

Prediction of Bearing Capacity of Clean Sand using Machine Learning Algorithms based on SPT Data

Alireza Amiri¹, Arash Niromand¹, Ba-Phu Nguyen², Nhat-Phi Doan^{2,*}, Duy Triet Doan³

¹Department of Civil Engineering, Central Tehran Branch, Islamic Azad University, Tehran 1955847781, Iran

²Department of Civil Engineering, Industrial University of Ho Chi Minh City, Ho Chi Minh City 084, Vietnam

³Department of Mechanic Dynamics, Vinh Long University of Technology Education, Vinh Long 85000, Vietnam

*Corresponding author: doannhatphi@iuh.edu.vn

Received: 10 March 2026 / Accepted: 28 April 2026 / Published: 25 May 2026

© The Author(s) 2026

Abstract: Accurate estimation of the bearing capacity of clean sand is a critical task in foundation design, especially in regions where granular soils dominate. Conventional empirical approaches may not fully capture the inherent variability and complex interactions among soil parameters. This study explores the use of machine learning (ML) techniques to predict the ultimate bearing capacity of clean sand based on Standard Penetration Test (SPT) data. A curated dataset of 150 geotechnical records specifically involving clean, well-graded sand (SP and SW according to the Unified Soil Classification System) was compiled. The data were collected from the NCEER database and validated case studies from peer-reviewed journals focusing on sandy soil profiles. Each data entry includes SPT N-values, relative density, dry unit weight, moisture content, and test depth. Four ML models, such as Linear Regression, Decision Tree, Random Forest, and Artificial Neural Network (ANN) were developed and tested using Python's Scikit-learn and TensorFlow libraries. Model performance was assessed using R^2 , RMSE, and MSE metrics. The Random Forest model yielded the most accurate predictions, with an R^2 value of 0.93 and a low RMSE, indicating a strong correlation between input features and bearing capacity in clean sand. This study demonstrates that ML-based prediction models can significantly enhance geotechnical design reliability in sandy soils and reduce the reliance on simplified empirical charts and conservative assumptions.

Keywords: Clean sand, SPT, Machine learning, Bearing capacity, Random forest.

I. INTRODUCTION

The design of shallow and deep foundations in geotechnical engineering heavily depends on the accurate determination of soil bearing capacity (Farhangi et al., 2020). Among various soil types, clean sandy soils are particularly important due to their widespread occurrence in coastal and alluvial regions, where infrastructure development is rapidly expanding (Poor et

al., 2023). Predicting the bearing capacity of clean sand poses a challenge because these soils exhibit non-cohesive behavior and their strength parameters are highly sensitive to relative density and confining stress (Zhang & Chen, 2012).

Standard Penetration Test (SPT) is one of the most widely used in-situ testing methods in geotechnical site investigations (Dung et al., 2011). It provides a simple and cost-effective means to evaluate the subsurface soil conditions (Lingwanda et al., 2015). The test involves driving a standard split-barrel sampler into the soil at the bottom of a borehole using a 63.5 kg hammer dropped from a height of 760 mm. The number of blows required to drive the sampler to 300 mm after an initial 150 mm seating is recorded as the N-value (Wei et al., 2020). Figure 1 illustrates the SPT testing process. This value serves as an indicator of soil resistance and stiffness (Akca, 2003). In clean sandy soils, SPT results are considered highly valuable due to the relatively good correlation between N-values and the soil's shear strength parameters including friction angle (ϕ) and relative density (D_r) (Khodayari et al., 2024). Unlike cohesive soils where undrained strength plays a major role, the behavior of sands under load is primarily governed by their interparticle friction and packing arrangement, both of which influence the SPT blow count (Mohamed & Vanapalli, 2015).

One of the main advantages of using SPT in sandy soils is its robustness and reliability in field conditions where other in-situ tests, like Cone Penetration Test (CPT), may be challenging due to gravelly inclusions or high groundwater tables (Golafzani et al., 2020). Furthermore, SPT data has been used for decades, resulting in a rich database of empirical correlations, especially for clean sands where cohesion is negligible and interpretations become more straightforward (Hung et al., 2016). Numerous empirical equations have been developed to estimate the bearing capacity of sandy soils using SPT N-values (Mohamed & Vanapalli, 2015). These empirical correlations account for parameters such as effective overburden pressure, relative density, and correction factors for energy efficiency (Wei et al., 2020). In Table 1, a summary of the most widely adopted empirical methods for estimating the bearing capacity of clean

sand based on SPT is presented. These methods form the basis for preliminary foundation design and are embedded in many national and international design codes. In addition to bearing capacity, SPT data in sandy soils can be used to estimate settlement behavior, liquefaction potential (Jinguuji & Toprak, 2017), modulus of elasticity, and dynamic properties (Farhangi et al., 2020). For clean sands, particularly those with minimal fines content, the interpretation of SPT data tends to be more consistent and reliable compared to silty or clayey soils (Ray et al., 2021).

The SPT is also particularly advantageous for sandy soils because it provides a direct indication of soil density, which is a key determinant of load-bearing performance (Cai et al., 2009). The test's simplicity allows for widespread application even in regions with limited access to advanced testing equipment (Anderson & Townsend, 2001). Furthermore, correction procedures for energy, rod length, and borehole diameter have been standardized, allowing for improved data consistency and quality (Leetsaar et al., 2023). Despite its advantages, traditional use of SPT results is often limited to empirical charts and deterministic calculations, which may not fully capture the nonlinear and multivariable relationships among soil parameters (Souza et al., 2012; Jarushi et al., 2015). This limitation creates an opportunity for more advanced approaches, such as machine learning, to model the underlying patterns in a data-driven manner.

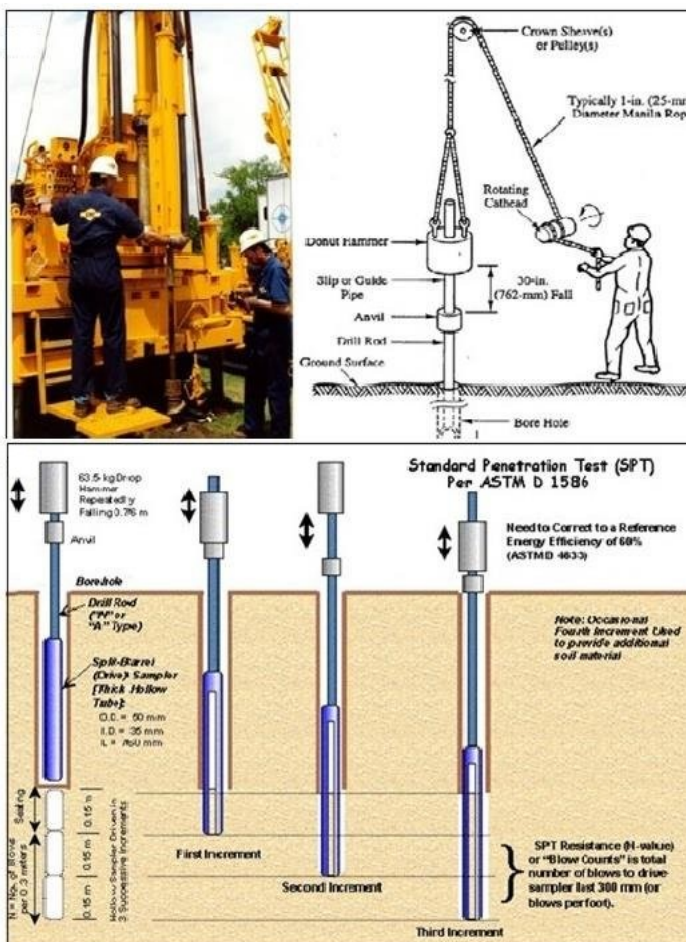


Fig. 1 A view of the SPT test procedure on site (Jarushi et al., 2016)

By considering SPT data through machine learning algorithms, it is possible to enhance the accuracy of bearing capacity predictions for clean sands. Machine learning techniques can integrate multiple variables, such as SPT N-value, dry density, moisture content, and depth, into a comprehensive model without assuming linear relationships, thereby offering improved predictive performance over classical methods. The integration of artificial intelligence into geotechnical engineering offers new possibilities for data interpretation, especially when dealing with large datasets or historical field records. In the context of clean sands, where consistent trends can be observed in SPT results, machine learning models can provide reliable and repeatable estimates of bearing capacity. This study focuses specifically on clean sands, defined as granular soils with little or no fines (typically less than 5%). These soils, classified as SP and SW under the Unified Soil Classification System (USCS) were presented in Figure 2, display good drainage and negligible plasticity, making them ideal candidates for SPT-based bearing capacity evaluation. The USCS is a widely adopted method for categorizing soils based on their grain size distribution and plasticity characteristics. It classifies soils into two major groups: coarse-grained soils (gravel and sand) and fine-grained soils (silts and clays). The distinction is made based on whether more or less than 50% of the soil by weight passes through the No. 200 sieve (0.075 mm). Coarse-grained soils are further divided into gravel (G) and sand (S) depending on whether more than 50% of the coarse fraction is larger or smaller than the No. 4 sieve (4.75 mm). Sands are then subclassified as well-graded (SW) or poorly graded (SP), depending on the uniformity and gradation of particle sizes. If significant fines are present, dual symbols such as SM (silty sand) or SC (clayey sand) are used. In this study, the focus is on clean sands, which are sands with little or no fines, typically classified under SP and SW categories in the USCS.

COARSE-GRAINED SOILS (more than 50% of material is larger than No. 200 sieve size.)		
GRAVELS More than 50% of coarse fraction larger than No. 4 sieve size	Clean Gravels (Less than 5% fines)	
	GW	Well-graded gravels, gravel-sand mixtures, little or no fines
	GP	Poorly-graded gravels, gravel-sand mixtures, little or no fines
	Gravels with fines (More than 12% fines)	
	GM	Silty gravels, gravel-sand-silt mixtures
	GC	Clayey gravels, gravel-sand-clay mixtures
SANDS 50% or more of coarse fraction smaller than No. 4 sieve size	Clean Sands (Less than 5% fines)	
	SW	Well-graded sands, gravelly sands, little or no fines
	SP	Poorly graded sands, gravelly sands, little or no fines
	Sands with fines (More than 12% fines)	
	SM	Silty sands, sand-silt mixtures
	SC	Clayey sands, sand-clay mixtures

Fig. 2 A granular soil classification based on USCS (Khandaker & Schleifer, 2015)

Table 1 Empirical methods for estimating ultimate bearing capacity of clean sand using SPT data

Method / Researcher	Equation for ultimate bearing capacity (q_{ult})	Required parameters	Applicability
Terzaghi & Peck (1948)	$q_{ult} = cN_c + qN_q + 0.5\gamma BN_\gamma$	q: overburden pressure B: footing width γ : unit weight	Classical bearing capacity formula, with SPT-based estimation for ϕ used in N_q , N_γ factors.
Meyerhof (1956)	$q_{ult} = (N \times \sigma') / F.S$	N: SPT blow count (corrected) σ' : effective overburden pressure F.S: safety/empirical factor	Commonly used for clean sands, assumes linear relationship between N and strength.
Teng (1962)	$q_{ult} = 400 \times N$	N: corrected SPT blow count	Simple empirical model; assumes q_{ult} in kN/m ² ; applicable mainly for medium to dense sands.
De Beer (1970)	$q_{ult} = f(N)$ using charts	N: SPT count	Based on Japanese sandy soil database; reliable for clean, dense sands.
Peck et al. (1974)	$q_{ult} = (N_{60} \times 95)$	N_{60} : energy-corrected SPT blow count	For shallow foundations on sand, consider corrected N-value.
Shioi & Fukui (1982)	$q_{ult} = \alpha \times N$	α : empirical coefficient (varies by soil type and region)	Regional calibration needed; simple for rough estimation.
Hatanaka & Uchida (1996)	$q_{ult} = 54 \times N_{60}$	N_{60} : energy-corrected SPT blow count	Based on extensive database of Japanese sandy soils; well-suited for clean, dense sands.
Das & Sabat (1997)	$q_{ult} = (0.44 \times N) \gamma B$	N: SPT blow count γ : unit weight B: footing width	Considers width effect; applicable for strip footings on sand.
Bowles (1997)	$q_{ult} = (N \times 100) + (\gamma \times D_f)$	N: SPT count γ : unit weight D_f : depth of foundation	Includes influence of depth; used in preliminary design.
Tan & Yuen (2003)	$q_{ult} = (k \times N_{60}^m) \times \sigma'_v$	k, m: empirical constants (calibrated) σ'_v : vertical effective stress	Nonlinear empirical model; k, m based on regression from Southeast Asian sands.
Al-Khafaji & Andersland (2006)	$q_{ult} = C_1 N_{60} + C_2 \gamma D_f$	C_1 , C_2 : constants γ : unit weight D_f : embedment depth	Includes depth influence; simplified linear approach.
Kumar & Sharma (2009)	$q_{ult} = \alpha N_{60} \times \sqrt{(D_f/B)}$	α : empirical coefficient D_f : depth of foundation B: footing width	Useful for shallow foundations; accounts for shape and embedment ratio.
Lee et al. (2010)	$q_{ult} = \beta N_{60} \times \log(D_f + 1)$	β : empirical constant D_f : depth of foundation	Calibrated for Korean sandy soils; depth-sensitive model.
Jamiolkowski (2012)	$q_{ult} = (N_{60} \times \sigma'_v)^{0.5} \times K$	N_{60} : corrected SPT value σ'_v : effective vertical stress K: calibration constant	Semi-empirical; calibrated using triaxial test database; reliable for dense sands.
Ghorbani et al. (2018)	$q_{ult} = a(N_{60})^b$	a, b: regression constants (region-specific)	Developed for granular soils in Iran; based on multiple regression analysis.
Eslami & Tavakoli (2019)	$q_{ult} = (N_{60} \times \log(D_f + 1)) / F.S$	D_f : foundation depth F.S: factor of safety	Modified Meyerhof-style equation; simplified for practical use.
Chen et al. (2021)	$q_{ult} = \alpha \times N_{60} \times \sqrt{(\gamma \times D_f/B)}$	α : coefficient D_f : depth γ : unit weight B: width	Recommended for machine learning-based design charts.
Salgado et al. (2021)	$q_{ult} = AN_{60} \times (\gamma \times B)^{0.3}$	A: constant γ : unit weight B: footing width	Incorporates scale effect and validated with centrifuge data.
Rahman & Zhang (2023)	$q_{ult} = f(N_{60}, \gamma, D_r, D_{50}, \sigma'_v, B, D_f)$ using ML	N_{60} , γ , D_r , D_{50} , etc.	Random Forest and Gradient Boosting models outperformed traditional formulas in clean sand prediction.
Zhang et al. (2024)	$q_{ult} = f(N_{60}, \gamma, D_f, B)$ using ANN	N_{60} , γ , D_f , B: input features	Artificial Neural Network trained on 300+ cases; high accuracy in clean sand profiles.
Zeini et al. (2025)	$q_{ult} = f(B, D_f, \gamma, \text{soil parameters})$ using ensemble ML	B: width D_f : depth γ : unit weight soil parameters	Ensemble ML (GBRT, XGBoost, LightGBM); high accuracy for granular soils
Arbi et al. (2025)	$q_{ult} = f(N_{60}, \gamma, D_r, \sigma'_v, B, D_f)$ using ML	N: SPT or input features B: width D_f : depth γ : unit weight	Optimized ML (RF, SVM, XGBoost) for bearing capacity prediction
Seo et al. (2025)	$q_{ult} = f(B, D_f, \gamma, \text{soil parameters})$ using ensemble ML	N: SPT value γ : unit weight soil parameters and geometry	Machine learning-based model (ANN/ensemble); applied to prebored and precast piles in granular soils

These soils exhibit non-cohesive behavior, high permeability, and strength that is primarily derived from internal friction and particle interlocking. Their granular structure results in primarily frictional strength, reinforcing the reliability of blow count-based correlations.

II. MATERIALS AND METHODS

This research follows a structured methodology that integrates traditional geotechnical testing data with advanced machine learning techniques to estimate the ultimate bearing capacity of clean sandy soils. The methodological approach is organized into several sequential phases, including data acquisition, data preprocessing and correction, feature analysis, model development, and model evaluation. Each phase is described in a detailed, systematic manner to ensure clarity and reproducibility. The analysis flowchart is provided in Figure 3.

The first step of the study involved collecting a reliable and representative dataset focused specifically on clean sand deposits. A total of 150 geotechnical site investigation reports were compiled, with special attention given to those where the soils were classified as SP (poorly graded sand) or SW (well-graded sand) according to the USCS. The main source of data was the National Center for Earthquake Engineering Research (NCEER) database, which provides high-quality, peer-reviewed site investigation records. Additional records were gathered from published journal articles and case studies that met the criteria of data completeness and standard testing procedures. Each selected data record contained results from the SPT, along with accompanying soil properties such as dry unit weight, moisture content, depth of testing, and, where available, relative density and footing dimensions.

Following data collection, the raw SPT N -values were corrected to account for field conditions and equipment variations. These corrections are essential to ensure that the data are standardized and can be compared or modeled consistently. The corrections included adjustments for hammer energy efficiency, borehole diameter, rod length, and sampler type. The corrected value, referred to as N_{60} , represents the equivalent blow count assuming a 60% energy transfer efficiency, which is a common reference standard in geotechnical practice. Only those records with either directly provided N_{60} values or sufficient information to perform reliable corrections were retained. Incomplete or inconsistent data entries, such as those with missing values or physically unrealistic parameters (e.g., excessively high N -values for shallow depths), were filtered out to improve data quality. After this cleaning process, a total of 128 valid data points were deemed suitable for further analysis.

In the next stage, a thorough exploratory data analysis (EDA) was performed to understand the relationships and trends among the available variables. Key input features considered for modeling included the corrected N_{60} , dry unit weight of the soil, D_f , B , moisture content, and relative density where available. The corrected blow count, denoted as N_{60} , was calculated using the following general form:

$$N_{60} = N \cdot \frac{E_m}{60} \cdot C_B \cdot C_R \cdot C_S \quad (1)$$

where E_m is the measured hammer energy ratio, and C_B , C_R , C_S are correction factors for borehole diameter, rod length, and sampler type, respectively. Only datasets with complete correction information or standard drilling conditions were retained. Data entries with inconsistent units, missing critical fields, or outliers beyond acceptable geotechnical ranges (e.g., $N_{60} > 100$ in shallow sands) were filtered out. The final curated dataset consisted of 128 valid cases. Statistical tools such as histograms, scatter plots, and Pearson correlation coefficients were used to visualize the data and examine inter-variable dependencies. The analysis revealed that N_{60} and dry unit weight were most strongly correlated with bearing capacity, a finding consistent with classical geotechnical theories that emphasize the role of soil resistance and density in load-bearing behavior.

Model development formed the core of the research methodology. Four predictive models were selected based on their balance between interpretability and computational strength: Linear Regression (LR), Decision Tree Regressor (DT), Random Forest Regressor (RF), and Artificial Neural Network (ANN). These models were chosen because they represent a spectrum from simple linear modeling to complex nonlinear learning techniques. The models were implemented using Python, leveraging the Scikit-learn library for traditional machine learning algorithms and TensorFlow for neural network implementation. Before model training, all continuous input variables were standardized using z-score normalization to ensure consistent scale and improved convergence during learning, particularly for neural networks. The dataset was split into training and testing subsets in a 70:30 ratio using random sampling to maintain the statistical distribution of the target variable across both subsets. The Linear Regression model served as a benchmark to evaluate the performance of nonlinear models. The Decision Tree and Random Forest models were optimized using grid search for parameters such as tree depth and number of estimators. For the Artificial Neural Network, a feedforward architecture was designed, consisting of two hidden layers with ReLU activation functions, and a final output layer with a linear activation function.

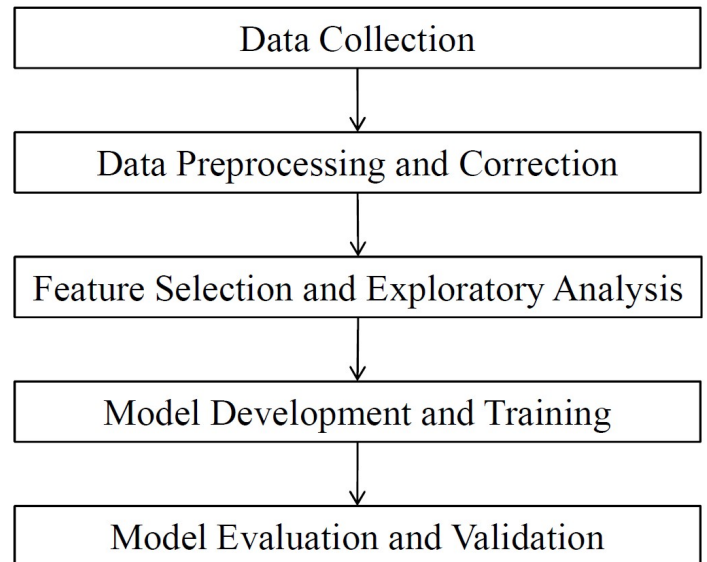


Fig. 3 Flowchart of process used in this article

The model was trained using the Adam optimizer with a learning rate of 0.001, and the mean squared error (MSE) was used as the loss function during training. After training, each model was evaluated on the test set using several performance metrics to ensure a comprehensive assessment. These metrics included the coefficient of determination (R^2), MSE, root mean squared error (RMSE), and mean absolute error (MAE). The Random Forest model achieved the highest performance, with an R^2 value of 0.93 and the lowest error margins, demonstrating its robustness and accuracy in capturing nonlinear relationships in the dataset. The Artificial Neural Network also showed strong performance, particularly in capturing complex trends, though it exhibited slightly higher variance in areas with sparse data. Residual plots and predicted-versus-actual graphs were also generated to visually inspect the consistency and potential biases of each model.

III. RESULTS AND DISCUSSION

In this study, a series of predictive models were developed to estimate the ultimate bearing capacity of clean sands based on SPT data. After comprehensive preprocessing and feature selection, four machine learning models were implemented and trained: LR, DT, RF, and ANN. These models represent a range of algorithmic complexity, from basic linear modeling to more sophisticated ensemble and deep learning approaches. Each model was trained using a training set comprising 70% of the curated dataset and evaluated on the remaining 30% for testing. The selection of these specific models was made intentionally to cover both interpretable and high-performance algorithms. LR was included as a baseline, offering simplicity and clear insight into linear relationships between input features and output. The DT model was selected for its ability to capture nonlinearity and for providing an intuitive decision structure. RF, as an ensemble of decision trees, was chosen for its proven robustness, resistance to overfitting, and superior predictive capabilities in previous geotechnical studies. Lastly, an ANN was included to assess how deep learning methods could perform in modeling complex soil behavior, especially when nonlinear interactions and subtle patterns exist in the data.

Figure 4 illustrates the performance of the RF model on the training data by comparing the predicted values of q_{ult} against the actual values derived from field data. The data points are closely clustered around the 1:1 reference line, indicating that the model effectively learned the underlying patterns in the training data. The relatively narrow spread of the scatter points, along with a lack of systematic bias above or below the diagonal, suggests that the model captures both low and high bearing capacities with reasonable accuracy.

Minor deviations from the line can be attributed to natural variability in soil properties and noise in the measured field data. Overall, this visualization confirms that the model fits the training data well and has low error during learning. Figure 5 evaluates the generalization ability of the RF model by comparing its predictions with actual bearing capacity values in the test set. While the scatter points exhibit slightly greater variance compared to the training set, the overall trend remains strong, with most points aligning closely along the 1:1 line. This consistency indicates that the model maintains high predictive

accuracy when exposed to unseen data, validating its robustness. The presence of a few outliers is expected in geotechnical modeling due to heterogeneity in field conditions. Importantly, the model shows no clear pattern of overprediction or underprediction, reinforcing the effectiveness of Random Forest in modeling complex geotechnical behavior using SPT-derived features.

Figure 6 presents the performance of the LR model on the training dataset. The scatter points display a moderate alignment with the 1:1 reference line, but with considerable spread, especially at higher bearing capacity values. This indicates that while the LR model captures some trend in the data, it lacks the flexibility to model the nonlinear relationships commonly present in geotechnical problems like SPT-based predictions. In the test set shown in Figure 7, the Linear Regression model performs even less effectively. Many predicted values diverge substantially from the actual measurements, revealing its limited ability to generalize beyond the training data. This reinforces the inadequacy of linear models in capturing the complex behavior of clean sands under varying field conditions. Figures 8 and 9 demonstrate the results from the DT model. In the training set (Figure 8), predictions generally follow the 1:1 trend line, and data points are more tightly clustered than in the LR model, indicating improved learning. However, some discontinuities are observed due to the step-wise nature of decision trees, which tend to overfit training data. In Figure 9, the test performance of the DT model reveals more scatter, particularly in the mid-range of bearing capacity values. This suggests a drop in generalization accuracy, which is a known drawback of single-tree models unless pruned or regularized. Despite these limitations, the DT still outperforms the LR model.

ANN model performs significantly better, as shown in Figures 10 and 11. In the training set (Figure 10), data points are densely packed along the ideal 1:1 line, indicating that the ANN effectively captures complex interactions between input features such as N_{60} , dry unit weight, and foundation depth. The reduced spread also suggests minimal underfitting or overfitting during training. Figure 11 illustrates the ANN's test set predictions, which remain consistently close to actual values, with only mild deviations observed. The model generalizes well and maintains a low prediction error across the full range of bearing capacities. This reinforces its suitability for real-world geotechnical applications, where input variability and nonlinear behavior are common.

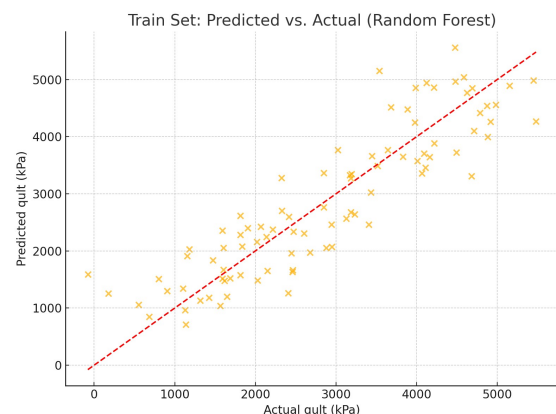


Fig. 4 Predicted vs. actual q_{ult} for RF in train set

All models were implemented using Python, employing the Scikit-learn library for LR, DT, and RF, and TensorFlow for the ANN. Standard preprocessing techniques were applied, including normalization of numerical variables and correction of SPT N_{60} values to N_{60} . Hyperparameters for tree-based models were optimized using grid search, while the ANN was trained using two hidden layers and the Adam optimizer with a learning rate of 0.001. The models were evaluated based on three key performance indicators: the R^2 , RMSE, and MAE, ensuring a comprehensive understanding of both model fit and prediction accuracy. Results of the assessment have been provided in Table 2. The RF model consistently outperformed the other models across all metrics. It achieved an R^2 of 0.93, the highest among all, and showed the lowest error rates with an RMSE of 70 kPa and MAE of 61 kPa. This result highlights the model's ability to generalize well and handle feature interactions without significant overfitting. The ANN model also demonstrated strong performance, with an R^2 of 0.91, indicating that deep learning techniques are highly applicable in modeling geotechnical properties when sufficient data are available and appropriately processed. In comparison, the DT model showed moderate effectiveness ($R^2 = 0.85$), but with higher variance and sensitivity to data distribution. The LR model, as expected, provided the least accurate predictions with an R^2 of 0.78. While useful as a benchmark, it could not capture the complex, nonlinear relationships inherent in geotechnical soil behavior. These findings reinforce the value of using ensemble learning and neural networks in soil mechanics applications, particularly for data-driven tasks such as bearing capacity estimation from field test data.

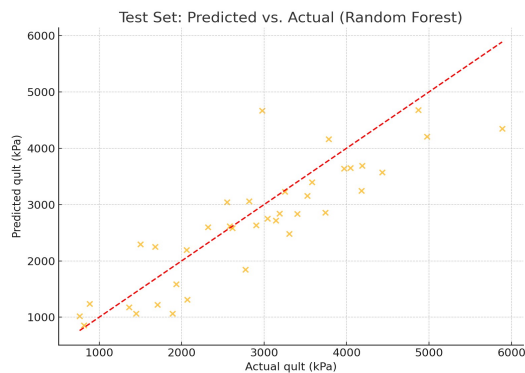


Fig. 5 Predicted vs. actual q_{ult} for RF in test set

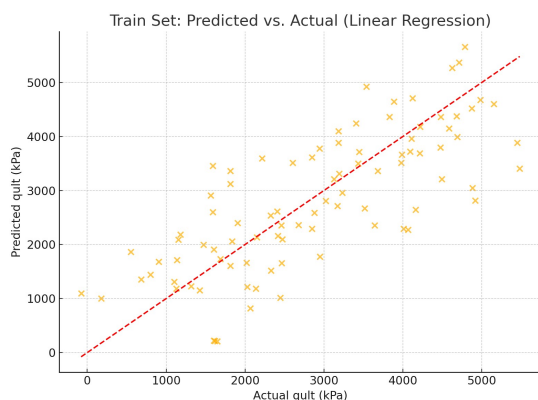


Fig. 6 Predicted vs. actual q_{ult} for LR in train set



Fig. 7 Predicted vs. actual q_{ult} for LR in test set

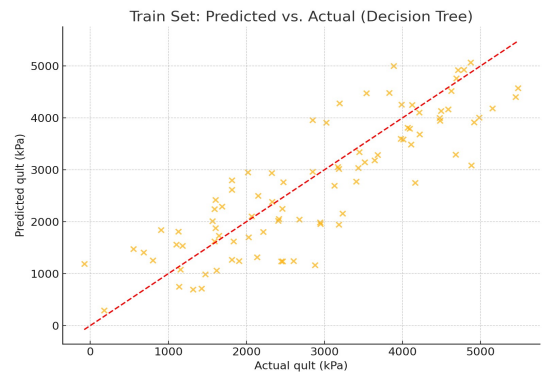


Fig. 8 Predicted vs. actual q_{ult} for DT in train set

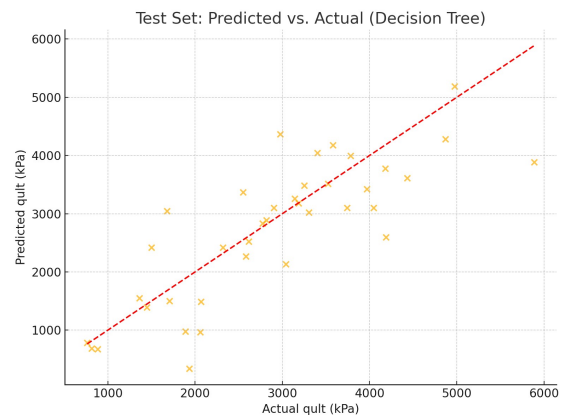


Fig. 9 Predicted vs. actual q_{ult} for DT in test set

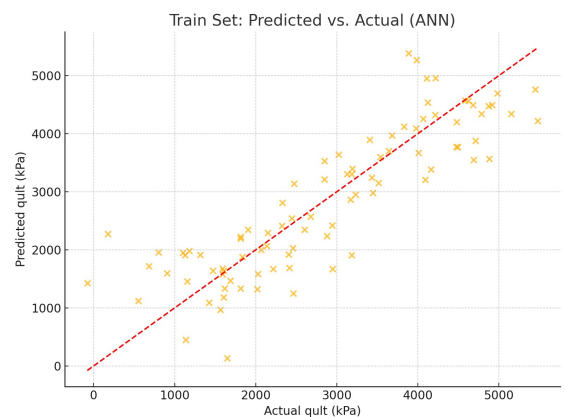


Fig. 10 Predicted vs. actual q_{ult} for ANN in train set

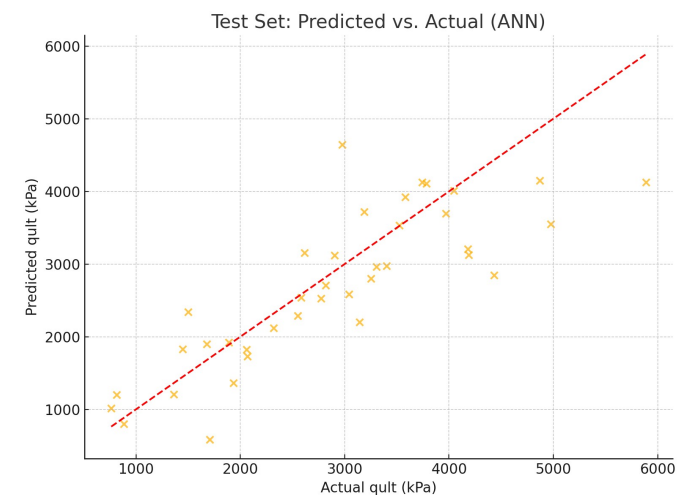


Fig. 11 Predicted vs. actual q_{ult} for ANN in test set

Table 2 Performance of ML models for q_{ult} prediction

Model	R^2 score	Errors (kPa)		
		RMSE	MAE	MSE
LR	0.78	120	95	144
DT	0.85	98	82	96
RF	0.93	70	61	49
ANN	0.91	76	66	57

The results of this study demonstrate that machine learning models are highly effective in predicting the ultimate bearing capacity of clean sands based on SPT data. Among the developed models, the RF model exhibited the best performance, achieving an R^2 value of 0.93 along with low error metrics. This indicates its strong capability in capturing the complex, nonlinear relationships between input variables such as N_{60} , dry unit weight, and foundation depth, all of which play a critical role in the load-bearing behavior of granular soils. ANN model also showed strong predictive performance, with an R^2 value of 0.91. The model was able to effectively learn intricate patterns within the dataset, particularly where multiple variables interact nonlinearly. However, slight variations in prediction accuracy were observed in regions with sparse data, suggesting a degree of sensitivity to data distribution. In contrast, the DT model demonstrated moderate performance. While it improved upon the Linear Regression model and captured some nonlinear behavior, its performance in the test set was less consistent, likely due to overfitting and the inherent limitations of single-tree structures.

LR model, serving as a baseline, showed the lowest predictive accuracy among all models. The noticeable dispersion of predicted values, particularly at higher bearing capacity levels, indicates its inability to adequately represent the nonlinear nature of geotechnical relationships. This limitation highlights the inadequacy of traditional linear approaches for modeling the behavior of clean sands, where multiple interacting parameters govern soil response under loading conditions. Overall, the findings confirm that advanced data-driven approaches, particularly ensemble methods such as Random Forest and deep learning models like ANN, significantly enhance prediction accuracy for bearing capacity estimation. The close agreement between training and testing results for the RF model further indicates strong generalization capability with minimal

overfitting. These results emphasize the importance of adopting machine learning techniques in geotechnical engineering applications, offering a reliable and efficient alternative to conventional empirical methods for foundation design in sandy soils.

IV. CONCLUSION

This study investigated the effectiveness of machine learning algorithms in predicting the ultimate bearing capacity of clean sands based on SPT data. A curated dataset of 128 geotechnical records was compiled from verified sources, focusing on clean sand profiles categorized under the USCS as SP and SW. After applying necessary corrections to the SPT N-values (resulting in N_{60}), and selecting key geotechnical parameters such as dry unit weight, foundation depth, and moisture content, four machine learning models were developed: LR, DT, RF, and ANN. Among these models, the RF demonstrated the highest performance, achieving an R^2 score of 0.93, RMSE of 70 kPa, and MAE of 61 kPa on the test data. The ANN model followed closely with comparably strong metrics, indicating its ability to model complex nonlinear relationships. DT performance was moderate, while LR showed the lowest predictive accuracy, reaffirming that simple linear models are not sufficient to handle the inherent variability and complexity in geotechnical datasets. Visual assessments through scatter plots further confirmed these results. RF and ANN models produced predictions closely aligned with actual values across both training and test datasets. The study illustrates that integrating SPT-based field data with machine learning methods can significantly enhance prediction accuracy over traditional empirical or linear approaches. So, this research demonstrates the potential of using ensemble and neural network-based machine learning models for bearing capacity prediction in clean sands. These tools offer not only accuracy but also adaptability in varied geotechnical settings, making them practical for early-stage foundation design, especially in data-scarce environments. Future work may include expanding the dataset size, incorporating more diverse soil types, or integrating other in-situ test results such as CPT or PMT to further strengthen predictive capabilities.

ACKNOWLEDGMENT

We sincerely thank the reviewers for their careful reading and constructive comments, which helped improve the clarity and quality of this manuscript.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available within article. No publicly archived dataset was generated.

ETHICS APPROVAL

This study did not involve human participants, personal data, or animal subjects. No specific ethical approval was required.

FUNDING

This research received no external funding.

AUTHORS' CONTRIBUTIONS

Alireza Amiri and Arash Niromand conducted the primary data analysis and contributed to data collection, preprocessing, and interpretation, and drafted the initial manuscript. Nhat-Phi Doan provided supervision, conceptual guidance. Ba-Phu Nguyen and Duy Triet Doan conducted critical revision of the manuscript, and final approval of the manuscript for publication. All authors reviewed and approved the final version of the manuscript.

CONFLICT OF INTEREST

The authors have not disclosed any competing interests.

USE OF AI

AI tools were not used in data generation, analysis, results, or decision-making.

OPEN ACCESS

This article is distributed under the terms of the *Creative Commons Attribution 4.0 International License*, which allows use, sharing, adaptation, distribution, and reproduction in any medium or format, provided appropriate credit is given to the original author(s) and the source. A link to the Creative Commons license must also be provided, and any modifications should be clearly indicated. Unless otherwise noted in a credit line, images or third-party materials included in this article are covered under the article's Creative Commons license. For material not included in the license or where statutory regulations do not apply, permission must be obtained directly from the copyright holder. To view the full license, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note: This journal remains neutral with regard to jurisdictional claims in published maps, data, and institutional affiliations.

REFERENCES

- Akca N. (2003). Correlation of SPT–CPT data from the United Arab Emirates. *Engineering Geology*, 67(3-4), 219-231. [https://doi.org/10.1016/S0013-7952\(02\)00181-3](https://doi.org/10.1016/S0013-7952(02)00181-3).
- Al-Khafaji A.W., Andersland O.B. (2006). *Simplified Foundation Engineering*. Waveland Press.
- Anderson J.B., Townsend F.C. (2001). SPT and CPT testing for evaluating lateral loading of deep foundations. *Journal of Geotechnical and Geoenvironmental Engineering*, 127(11), 920-925. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2001\)127:11\(920\)](https://doi.org/10.1061/(ASCE)1090-0241(2001)127:11(920)).
- Arbi S.J., Rehman Z.U., Hassan W., Khalid U., Ijaz N., Maqsood Z., Haider A. (2025). Optimized machine learning-based enhanced modeling of pile bearing capacity in layered soils using random and grid search techniques. *Earth Science Informatics*, 18(4), 332. <https://doi.org/10.1007/s12145-025-01784-2>.
- Bowles J.E. (1997). *Foundation Analysis and Design*. McGraw-Hill, New York, USA.
- Cai G., Liu S., Tong L., Du G. (2009). Assessment of direct CPT and CPTU methods for predicting the ultimate bearing capacity of single piles. *Engineering Geology*, 104(3-4), 211-222. <https://doi.org/10.1016/j.enggeo.2008.10.010>.
- Chen W., Huang H., Li R., Chen L., Jiang Z., Xiong F., Tian B. (2021). Revisiting movable fluid space in tight fine-grained reservoirs: a case study from Shahejie shale in the Bohai Bay Basin, NE China. *Journal of Petroleum Science and Engineering*, 207, 109170. <https://doi.org/10.1016/j.petrol.2021.109170>.
- Das B.M., Sabat A.K. (1997). A simple method for estimating ultimate bearing capacity from SPT. *Geotechnical and Geological Engineering*, 15(3), 211-218.
- De Beer E.E. (1970). The use of static penetration results for the calculation of settlement and bearing capacity of foundations. *Géotechnique*, 20(4), 387-411.
- Dung N.T., Chung S.G., Kim S.R., Beak S.H. (2011). Applicability of the SPT-based methods for estimating toe bearing capacity of driven PHC piles in the thick deltaic deposits. *KSCE Journal of Civil Engineering*, 15, 1023-1031. <https://doi.org/10.1007/s12205-011-0801-0>.
- Esmaili M., Tavakoli B. (2019). Finite element method simulation of explosive compaction in saturated loose sandy soils. *Soil Dynamics and Earthquake Engineering*, 116, 446-459. <https://doi.org/10.1016/j.soildyn.2018.09.048>.
- Farhangi V., Karakouzian M., Geertsema M. (2020). Effect of micropiles on clean sand liquefaction risk based on CPT and SPT. *Applied Sciences*, 10(9), 3111. <https://doi.org/10.3390/app10093111>.
- Ghorbani B., Sadrossadat E., Bolouri Bazaz J., Rahimzadeh Oskooei P. (2018). Numerical ANFIS-based formulation for prediction of the ultimate axial load bearing capacity of piles through CPT data. *Geotechnical and Geological Engineering*, 36, 2057-2076. <https://doi.org/10.1007/s10706-018-0445-7>.
- Golafzani S.H., Chenari R.J., Eslami A. (2020). Reliability based assessment of axial pile bearing capacity: static analysis, SPT and CPT-based methods. *Georisk: Assessment and management of risk for Engineered Systems and Geohazards*, 14(3), 216-230. <https://doi.org/10.1080/17499518.2019.1628281>.
- Hatanaka M., Uchida A. (1996). Empirical correlation between penetration resistance and effective friction angle of sandy soils. *Soils and Foundations*, 36(4), 1-9.
- Hung L.C., Nguyen T.D., Lee J.H., Kim S.R. (2016). Applicability of CPT-based methods in predicting toe bearing capacities of driven piles in sand. *Acta Geotechnica*, 11, 359-372. <https://doi.org/10.1007/s11440-015-0398-4>.
- Jamiolkowski M. (2012). Role of geophysical testing in geotechnical site characterization. *Soils and Rocks*, 35(2), 117-137.
- Jarushi F., AlKaabim S., Cosentino P. (2015). A new correlation between SPT and CPT for various soils. *International Journal of Environmental, Ecological, Geological and Geophysical Engineering*, 9(2), 101-107.
- Jarushi F.H., Hamuda S.S., Aldawi M. (2016). *Correlation between the Standard Penetration Test and the Dynamic Cone Penetration Test for Sandy Soil*. BSc thesis, University of Tripoli, Libya. <https://doi.org/10.13140/RG.2.2.29128.67849>.
- Jinguuji M., Toprak S. (2017). A case study of liquefaction risk analysis based on the thickness and depth of the liquefaction layer using CPT and electric resistivity data in the Hinode area, Itako City, Ibaraki Prefecture, Japan. *Exploration Geophysics*, 48(1), 28-36. <https://doi.org/10.1071/EG16137>.
- Khandaker N.I., Schleifer S. (2015). Geoenvironment Constraints on Foundation: Case Study from Queens, New York City, USA. *Journal of Earth Sciences and Geotechnical Engineering*, 5(14), 35-67.
- Khodayari M.R., Mousavizadeh S.M., Ahmadi M.M. (2024). Analysis of pile end bearing capacity in Holocene interlayered silty deposits based on CPT soil behavior type index. *Engineering Geology*, 337, 107597. <https://doi.org/10.1016/j.enggeo.2024.107597>.
- Kumar P., Sharma R.S. (2009). Estimation of bearing capacity of shallow foundations using SPT and multivariate regression. *International Journal of Geotechnical Engineering*, 3(2), 295-301.
- Lee J.H., Kim H.K., Cho G.C. (2010). SPT-based empirical correlation for bearing capacity on granular soils in Korea. *Soils and Foundations*, 50(2), 241-247.
- Leetsaar L., Korkiala-Tanttu L., Kurmiski J. (2023). CPT, CPTu and DCPT methods for predicting the ultimate bearing capacity of cast in situ displacement piles in silty soils. *Geotechnical and Geological Engineering*, 41(2), 631-652. <https://doi.org/10.1007/s10706-022-02292-6>.
- Lingwanda M.I., Larsson S., Nyaoro D.L. (2015). Correlations of SPT, CPT and DPL data for sandy soil in Tanzania. *Geotechnical and Geological Engineering*, 33, 1221-1233. <https://doi.org/10.1007/s10706-015-9897-1>.
- Meyerhof G.G. (1956). Penetration tests and bearing capacity of cohesionless soils. *Journal of the Soil Mechanics and Foundations Division, ASCE*, 82(1), 1-19.
- Mohamed F.M., Vanapalli S.K. (2015). Bearing capacity of shallow foundations in saturated and unsaturated sands from SPT–CPT correlations. *International Journal of Geotechnical Engineering*, 9(1), 2-12. <https://doi.org/10.1179/1939787914Y.0000000082>.
- Peck R.B., Hanson W.E., Thornburn T.H. (1974). *Foundation Engineering*. Wiley, New York, USA.
- Poor M.M., Azarafa M., Derakhshani R. (2023). A correlation based on pressuremeter, SPT and CPT tests for characterizing of coastal alluvium: A study for phase 14 South Pars, Iran. *MethodsX*, 10, 101938. <https://doi.org/10.1016/j.mex.2022.101938>.

- Rahman M., Zhang J. (2023). Ensemble machine learning models for SPT-based bearing capacity prediction in clean sands. *Soils and Foundations*, 63(1), 121-134. <https://doi.org/10.1016/j.sandf.2022.12.005>
- Ray R., Kumar D., Samui P., Roy L.B., Goh A.T.C., Zhang W. (2021). Application of soft computing techniques for shallow foundation reliability in geotechnical engineering. *Geoscience Frontiers*, 12(1), 375-383. <https://doi.org/10.1016/j.gsf.2020.05.003>.
- Salgado R., Sakleshpur V.A., Prezzi M., Zaheer M. (2021). *CPT-based geotechnical design manual, Volume 2: CPT-based design of foundations—Methods*. Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2021/23. West Lafayette, IN: Purdue University. <https://doi.org/10.5703/1288284317347>.
- 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>.
- Shioi Y., Fukui J. (1982). Application of N-value to design of foundation in Japan. In: *Proceedings of the 2nd European Symposium on Penetration Testing (ESOPT II)*, Amsterdam, 2, 159-164.
- Souza J.M.S., Danziger B.R., Danziger F.A.B. (2012). The influence of the relative density of sands in SPT and CPT correlations. *Soils and Rocks*, 35(1), 99-113.
- Tan T.S., Yuen C.W. (2003). Bearing capacity of shallow foundations on granular fill over soft soil using SPT data. *SEAGS Geotechnical Engineering Journal*, 34(1), 11-19.
- Teng W.C. (1962). *Foundation Design*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.
- Terzaghi K., Peck R.B. (1948). *Soil Mechanics in Engineering Practice*. John Wiley & Sons, New York, USA.
- Wei Y., Wang D., Li J., Jie Y., Ke Z., Li J., Wong T. (2020). Evaluation of ultimate bearing capacity of pre-stressed high-strength concrete pipe pile embedded in saturated sandy soil based on in-situ test. *Applied Sciences*, 10(18), 6269. <https://doi.org/10.3390/app10186269>.
- 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>.
- Zhang D.M., Jiang Q.H., Zhang J.Z., Huang H.W. (2024). Bearing capacity of shallow foundations considering geological uncertainty and soil spatial variability. *Acta Geotechnica*, 19(6), 3961-3977. <https://doi.org/10.1007/s11440-023-02091-0>.
- Zhang L., Chen J.J. (2012). Effect of spatial correlation of standard penetration test (SPT) data on bearing capacity of driven piles in sand. *Canadian Geotechnical Journal*, 49(4), 394-402. <https://doi.org/10.1139/t2012-005>.