the most consistent directional pattern across all the seven models with the same sign and similar magnitude irrespective of the type of algorithm. This finding is in line with Mase et al. (2025), who reported that increasing slope angle (β) consistently reduced the predicted bearing capacity output across all six machine learning models they evaluated (Lasso, SVR, kNN, Decision Tree, Random Forest, Polynomial, FFNN, and XGBoost), attributing this to the greater horizontal thrust force generated on steeper slopes.

In contrast, the unit weight (γ) and additional load (P) consistently show the narrowest SHAP distribution among the six parameters, with values tightly clustered around zero across all the seven models. This suggests that these two parameters have a relatively minor influence on the model predictions as compared to D_f, φ, and β, for the range of values examined in this study. This observation is consistent with the findings of Mase et al. (2025) who also concluded that the most significant parameters impacting bearing capacity were D_f, B and φ, whereas P and γ had relatively little effect for the models examined.

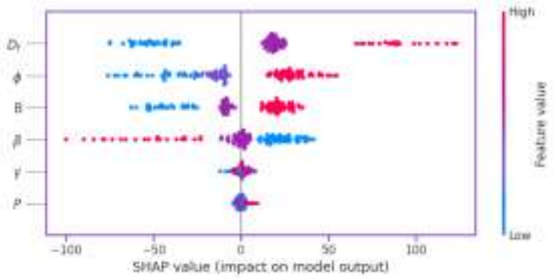
A clear pattern emerges in the comparison of the linear model (MLR) and the six nonlinear models. The MLR SHAP plot shows narrow, distinct ribbon-shaped bands for each feature, with the SHAP values for a particular feature lying along near straight line segments. This is because in the linear model the coefficients are fixed and constant and cannot have any impact on the sensitivity of a parameter to the values of other parameters. Among the nonlinear models explored (XGBoost, RF, DT, KNN, SVR, and Polynomial), the SHAP distributions of significant features including D_f, φ and β are consistently wide and continuous, indicating that these models are effective in recognizing complex nonlinear relationships and interaction effects among input variables. Interestingly, the SVR model's SHAP plot has a significantly different visual character from the other six models: rather than a dense, continuous cluster of points, the values are organized as sparse, discrete vertical clusters for each feature. This indicates that the SVR model's decision function produces a more limited number of distinct output levels per feature compared to tree-based and ensemble models, although the overall direction and feature importance rankings remain consistent with the other models.



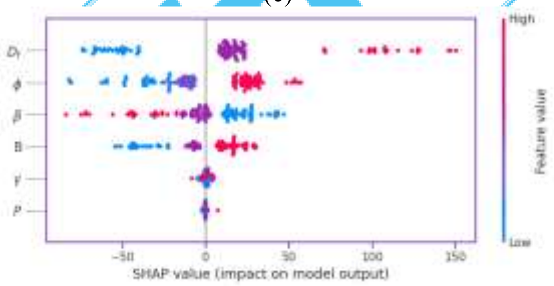
(a)



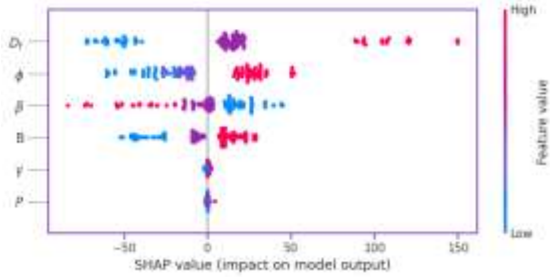
(b)



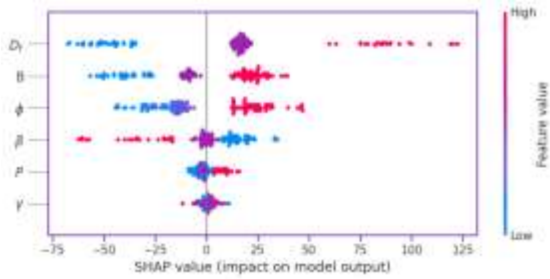
(c)



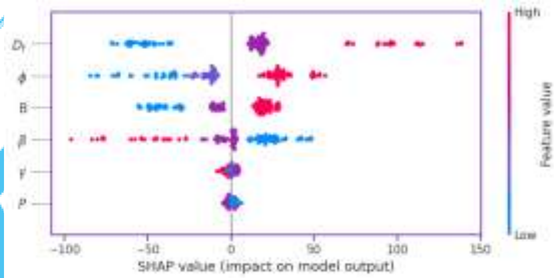
(d)



(e)



(f)



(g)

Figure 5. SHAP summary plots of feature contributions to predicted q*ult* for: (a) Multiple Linear Regression, (b) Polynomial Regression, (c) Support Vector Regression, (d) Decision Tree, (e) Random Forest, (f) K-Nearest Neighbors, and (g) XGBoost. Points represent individual samples; horizontal position indicates SHAP value (positive = increases q*ult*; negative = decreases q*ult*); color shows feature value (red = high, blue = low). Features are sorted by relevance top to bottom.

IV. CONCLUSION

The present study introduces a combined Finite Element Analysis (FEA) and Machine Learning (ML) approach for predicting the ultimate bearing capacity (q_{ult}) of strip footings resting on stiff clay slope conditions based on an existing database of 600 numerical simulations. The XGBoost model was the highest performing and most stable of the seven ML models generated with high accuracy for the training and testing datasets. Support vector regression came in second place, displaying the least drop in accuracy between the two subgroups and so offering a robust and efficient computational alternative for real applications. In contrast, multiple linear regression consistently underperformed, confirming its limited ability to capture the nonlinear interactions among slope geometry, footing dimensions, and bearing capacity. Moreover, the SHAP analysis provided insight into the physical basis of these predictions, as embedment depth was the most dominant parameter in all model architectures, followed by friction angle and footing width, whereas slope angle always reduced the predicted bearing capacity and unit weight and applied load only marginally contributed. These findings answer the research questions of how foundation, soil and slope parameters influence the bearing capacity of strip footings on slopes and show that the proposed FEA-ML hybrid approach is able to reliably reproduce this behaviour, while providing an easily interpretable and time-efficient alternative to iterative finite element analysis. This framework has to be extended in future research by including groundwater conditions, layered or anisotropic soil profiles, three-dimensional foundation-slope geometries and other advanced algorithms such as deep learning.

AUTHOR'S CONTRIBUTION

Conceptualization: ADK, LZM;
Data curation: ADK, LZM;
Formal analysis: ADK, LZM;
Methodology: LZM, ADK;
Validation: ADK, LZM;
Visualization: ADK, LZM;
Writing: ADK, LZM, MNF, RM, AF.

REFERENCES

Acharyya, R. (2019). Finite element investigation and ANN-based prediction of the bearing capacity of strip footings resting on sloping ground. *International Journal of Geo-Engineering*, *10*(1). https://doi.org/10.1186/s40703-019-0100-z

Bowles, J. E. (1997). *Foundation analysis and design* (5th ed.). McGraw-Hill.

Das, B. M., & Sivakugan, N. (2018). *Principles of strip footing engineering*. Cengage Learning.

El-Emam, M., El Berizi, Y., Mabrouk, A. B., & Tabsh, S. W. (2023). Bearing capacity of strip footing on top of slope: Numerical parametric study. *Ain Shams Engineering Journal*, *14*(11), 102522. https://doi.org/10.1016/j.asej.2023.102522

El Gendy, M. (2025). A comparative machine and deep learning approach for predicting ultimate bearing capacity of shallow foundations in cohesionless soil. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-22812-7

Khatti, J., & Grover, K. S. (2025). Estimation of uniaxial strength of rock: A comparison between Bayesian-optimized machine learning models. *Mining, Metallurgy & Exploration*, *42*(1), 133–154. https://doi.org/10.1007/s42461-024-01168-y

Kumar, D. R., Samui, P., Wipulanusat, W., Keawsawasvong, S., Sangjinda, K., & Jitchaijaroen, W. (2023). Soft-computing techniques for predicting seismic bearing capacity of strip footings in slopes. *Buildings*, *13*(6), 1371. https://doi.org/10.3390/buildings13061371

Le, T. N., Tran, D. T., Bui, H. N., Duong, N. T., Pham, G. H., & Lai, V. Q. (2025). Hybrid XGBoost–FELA Approach for Bearing Capacity of Rectangular Foundations on Slopes. *Geotechnical and Geological Engineering*, *43*(5). https://doi.org/10.1007/s10706-025-03155-6

Liu, Y., & Liang, Y. (2024). Integrated machine learning for modeling bearing capacity of shallow foundations. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-58534-5

Mase, L. Z., Misliniyati, R., Muharama, N. A., Supriani, F., Ahmad, D. A., Fernanda, R., Chauhan, V. B., & Chaiyaput, S. (2025). Integrating Multiple Linear Regression Analysis and Machine Learning Models to Predict the Bearing Capacity of Strip Footings on Sandy Clay Slopes. In *Transportation Infrastructure Geotechnology* (Vol. 12, Number 2). https://doi.org/10.1007/s40515-025-00544-5

Mase, L. Z., Putri, M. A., Edriani, A. F., Lai, V. Q., & Keawsawasvong, S. (2024). Prediction of the Bearing Capacity of Strip Footing at the Homogenous Sandy Slope Based on the Finite Element Method and Multivariate Adaptive Regression Spline. *Transportation Infrastructure Geotechnology*, *11*(4), 1520–1546. https://doi.org/10.1007/s40515-023-00334-x

Mase, L. Z., Sunanda, R., Utami, F., Wahyuni, M. S., Syafrizal, M. D., Putri, M. A., & Keawsawasvong, S. (2023). Slope Stability of Riverbank during Flood in

Bengkulu City, Indonesia (A case study of Nakau Segment). *Science and Technology Asia*, *28*(2), 125–142. https://doi.org/10.14456/scitechasia.2023.32

Moayedi, H., & Hayati, S. (2018). Modelling and optimization of ultimate bearing capacity of strip footing near a slope by soft computing methods. *Applied Soft Computing Journal*, *66*, 208–219. https://doi.org/10.1016/j.asoc.2018.02.027

Otake, Y., & Honjo, Y. (2022). Challenges in geotechnical design revealed by reliability assessment: Review and future perspectives. *Soils and Foundations*, *62*(3), 101129. https://doi.org/10.1016/j.sandf.2022.101129

Pramudika, Y., & Hastuty, I. P. (2015). *Comparison of the bearing capacity of the continuous foundation on layered soils with analytical methods and finite element methods* (Research Report). University of South Sumatra.

Shah, J., Alam, M., Javed, F., Umar, M., & Ahmad, F. (2025). Comparative performance evaluation of machine learning models for predicting the ultimate bearing capacity of shallow foundations on granular soils. *Scientific Reports*, *15*(1), 1–21. https://doi.org/10.1038/s41598-025-13926-z

Yaghoubi, E., Yaghoubi, E., Khamees, A., & Vakili, A. H. (2024). A systematic review and meta-analysis of artificial neural network, machine learning, deep learning, and ensemble learning approaches in field of geotechnical engineering. In *Neural Computing and Applications* (Vol. 36, Number 21). Springer London. https://doi.org/10.1007/s00521-024-09893-7