International Journal Software Engineering and Computer Science (IJSECS)
6 (2), 2026, 739-751
Published Online August 2026 in IJSECS (http://www.journal.lembagakita.com/index.php/ijsecs)
P-ISSN: 2776-4869, E-ISSN: 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v6i2.7430.

RESEARCH ARTICLE **Open Access**

Machine Learning-Based Prediction of Shallow Foundation Bearing Capacity Incorporating Slope Inclination as a Predictive Feature

Afifah Yurisa Putri [1], **Lindung Zalbuin Mase** [2*], **Muharram Nur Fikri** [3], **Rena Misliniyati** [4], **Aidil Fitriansyah** [5]

[1] Departement of Civil Engineering, Faculty of Engineering, Universitas Bengkulu, Bengkulu City, Bengkulu Province, Indonesia.
[2*,3,4,5] Geotechnical Hazard Research Unit, Faculty of Engineering, Universitas Bengkulu, Bengkulu City, Bengkulu Province, Indonesia.

**Corresponding author*: lmase@unib.ac.id.

Received: June 5, 2026; Accepted: July 3, 2026; Published: August 1, 2026.

Abstract: Accurate prediction of the ultimate bearing capacity of shallow foundations is important for safe geotechnical design, particularly under varying slope and soil conditions where conventional analytical methods may be limited by simplifying assumptions. This study developed and evaluated machine learning models for predicting the ultimate bearing capacity of shallow foundations using slope inclination, footing width, foundation depth, and soil friction angle as input variables. A dataset of 399 samples obtained from a previously validated finite element method (FEM)-based investigation was used to train and test Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost) models. Model performance was evaluated using the coefficient of determination (R^2), Root Mean Square Error (RMSE), and 10-fold cross-validation. Unlike previous studies that primarily considered conventional foundation and soil parameters, this study incorporates slope inclination within a unified machine learning framework and compares statistical, kernel-based, and ensemble learning approaches under the same evaluation conditions. XGBoost achieved the highest predictive performance, with a testing R^2 of 0.9938 and an RMSE of 31.669, followed by SVR with an R^2 of 0.9534 and an RMSE of 86.900. MLR showed comparatively lower performance, with an R^2 of 0.8396 and an RMSE of 161.158. The 10-fold cross-validation results further indicated stable XGBoost performance, with a mean R^2 of 0.991 and a standard deviation of 0.003. These results indicate that XGBoost provides high predictive performance for the evaluated dataset and may support bearing capacity estimation for shallow foundation design.

Keywords: Ultimate Bearing Capacity; Shallow Foundation; XGBoost; Support Vector Regression; Machine Learning.

1. Introduction

The design of shallow foundations requires accurate evaluation of ultimate bearing capacity because it directly affects structural safety, serviceability, and construction cost. An underestimated bearing capacity may result in excessive settlement or shear failure, whereas an overestimated value may lead to unsafe foundation design. Therefore, accurate prediction of the ultimate bearing capacity of shallow foundations remains an important issue in geotechnical engineering. Terzaghi (1943) and Meyerhof (1963) developed analytical approaches that remain widely used in shallow foundation design, providing practical methods for estimating bearing capacity based on foundation geometry and soil strength parameters. However, their underlying assumptions regarding failure mechanisms, soil homogeneity, and stress distribution may limit their predictive capability under complex geotechnical conditions (Pantelidis, 2024). Consequently, differences may occur between analytical predictions and actual foundation behavior, particularly when soil and ground surface conditions differ from the assumptions used in conventional formulations.

Alongside these limitations, the development of artificial intelligence and machine learning has provided alternative approaches for addressing nonlinear problems in geotechnical engineering. Machine learning algorithms can identify nonlinear relationships and interactions among multiple variables without requiring an explicit mathematical formulation. Recent studies have applied machine learning to various geotechnical problems, including soil characterization, slope stability assessment, liquefaction prediction, and foundation engineering (Yaghoubi *et al.*, 2024), and recent reviews have also reported the growing application of artificial intelligence for prediction, decision support, and uncertainty assessment in geotechnical engineering (Sheil *et al.*, 2026). However, the performance of machine learning models depends not only on predictive accuracy but also on appropriate evaluation and transparent validation procedures (Bozorgzadeh & Feng, 2024), a consideration particularly relevant to bearing capacity prediction because the response is affected by interactions among soil properties, foundation dimensions, and ground surface conditions.

The ultimate bearing capacity of shallow foundations is affected by several interrelated factors, including soil strength, foundation width, embedment depth, and ground surface conditions. The interaction among these parameters may produce nonlinear responses that are difficult to represent consistently using conventional analytical equations (Liu & Liang, 2024; Yaghoubi *et al.*, 2024). Therefore, data-driven models have increasingly been investigated as an alternative approach for estimating bearing capacity under different geotechnical conditions.

Several studies have applied machine learning techniques to the prediction of ultimate bearing capacity. Ahmad *et al.* (2021) employed Gaussian Process Regression and reported that the data-driven model provided improved predictions compared with conventional analytical approaches for cohesionless soils. Liu and Liang (2024) developed a machine learning framework for bearing capacity prediction and reported improved predictive performance compared with traditional approaches. Shah *et al.* (2025) also evaluated several machine learning algorithms for predicting ultimate bearing capacity under different soil and foundation conditions. These studies demonstrate the potential of machine learning for representing the nonlinear relationships involved in bearing capacity prediction.

Ensemble learning methods have also received attention because of their ability to model complex relationships among input variables. El Gendy (2025) compared several machine learning and deep learning algorithms for shallow foundation bearing capacity prediction and reported strong performance from ensemble-based methods, including XGBoost. Zeini *et al.* (2025) similarly demonstrated the applicability of ensemble machine learning models for predicting the bearing capacity of shallow foundations on granular soils. Li *et al.* (2025) reported that advanced machine learning models can capture complex interactions among geotechnical parameters while maintaining high predictive performance. In addition, Mase *et al.* (2025) investigated bearing capacity under sloping ground conditions using machine learning techniques, indicating that slope conditions should be considered when predicting foundation performance. Despite these developments, several research gaps remain. Previous studies have investigated bearing capacity prediction using different combinations of foundation and soil parameters, while studies considering sloping ground have also demonstrated the relevance of slope conditions to foundation performance. However, a systematic comparison of different predictive approaches using the same dataset and evaluation conditions remains limited. In particular, the relative predictive contribution of slope inclination together with footing width, foundation depth, and soil friction angle requires further assessment. A comparative evaluation of statistical regression, kernel-based learning, and ensemble learning models may provide a clearer basis for identifying the most suitable predictive approach for the evaluated foundation conditions.

To address these gaps, this study develops and compares three predictive models, namely Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost), for estimating the ultimate bearing capacity of shallow foundations. The models use the sine of slope inclination (sin β), footing width (B), foundation depth (Df), and the tangent of soil friction angle (tan φ) as input variables, while ultimate bearing capacity (Qult) is used as the target variable. The contribution of this study lies in the systematic comparison of statistical regression, kernel-based machine learning, and ensemble learning approaches using a validated geotechnical dataset under the same evaluation framework. The study also assesses the relative importance of the input parameters using feature importance analysis, providing additional information on the predictive role of slope inclination and other input parameters in estimating shallow foundation bearing capacity. The objectives of this study are: (1) to develop MLR, SVR, and XGBoost models for predicting the ultimate bearing capacity of shallow foundations; (2) to evaluate and compare model performance using the coefficient of determination (R^2), root mean square error (RMSE), and 10-fold cross-validation; (3) to assess the relative importance of the geotechnical and geometric input parameters using feature importance analysis; and (4) to identify the model with the best predictive performance for the evaluated dataset.

2. Related Work

2.1 Previous Studies on Bearing Capacity Prediction

The prediction of ultimate bearing capacity remains an important topic in geotechnical engineering because of its direct relationship with foundation safety and design. Traditional bearing capacity theories developed by Terzaghi (1943) and Meyerhof (1963) have been widely used in engineering practice. However, their applicability may be affected by simplifying assumptions related to soil conditions, failure mechanisms, and loading conditions (Pantelidis, 2024). These limitations have encouraged researchers to investigate data-driven approaches for predicting bearing capacity under different geotechnical conditions. Recent developments in artificial intelligence and machine learning have provided alternative approaches for modeling nonlinear geotechnical problems. Yaghoubi *et al.* (2024) reviewed the application of machine learning, deep learning, and ensemble learning methods in geotechnical engineering and reported their increasing use in prediction and analysis. Ahmad *et al.* (2021) demonstrated that Gaussian Process Regression could be used to predict the ultimate bearing capacity of shallow foundations and reported improved predictive performance compared with conventional approaches. Similarly, Liu and Liang (2024) showed that machine learning models can represent complex relationships among soil and foundation parameters in bearing capacity prediction.

More recent studies have investigated different machine learning algorithms and their performance in bearing capacity prediction. Shah *et al.* (2025) evaluated several machine learning approaches and reported strong predictive performance from ensemble-based models. El Gendy (2025) and Zeini *et al.* (2025) also investigated boosting-based approaches, including XGBoost, and reported strong performance for shallow foundation bearing capacity prediction. Li *et al.* (2025) further examined the use of engineering knowledge in machine learning-based prediction, indicating the potential value of incorporating domain-specific information into data-driven models. These studies indicate that machine learning can provide useful predictive capabilities for bearing capacity estimation, although model performance may vary depending on the dataset, input parameters, and evaluation procedure. Recent research has also addressed the physical consistency and robustness of machine learning models. Kiany *et al.* (2023) investigated hybrid and grey-box artificial intelligence approaches for geotechnical prediction, while Pei and Luo (2025) incorporated physical constraints into neural network architectures. Niu *et al.* (2025) developed a neural operator-based deep learning approach for representing nonlinear relationships between soil properties and foundation geometry. These studies reflect a broader development toward combining data-driven methods with engineering knowledge. However, the approaches used in these studies differ from the conventional machine learning models evaluated in the present study. Mase *et al.* (2025) investigated bearing capacity under sloping ground conditions using machine learning techniques, demonstrating the relevance of slope conditions to foundation performance. This finding is particularly relevant because slope inclination introduces an additional geometric condition that may affect the relationship between foundation parameters and ultimate bearing capacity. Taken together, previous studies demonstrate the potential of machine learning for bearing capacity prediction while also indicating the need to evaluate different algorithms under consistent data and evaluation conditions.

2.2 Comparison of Machine Learning Approaches for Bearing Capacity Prediction

Different machine learning approaches have distinct characteristics that affect their suitability for bearing capacity prediction. Statistical regression models provide relatively simple and interpretable relationships between input variables and the target response. However, linear regression models may have difficulty representing complex nonlinear interactions among geotechnical parameters. This limitation is relevant to bearing capacity prediction because soil and foundation parameters may interact in nonlinear ways. Kernel-based approaches such as Support Vector Regression (SVR) provide an alternative for modeling nonlinear relationships. Through the use of appropriate kernel functions, SVR can represent nonlinear patterns while maintaining good generalization performance. This characteristic has made SVR applicable to prediction problems involving limited datasets and complex engineering relationships. Its performance, however, depends on the selected kernel and associated hyperparameters.

Ensemble learning approaches provide another strategy for modeling nonlinear relationships by combining multiple decision-tree-based learners. XGBoost is one such approach and has been applied to several geotechnical prediction problems. Shah *et al*. (2025), El Gendy (2025), and Zeini *et al*. (2025) reported strong predictive performance from ensemble-based approaches, including XGBoost, in their respective studies. The ability of XGBoost to represent nonlinear interactions and estimate the relative importance of input variables makes it relevant to bearing capacity prediction. Recent research has also emphasized the integration of engineering knowledge into machine learning models. Li *et al*. (2025), Kiany *et al*. (2023), Pei and Luo (2025), and Niu *et al*. (2025) investigated different strategies involving engineering knowledge, physical constraints, or hybrid modeling. Although these approaches can provide additional physical interpretation, they also involve modeling strategies that differ from the standard statistical, kernel-based, and ensemble approaches considered in the present study.

Machine learning has also been applied to other geotechnical problems, including liquefaction assessment, pile capacity estimation, and geotechnical stability analysis (Kumar *et al*., 2023; Hakan *et al*., 2025; Seo *et al*., 2025; Nguyen-Son *et al*., 2026). These applications demonstrate the ability of data-driven models to represent nonlinear relationships in geotechnical datasets. However, model performance remains dependent on the characteristics of the dataset and the evaluation framework used in each study. Overall, MLR, SVR, and XGBoost represent three different predictive approaches with distinct characteristics. MLR provides a simple statistical baseline, SVR provides nonlinear kernel-based modeling capability, and XGBoost uses ensemble tree-based learning to represent complex interactions. Comparative evaluation of these three approaches using the same dataset and evaluation procedure can therefore provide a more consistent basis for assessing their relative predictive performance. In addition, examining the relative importance of slope inclination alongside footing width, foundation depth, and soil friction angle can provide further information about the predictive role of these variables in the evaluated foundation conditions.

2.3 Overview of Machine Learning Algorithms

Three predictive approaches were employed in this study: Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost). These methods were selected to represent statistical regression, kernel-based learning, and ensemble learning, respectively. Multiple Linear Regression (MLR) was selected as a baseline statistical model because of its simplicity, interpretability, and widespread application in engineering analysis (Liao *et al*., 2024). MLR assumes a linear relationship between the input variables and the target variable. In the present study, MLR provides a reference for determining whether more flexible nonlinear models can provide better predictive performance for ultimate bearing capacity. Support Vector Regression (SVR) is a supervised learning method that can model nonlinear relationships through kernel functions. The method seeks a regression function that minimizes prediction error while controlling model complexity. SVR has demonstrated good generalization capability in various prediction problems and can be useful for engineering datasets with relatively limited sample sizes (Liao *et al*., 2024). Its predictive performance is influenced by the selected kernel function and hyperparameters. Extreme Gradient Boosting (XGBoost) is an ensemble learning method based on gradient-boosted decision trees. The algorithm builds a sequence of decision trees in which subsequent trees are used to reduce the errors of previous learners. This approach enables XGBoost to represent nonlinear relationships and interactions among input variables. XGBoost has demonstrated strong predictive performance in several geotechnical applications, including shallow foundation bearing capacity prediction (El Gendy, 2025; Zeini *et al*., 2025). Its ability to estimate feature importance also allows the relative predictive contribution of the input variables to be examined.

3. Methodology

3.1 Research Framework

This study was conducted to develop and evaluate machine learning models for predicting the ultimate bearing capacity (Qult) of shallow foundations under varying slope and soil conditions. The research framework consisted of six main stages: dataset preparation, data preprocessing, model development, hyperparameter optimization, model evaluation, and feature importance analysis. Four input variables, namely the sine of slope inclination (sin β), footing width (B), foundation depth (Df), and the tangent of soil friction angle (tan φ), were used to predict ultimate bearing capacity (Qult). The dataset was obtained from a previously validated finite element method (FEM)-based investigation and subsequently prepared for machine learning analysis. The dataset was divided into training and testing subsets using an 85:15 ratio. Three predictive approaches, namely Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Extreme Gradient Boosting

(XGBoost), were developed and compared. Hyperparameter optimization was performed using GridSearchCV with 10-fold cross-validation on the training data. Model performance was subsequently evaluated using the coefficient of determination (R^2) and Root Mean Square Error (RMSE) on the independent testing data. Finally, feature importance analysis was conducted using the XGBoost model to assess the relative predictive importance of the input variables.

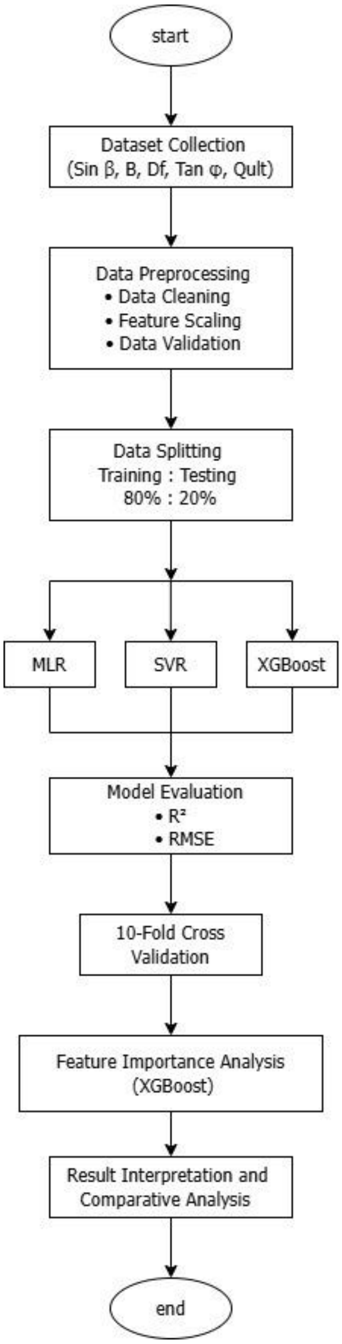

Figure 1. Research Framework of the Proposed Method

3.2 Dataset Description

The dataset used in this study consisted of 399 samples obtained from a previously validated finite element method (FEM)-based investigation of shallow foundation bearing capacity under sloping ground conditions. The database was generated by systematically varying four input parameters, namely the sine of slope inclination (sin β), footing width (B), foundation depth (Df), and the tangent of soil friction angle (tan φ), while the corresponding ultimate bearing capacity (Qult) values were recorded as the target output. The selected variables represent geometric and geotechnical parameters associated with shallow foundation bearing capacity and were considered based on previous studies (Ahmad *et al.*, 2021; Liu & Liang, 2024; Shah *et al.*, 2025; Mase *et al.*, 2025). The FEM-based dataset provides numerical relationships between the input variables and the corresponding bearing capacity values, making it suitable for the development and evaluation of machine learning models.

Table 1. Statistical Characteristics of the Dataset

Variable	Count	Mean	Std. Dev.	Minimum	Maximum
Sin β	399	0.403	0.172	0.174	0.643
B (m)	399	1.286	0.209	1.000	1.500
Df (m)	399	0.857	0.350	0.500	1.500
Tan φ	399	0.682	0.070	0.577	0.781
Qult (kPa)	399	402.730	258.640	44.550	1480.600

3.3 Dataset Preprocessing

Prior to model development, the dataset was examined to ensure data quality and consistency. The preprocessing stage included data verification and duplicate checking. No missing values were identified in the dataset. The dataset was randomly divided into training and testing subsets using an 85:15 ratio with a random state of 42. A total of 339 samples were used for training and model development, while 60 samples were reserved as an independent testing set. The testing data were not used during model training or hyperparameter optimization. Feature scaling was performed using the StandardScaler technique. The scaler was fitted using the training data and subsequently applied to both the training and testing data. Standardization was particularly relevant to Support Vector Regression (SVR), which is sensitive to differences in the magnitude of input variables.

3.4 Machine Learning Models

Three predictive models were developed in this study, namely Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost), representing statistical regression, kernel-based learning, and ensemble learning approaches, respectively. MLR was employed as a baseline model because of its simplicity and interpretability (Liao *et al.*, 2024; Seo *et al.*, 2025). The final regression equation obtained from the developed MLR model is presented as follows:

$$Qult = 638.9730 - 249.7300(Sin\beta) + 98.3389(B) + 164.9412(Df) + 210.3534(Tan\varphi)$$

Where Qult is the ultimate bearing capacity, sin β represents the sine of slope inclination, B denotes footing width, Df represents foundation depth, and tan φ denotes the tangent of the soil friction angle. SVR was adopted to model nonlinear relationships among the input variables through kernel-based learning, whereas XGBoost was employed to represent complex nonlinear interactions and provide feature importance information (El Gendy, 2025; Zeini *et al.*, 2025). Hyperparameter optimization was conducted using GridSearchCV with 10-fold cross-validation on the training dataset. For SVR, the search space included four kernel functions, namely linear, polynomial, radial basis function (RBF), and sigmoid, with C values ranging from 0.1 to 100 and epsilon (ε) values ranging from 0.01 to 1. The optimal configuration employed an RBF kernel with C = 100 and ε = 0.01. For XGBoost, the tuning process evaluated combinations of n_estimators, max_depth, learning_rate, subsample, and colsample_bytree. The optimal configuration consisted of 200 estimators, a maximum depth of 6, a learning rate of 0.1, a subsample ratio of 0.8, and a colsample_bytree ratio of 0.8.

3.5 Model Evaluation

The predictive performance of the developed models was evaluated using the coefficient of determination (R^2) and Root Mean Square Error (RMSE). R^2 measures the proportion of variance in the observed target variable that is explained by the model, whereas RMSE measures the magnitude of prediction errors between the observed and predicted values. To assess model performance during model selection and hyperparameter optimization, 10-fold cross-validation was applied to the training dataset. The mean validation R^2 was calculated across the ten folds. After the optimal hyperparameters were selected, the final models were trained using the complete training dataset and evaluated using the independent testing dataset. The coefficient of determination (R^2) is calculated as follows:

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

Where: y_i represents the actual value, $\hat{y}_i$ denotes the predicted value, $\bar{y}$ is the mean of observed values, and n is the number of samples. The Root Mean Square Error (RMSE) is calculated as:

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

Where: y_i is the observed value, $\hat{y}_i$ is the predicted value, and n denotes the total number of samples.

3.6 Feature Importance Analysis

Feature importance analysis was conducted using the built-in feature importance mechanism of the final XGBoost model. The analysis was used to assess the relative predictive importance of each input variable for ultimate bearing capacity prediction. Particular attention was given to slope inclination because one of the objectives of this study was to compare its predictive importance with footing width, foundation depth, and soil friction angle. The resulting importance scores were used to assess the relative contribution of the input variables to the model predictions and to improve the interpretation of the XGBoost model. The feature importance results were interpreted as model-based predictive importance rather than as evidence of a causal relationship between the input variables and ultimate bearing capacity.

4. Result and Discussion

4.1 Results

4.1.1 Predictive Performance of Machine Learning Models

The coefficient of determination (R^2) and Root Mean Square Error (RMSE) were used to evaluate the predictive performance of Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost). Table 2 presents the performance of the three models on the training and testing datasets.

Table 2. Performance Comparison of the Developed Models for Ultimate Bearing Capacity Prediction

Model	Phase	R^2	RMSE (kPa)
Multiple Linear Regression	Train	0.8819	140.619
Multiple Linear Regression	Test	0.8396	161.158
Support Vector Regression	Train	0.9416	98.838
Support Vector Regression	Test	0.9534	86.900
XGBoost	Train	0.9954	27.885
XGBoost	Test	0.9938	31.669

As shown in Table 2, XGBoost achieved the highest predictive performance among the evaluated models, with a testing R^2 of 0.9938 and the lowest testing RMSE of 31.669 kPa. An R^2 of 0.9938 indicates that the model explained approximately 99.38% of the variance in the testing data. SVR also produced strong predictive performance, with a testing R^2 of 0.9534 and an RMSE of 86.900 kPa. In contrast, MLR obtained the lowest testing performance, with an R^2 of 0.8396 and an RMSE of 161.158 kPa. The lower performance of MLR suggests that a linear relationship was less effective in representing the relationships among the investigated geotechnical variables. The superior performance of XGBoost indicates that the ensemble-based approach was better able to model the nonlinear relationships represented in the evaluated dataset. Its tree-based structure allows the model to capture interactions among slope inclination, footing width, foundation depth, and soil friction angle without requiring a predefined linear relationship between the predictors and Qult. The difference between the training and testing performance of XGBoost was relatively small. The training R^2 was 0.9954, compared with 0.9938 for testing, while the training and testing RMSE values were 27.885 and 31.669 kPa, respectively. This relatively small difference indicates that substantial overfitting was not evident in the evaluated dataset. The cross-validation results presented in the following section provide additional information regarding model stability across different data partitions.

4.1.2 Model Validation Using 10-Fold Cross-Validation

Ten-fold cross-validation was performed on the training dataset to assess the stability of the developed models during model evaluation and selection. Table 3 presents the mean R^2 and standard deviation obtained across the ten validation folds.

Table 3. 10-Fold Cross-Validation Results

Model	Mean CV R^2	Standard Deviation
Multiple Linear Regression	0.877	0.028
Support Vector Regression	0.952	0.029
XGBoost	0.991	0.003

The cross-validation results further demonstrate the strong performance of XGBoost. The model achieved the highest mean CV R^2 of 0.991 with a standard deviation of 0.003. The small standard deviation indicates that the XGBoost model produced relatively consistent R^2 values across the validation folds. SVR achieved a mean CV R^2 of 0.952 with a standard deviation of 0.029, whereas MLR obtained a mean CV R^2 of 0.877 with a standard deviation of 0.028. The smaller variation observed for XGBoost compared with MLR and SVR indicates greater stability across the evaluated training partitions. The similarity between the mean CV R^2 of 0.991 and the testing R^2 of 0.9938 also indicates that the model maintained a comparable level of predictive performance on the independent testing data. However, this result should be interpreted within the scope of the FEM-based dataset used in this study.

4.1.3 Comparison Between Actual and Predicted Values

Figures 2–4 illustrate the relationships between the actual and predicted ultimate bearing capacity values for the developed models. Points located closer to the ideal prediction line indicate smaller differences between the observed and predicted values.

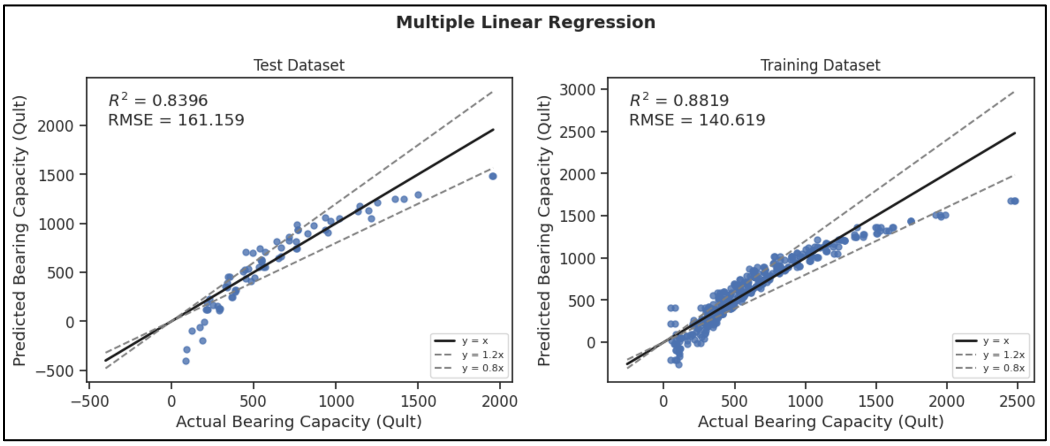

Figure 2. MLR Prediction Results

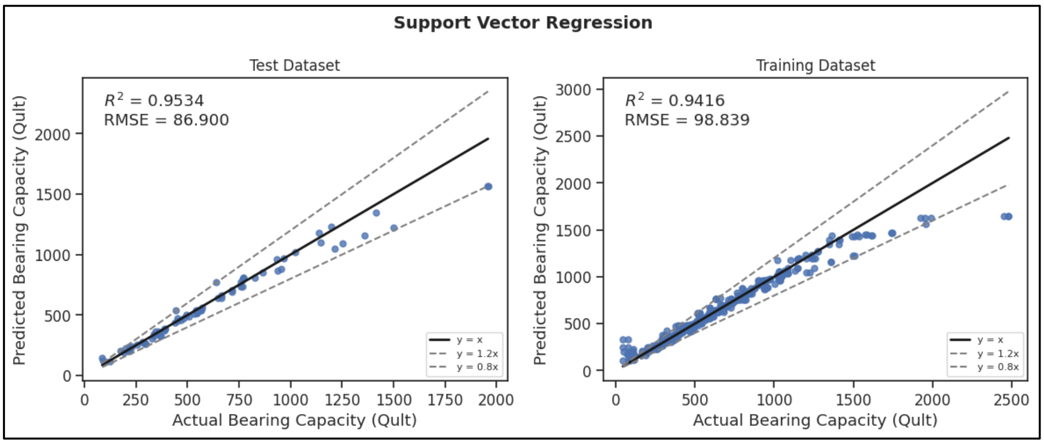

Figure 3. SVR Prediction Results

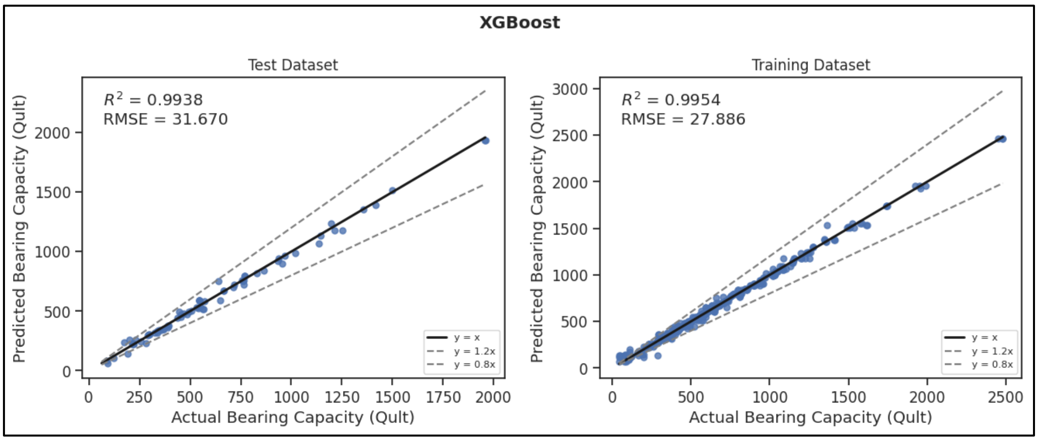

Figure 4. XGBoost Prediction Results

The scatter plots show that the XGBoost predictions were closely distributed around the ideal prediction line for both the training and testing datasets. The deviations between actual and predicted values were relatively small compared with those observed for MLR and SVR. SVR also showed a relatively close relationship between actual and predicted values, although several observations exhibited larger deviations from the ideal line. MLR showed the greatest dispersion, particularly for observations with higher bearing capacity values.
The visual observations in Figures 2–4 are consistent with the quantitative results presented in Table 2. XGBoost produced the highest R^2 values and the lowest RMSE values among the three models, while MLR produced the lowest R^2 and highest RMSE on the testing dataset.

4.1.4 Feature Importance Analysis

Feature importance analysis was performed using the XGBoost model to assess the relative predictive importance of the four input variables. Figure 5 presents the resulting importance scores.

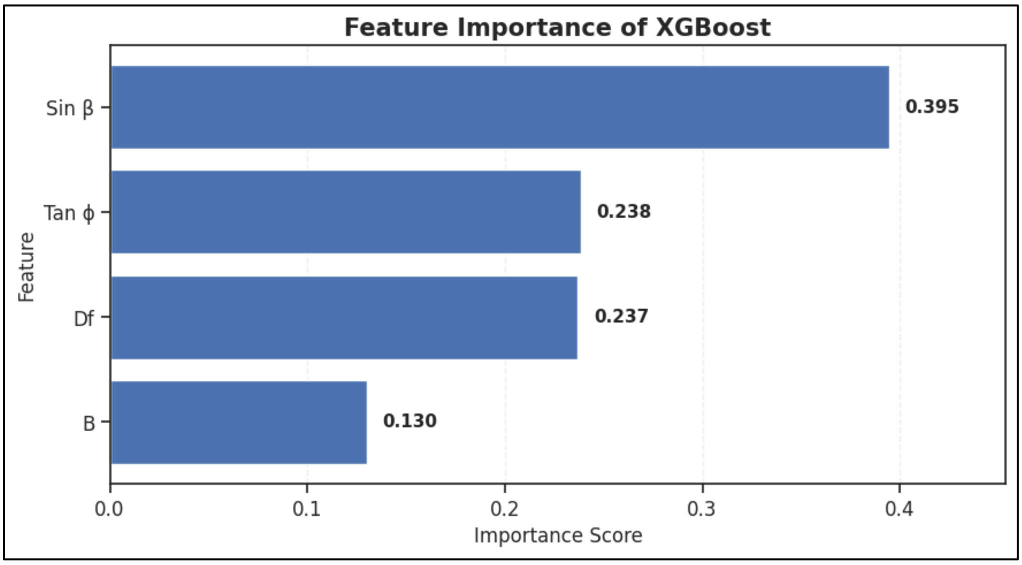

Figure 5. Feature Importance Scores Obtained from XGBoost

The results showed that the sine of slope inclination (sin β) had the highest feature importance score, with a value of 0.395. This result indicates that slope inclination had the greatest relative predictive importance among the four input variables in the developed XGBoost model. The finding supports the inclusion of slope-related parameters in the prediction of ultimate bearing capacity under the conditions represented in the dataset. The soil friction angle (tan φ) and foundation depth (Df) showed similar importance scores of 0.238 and 0.237, respectively. These results indicate that both variables contributed substantially to the model predictions. Footing width (B) had the lowest feature importance score, with a value of 0.130, indicating a smaller relative contribution to the XGBoost predictions compared with the other three variables. From a geotechnical perspective, the high importance of slope inclination is consistent with the role of ground surface geometry in shallow foundation behavior on sloping ground. Changes in slope geometry can alter the stress distribution and potential failure mechanism around the foundation. [tambahkan referensi] Therefore, the relatively high feature importance of sin β provides a basis for further examining slope conditions in bearing capacity prediction. The feature importance results should be interpreted as model-based predictive importance rather than direct evidence of causal influence. The scores indicate the relative contribution of each variable to the predictions generated by XGBoost and do not independently establish the physical causality of the observed relationships.

4.2 Discussion

The results demonstrate that the evaluated machine learning models were able to predict the ultimate bearing capacity represented in the FEM-based dataset, although their predictive performance differed considerably. XGBoost achieved the highest testing R^2 of 0.9938 and the lowest testing RMSE of 31.669 kPa. SVR produced a testing R^2 of 0.9534 and an RMSE of 86.900 kPa, whereas MLR obtained a testing R^2 of 0.8396 and an RMSE of 161.158 kPa. These results indicate that the nonlinear and ensemble-based approach of XGBoost was better suited to the relationships represented in the evaluated dataset than the linear regression and kernel-based approaches. The high performance of XGBoost may be related to the structure of the FEM-based dataset, in which the relationships between the input variables and ultimate bearing capacity were generated under systematically varied geotechnical conditions. XGBoost can model nonlinear relationships and interactions among the predictors through a sequence of decision trees. The small difference between its training and testing performance also indicates that substantial overfitting was not evident in the evaluated data. The training R^2 was 0.9954, while the testing R^2 was 0.9938.

The cross-validation results provide additional evidence of model stability. XGBoost achieved a mean CV R^2 of 0.991 with a standard deviation of 0.003, which was higher and less variable than the corresponding results for MLR and SVR. The mean CV R^2 was also close to the testing R^2, suggesting that the model maintained a similar predictive performance across the evaluated validation folds and independent testing data. Nevertheless, this consistency should not be interpreted as external validation because the dataset was derived from FEM simulations. Validation using experimental or field observations would be required to assess whether the model maintains similar performance under real foundation conditions. The testing RMSE of XGBoost was 31.669 kPa, while the Qult values in the dataset ranged from 44.55 to 1480.60 kPa. The RMSE is therefore relatively small compared with the range of the target values. This result indicates a low prediction error within the evaluated dataset. However, the RMSE alone does not establish the suitability of the model for final foundation design decisions, particularly when field uncertainties and variations in soil conditions are not represented in the dataset.

The superior performance of XGBoost is consistent with previous studies reporting strong performance from boosting-based machine learning models in geotechnical prediction applications (El Gendy, 2025; Zeini *et al.*, 2025). Similar findings have also been reported in studies applying advanced machine learning techniques to nonlinear geotechnical problems (Kiany *et al.*, 2023; Niu *et al.*, 2025). However, direct comparisons of R^2 values across different studies should be made cautiously because differences in datasets, input variables, target definitions, sample sizes, and validation procedures can affect model performance. The results of the present study therefore demonstrate the effectiveness of XGBoost specifically under the conditions represented by the evaluated FEM-based dataset. The feature importance analysis showed that sin β had the highest importance score at 0.395, followed by tan φ at 0.238, Df at 0.237, and B at 0.130. The results indicate that slope inclination had the highest relative predictive importance among the variables included in the XGBoost model. This finding is relevant to the research objective because slope inclination was explicitly incorporated together with foundation geometry and soil parameters.

The relatively high predictive importance of slope inclination is also consistent with the geotechnical role of slope geometry in shallow foundation behavior. Changes in slope inclination can modify the stress distribution and potential failure mechanism around the footing and may affect the soil volume available to resist the applied load. [tambahkan referensi] The feature importance result therefore supports further consideration of slope conditions in predictive models for shallow foundations on sloping ground.

The difference between the importance score of sin β (0.395) and that of footing width (0.130) indicates that slope inclination contributed more strongly to the XGBoost predictions within the evaluated dataset. However, this difference should not be interpreted as evidence that slope inclination is universally more important than footing width in all foundation conditions. The relative importance is specific to the range and distribution of variables represented in the dataset. The feature importance results also provide an additional interpretation of the proposed predictive framework. Unlike approaches that focus primarily on conventional foundation and soil parameters, this study explicitly included slope inclination as an input variable and assessed its relative predictive importance together with footing width, foundation depth, and soil friction angle. The result indicates that slope inclination was an important predictor within the evaluated dataset, providing quantitative support for its inclusion in machine learning-based bearing capacity prediction.

The feature importance scores were obtained from the built-in gain-based importance mechanism of XGBoost. Although this approach provides information regarding the relative predictive importance of the input variables, the scores may be affected by relationships among the predictors and should not be interpreted as direct measures of causality. Future studies may use explainable artificial intelligence methods, such as SHAP, to examine the contribution of individual variables in greater detail. Such analysis could also provide information about the direction and magnitude of individual feature contributions. Overall, the results show that XGBoost provided the strongest predictive performance among the three evaluated approaches, followed by SVR and MLR. The model achieved high testing performance and stable cross-validation results within the FEM-based dataset. The findings also indicate that slope inclination had the highest relative predictive importance among the four investigated variables. However, further validation using experimental or field data is needed before the model can be considered for broader engineering applications.

5. Conclusion and Recommendations

This study developed and compared three predictive models, namely Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost), for estimating the ultimate bearing capacity of shallow foundations. The models used four input variables, namely slope inclination (Sin β), footing width (B), foundation depth (Df), and soil friction angle (Tan φ). The results showed that all three models were able to predict ultimate bearing capacity within the investigated dataset, although their predictive performance differed. XGBoost achieved the best testing performance, with an R^2 value of 0.9938 and an RMSE of 31.669 kPa. The 10-fold cross-validation also produced a high mean R^2 of 0.991 with a standard deviation of 0.003, indicating stable predictive performance across the evaluated validation folds. Therefore, among the three approaches evaluated in this study, XGBoost provided the strongest predictive performance under the investigated conditions. The main contribution of this study is the systematic comparison of statistical regression, kernel-based learning, and ensemble learning approaches using the same dataset and evaluation framework. The study also explicitly incorporated slope inclination together with footing width, foundation depth, and soil friction angle. The XGBoost feature importance analysis showed that slope inclination had the highest relative predictive importance, with a score of 0.395, followed by soil friction angle (0.238), foundation depth (0.237), and footing width (0.130). This result indicates that slope inclination contributed most strongly to the predictions generated by the XGBoost model within the investigated dataset. However, the feature importance values represent model-based predictive importance and should not be interpreted as direct evidence of causal influence.

Several limitations should be acknowledged. First, the models were developed and evaluated using a single FEM-based dataset consisting of 399 samples and four predictor variables. Therefore, the reported predictive performance applies primarily to the ranges of slope inclination, footing width, foundation depth, soil friction angle, and foundation conditions represented in the dataset. Predictions outside these conditions should be interpreted with caution. Second, the study compared only three machine learning approaches, so other modeling techniques may produce different results. Third, the feature importance analysis was based on the gain-based importance metric of XGBoost. This metric provides relative predictive importance but may be affected by relationships among predictor variables and does not establish causal relationships. Future research should focus on developing and evaluating larger and more diverse datasets that cover wider ranges of soil properties, foundation geometries, and slope conditions. External validation using experimental or field data is also recommended to assess whether the predictive performance observed in the FEM-based dataset can be maintained under real geotechnical conditions. Other machine learning approaches, including LightGBM, CatBoost, Artificial Neural Networks, and Gaussian Process Regression, may also be evaluated to provide a broader comparison of predictive performance. In addition, explainable artificial intelligence methods such as SHAP, together with sensitivity and uncertainty analyses, could be applied to obtain a more detailed interpretation of model predictions. These approaches may provide additional information regarding the contribution and response of individual geotechnical variables. Overall, the XGBoost model demonstrated strong predictive performance within the investigated dataset and provides a basis for further development and external validation of machine learning-based approaches for shallow foundation bearing capacity prediction.

References

Ahmad, M., Ahmad, F., Wróblewski, P., Al-Mansob, R. A., Olczak, P., Kamiński, P., Safdar, M., & Rai, P. (2021). Prediction of ultimate bearing capacity of shallow foundations on cohesionless soils: A Gaussian process regression approach. *Applied Sciences, 11*(21). https://doi.org/10.3390/app112110317

Bozorgzadeh, N., & Feng, Y. (2024). Evaluation structures for machine learning models in geotechnical engineering. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 18*(1), 52–59. https://doi.org/10.1080/17499518.2024.2313485

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

Hakan, E., Talas, Ş., Kurnaz, F., & Erden, C. (2025). Estimation of soil liquefaction using artificial intelligence techniques: An extended comparison between machine and deep learning approaches. *Environmental Earth Sciences, 84*(5), 1–22. https://doi.org/10.1007/s12665-025-12116-4

Hamdaoui, K., Benzaamia, A., Sari Ahmed, B., *et al*. (2025). Interpretable machine learning for predicting compression index of clays using SHAP and gradient boosting models. *Journal of Engineering and Applied Sciences, 72*, 148. https://doi.org/10.1186/s44147-025-00727-4

Hsiao, C.-H., Kumar, K., & Rathje, E. M. (2024). Explainable AI models for predicting liquefaction-induced lateral spreading. *Frontiers in Built Environment, 10*, 1387953. https://doi.org/10.3389/fbuil.2024.1387953

Kiany, K., Baghbani, A., Abuel-Naga, H., Baghbani, H., Arabani, M., & Shalchian, M. M. (2023). Enhancing ultimate bearing capacity prediction of cohesionless soils beneath shallow foundations with grey box and hybrid AI models. *Algorithms, 16*(10), 456. https://doi.org/10.3390/a16100456

Kumar, D. R., Samui, P., Burman, A., Wipulanusat, W., & Keawsawasvong, S. (2023). Liquefaction susceptibility using machine learning based on SPT data. *Intelligent Systems with Applications, 20*, 200281. https://doi.org/10.1016/j.iswa.2023.200281

Li, H., Hosseini, S., Gordan, B., *et al.* (2025). Dimensionless machine learning: Dimensional analysis to improve LSSVM and ANN models and predict bearing capacity of circular foundations. *Artificial Intelligence Review, 58*, 117. https://doi.org/10.1007/s10462-024-11099-1

Liao, Z., Dai, S., & Kuosmanen, T. (2024). Convex support vector regression. *European Journal of Operational Research, 313*(3), 858–870. https://doi.org/10.1016/j.ejor.2023.05.009

Liu, Y., & Liang, Y. (2024). Integrated machine learning for modeling bearing capacity of shallow foundations. *Scientific Reports, 14*(1), 1–15. 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. *Transportation Infrastructure Geotechnology, 12*(2). https://doi.org/10.1007/s40515-025-00544-5

Nguyen-Son, L., Thai-Phuong, T., & Vo-Minh, T. (2026). Advanced machine learning models for predicting the seismic bearing capacity factors of shallow strip footings considering soil and superstructure inertia effects. *Transportation Infrastructure Geotechnology, 13*(4). https://doi.org/10.1007/s40515-026-00854-2

Niu, T., Huang, M., & Yu, J. (2025). Parametric deep learning model for predicting bearing capacity of strip foundation via neural operator. *AI in Civil Engineering, 4*, 11. https://doi.org/10.1007/s43503-025-00056-8

Pantelidis, L. (2024). Bearing capacity of shallow foundations: A focus on the depth factors in combination with the respective N-factors. *Arabian Journal of Geosciences, 17*(6). https://doi.org/10.1007/s12517-024-11976-7

Pei, T., & Luo, N. (2025). Prediction of ultimate bearing capacity of shallow foundations using physically constrained neural networks. In *Geotechnical Frontiers 2025* (Geotechnical Special Publication No. 365, pp. 11–20). American Society of Civil Engineers. https://doi.org/10.1061/9780784485989.002

Seo, S., Kim, G., Park, J. B., Kim, J., Park, Y. B., & Chung, M. (2025). Leveraging data-driven machine learning techniques to enhance bearing capacity estimation in prebored and precast piles. *Expert Systems with Applications, 285*, 128070. https://doi.org/10.1016/j.eswa.2025.128070

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

Sheil, B., Anagnostopoulos, C., Buckley, R., Ciantia, M. O., Febrianto, E., Fu, J., Gao, Z., Geng, X., Gong, B., Hanley, K., He, P., Kolomvatsos, K., Lopes, B. de C. F. L., Ninić, J., Previtali, M., Rezania, M., Ruiz-Lopez, A., Sun, J., Suryasentana, S., ... Zhang, P. (2026). Artificial intelligence transformations in geotechnics: Progress, challenges and future enablers. *Computers and Geotechnics, 189*, 107604. https://doi.org/10.1016/j.compgeo.2025.107604

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 the field of geotechnical engineering. *Neural Computing and Applications, 36*(21). https://doi.org/10.1007/s00521-024-09893-7

Zeini, H. A., Seno, M. E., Shehab, E. Q., Abood, E. A., Imran, H., Bernardo, L. F. A., & Ribeiro, T. P. (2025). Predicting the bearing capacity of shallow foundations on granular soil using ensemble machine learning models. *Geotechnics, 5*(3), 57. https://doi.org/10.3390/geotechnics5030057