Nikola Božović[1], Andrija Petrović[2], Marija Krstić[3], Mladen Ćosić[1], Sanja Jocković[4], Miloš Marjanović[4]*

PILE BEARING CAPACITY AND SETTLEMENT ESTIMATION BASED ON OPTIMISED MACHINE LEARNING MODELS

PROCENA NOSIVOSTI I SLEGANJA ŠIPOVA OPTIMIZOVANIM MODELIMA MAŠINSKOG UČENJA

Originalni naučni rad / Original scientific paper
Rad primljen / Paper received: 29.12.2025
https://doi.org/10.69644/ivk-2026-02-0233
Adresa autora / Author's address:
[1)] Institute for Testing of Materials-IMS, Belgrade, Serbia
N. Božović https://orcid.org/0009-0006-3804-6524 ,
*email: nikola.bozovic@institutims.rs
M. Ćosić https://orcid.org/0000-0002-2581-2692
[2)] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia
[3)] Elixir Engineering, Šabac, Serbia
[4)] University of Belgrade, Faculty of Civil Engineering, Belgrade, Serbia S. Jocković https://orcid.org/0000-0002-3896-4791 ;
M. Marjanović https://orcid.org/0000-0002-7968-3873

Keywords
- pile bearing capacity
- pile settlement
- pile testing
- machine learning
- geotechnical data

Ključne reči
- nosivost šipa
- sleganje šipa
- ispitivanje šipa
- mašinsko učenje
- geotehnički podaci

Abstract

Determining pile bearing capacity and settlement using analytical procedures commonly applied in engineering practice requires validation through field testing, either by static or dynamic methods. However, test results frequently exhibit significant deviations from calculated values. Since pile testing is both costly and time-consuming, this study investigates the application of eleven machine-learning algorithms to predict pile bearing capacity and settlement based on soil data at pile locations. The soil dataset is divided into laboratory and field data, forming four prediction approaches. The applied machine-learning algorithms include linear regression models, tree-based models, distance- and kernel-based models, as well as ensemble and boosting models. The database contains information on 170 piles tested both by static and dynamic methods within the Belgrade area. Algorithm performance is evaluated using R^2 values, with the CatBoost algorithm achieving the highest accuracy for both capacity and settlement prediction. For the best-performing model, the most influential parameters are identified for all four prediction approaches. The cross-sectional area of the pile is found to be the dominant factor in bearing-capacity prediction, while the applied load has the greatest influence on settlement prediction. Finally, conventional analytical methods for bearing capacity and settlement are performed for all piles in the database. The results obtained using analytical calculation methods often exhibit significant deviations from those derived from pile load testing. By incorporating results from SLT and DLT and linking them with field and laboratory data, machine learning provides more advanced solutions compared to conventional analytical methods.

Izvod

Određivanje nosivosti i sleganja šipova primenom standardnih analitičkih postupaka u praksi zahteva validaciju kroz terenska ispitivanja statičkim ili dinamičkim metodama. Međutim, rezultati ovih ispitivanja često pokazuju značajna odstupanja u odnosu na proračunske vrednosti, dok su sama ispitivanja skupa i vremenski zahtevna. U ovom radu primenjeno je jedanaest algoritama mašinskog učenja za predviđanje nosivosti i sleganja šipova na osnovu parametara tla u kojem su šipovi izvedeni. Podaci o tlu podeljeni su na laboratorijske i terenske, na osnovu čega su formirana četiri pristupa predviđanja. Primenjeni algoritmi mašinskog učenja obuhvataju linearne modele, modele zasnovane na drvetu odlučivanja, modele zasnovane na udaljenosti i jezgru, kao i ansambl i „boosting“ modele. Baza podataka korišćena u analizi obuhvata informacije o 170 šipova ispitanih statičkim i dinamičkim metodama na području Beograda. Efikasnost algoritama ocenjena je pomoću koeficijenta determinacije R^2, pri čemu je CatBoost algoritam postigao najveću tačnost u predviđanju nosivosti i sleganja. Za ovaj algoritam određeni su i najuticajniji ulazni parametri za sva četiri pristupa predviđanja. Kao dominantan parametar u predviđanju nosivosti izdvojila se površina poprečnog preseka šipa, dok se u predviđanju sleganja kao najuticajnija pokazala sila u šipu. Za sve šipove iz baze izvršen je i proračun nosivosti i sleganja primenom konvencionalnih analitičkih metoda. Rezultati proračuna analitičkim metodama često značajnije odstupaju od rešenja koja se dobijaju ispitivanjem šipova. Primenom rešenja dobijenih SLT i DLT ispitivanjem šipova i povezivanjem sa terenskim i laboratorijskim podacima, mašinsko učenje obezbeđuje naprednija rešenja u odnosu na analitičke metode.

List of abbreviations

Abbreviation	Full name
AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference system
ANN	Artificial Neural Network
BiLSTM	Bidirectional Long Short-Term memory
CB	CatBoost
CNN	Convolutional Neural Network
DLT	Dynamic Load Test

DNNs	Deep Neural Networks
DT	Decision Tree
ER	Elastic Net Regression
GB	Gradient Boosting
GBM	Gradient Boosting Machine
GEP	Gene Expression Programming
GNN	Graph Neural Network
KNN	K-Nearest Neighbours
LR	Linear Regression
LsR	Lasso Regression
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
MSE	Mean Squared Error
PINNs	Physics-Informed Neural Networks
PSO	Particle Swarm Optimization
R2	Coefficient of Determination
RF	Random Forest
RNN	Recurrent Neural Network
RR	Ridge Regression
RVMs	Relevance Vector Machines
SHAP	SHapley Additive exPlanations
SLT	Static Load Test
SVM	Support Vector Regressor
XGB	XGBoost

INTRODUCTION

Pile capacity prediction is a critical task in geotechnical engineering, where reliable predictions can significantly improve design efficiency and reduce costs. Traditional methods rely on empirical formulas, field testing, or laboratory experiments. However, these methods are often time-consuming, expensive and limited by the assumptions inherent in the models.

Over the years, various experimental and theoretical methods have been developed to evaluate pile capacity and settlement, yet each approach has limitations. Among these, the Static Load Test (SLT) is widely used for determining pile behaviour under working loads, although expensive and time-consuming. Dynamic load test (DLT) has gained global acceptance as an alternative, leveraging wave propagation theory. Numerous studies confirm a strong correlation between capacity estimates from DLT and SLT, demonstrating DLT reliability and efficiency (Likins and Rausche /28/; Souza et al. /1/, Cuajao et al. /2/). Nonetheless, DLT may require multiple tests per project which can increase costs and extend schedules.

To overcome these challenges, engineers and researchers increasingly employ artificial intelligence (AI) and machine learning methods that excel in identifying complex, nonlinear relationships within geotechnical data (e.g., Das and Basudhar /3/; Lee and Lee /4/). Recent studies reveal promising, ML-based solutions for enhancing pile capacity predictions. Reliable bearing capacity estimation is crucial for optimising geotechnical designs and minimising overall costs. While traditional techniques rely on empirical formulas and field tests with inherent assumptions, ML approaches offer adaptability and robustness, paving the way for more efficient and reliable prediction of pile capacity.

Artificial Neural Networks (ANN) rank among the earliest machine learning (ML) methods introduced in geotechnical engineering. Goh /5/ pioneered their use, demonstrating ANN's effectiveness in modelling complex nonlinear behaviour. Still, these models require meticulous hyperparameter tuning to prevent overfitting or underfitting. For example, Ghorbani et al. /6/ develops an adaptive neuro-fuzzy inference system (ANFIS, a form of ANN-based model) to correlate cone penetration test (CPT) results with ultimate pile capacity, achieving a correlation coefficient around 0.96 on a dataset of 108 case histories. Studies have also combined ANNs with metaheuristic optimisation to enhance accuracy. Armaghani et al. /7/ introduces a hybrid PSO-ANN (particle swarm optimisation optimised ANN) to forecast the settlement of rock-socketed piles, obtaining an R^2 of ~ 0.89 on test data. These results confirm that ANN-based approaches can capture the nonlinear relationships between soil properties, pile characteristics, and bearing behaviour with high fidelity. In parallel, Support Vector Machines (SVM) have gained attention for pile capacity prediction due to their robustness in high-dimensional spaces (Pal and Deswal /8/; Samui /9/).

While some researchers find SVM models to be outperformed by tree-based ensembles, others report competitive performance when SVM hyperparameters are well-tuned. For instances, recent research has demonstrated that SVM can effectively predict pile bearing capacity with competitive accuracy compared to conventional methods and ensemble models (Galib et al. /29/). With careful feature selection and parameter calibration, SVMs can be potent tools for pile capacity prediction even though they are somewhat less common in recent literature than ANN or ensemble models.

Several studies have shown that ensemble methods, such as Random Forest (RF), Gradient Boosting Machines (GBM), and XGBoost, offer high accuracy in predicting pile capacity. Kardani et al. /10/ optimised six ML algorithms using particle swarm optimisation, where PSO-tuned XGBoost achieved the best accuracy ($R^2 = 0.9615$). Amjad et al. /13/ confirmed the superiority of XGBoost over other ML models using a dataset of 200 driven piles. Arbi et al. /14/ extends this analysis to layered soils, using grid and random search techniques to tune XGBoost, RF and SVM models. XGBoost again emerges as the best performer.

Deep learning approaches have become increasingly popular. Pham et al. /11/ combines genetic algorithms with deep neural networks (DNNs) for architecture optimisation, achieving excellent results on a dataset of 472 driven piles. Benbouras et al. /12/ tested 11 ML algorithms and found DNNs outperform others in predicting driven pile capacity. Kumar et al. /15/ compares DNN, CNN, RNN, LSTM, and BiLSTM, with DNN providing the highest accuracy. PINNs (Physics-Informed Neural Networks) are introduced by Al-Atroush /16/ to integrate theoretical geotechnical equations into the training loss, improving interpretability and generalisation.

Hybrid ML models leveraging evolutionary algorithms have also demonstrated strong performance. Pham et al. /11/ used genetic algorithms for feature selection and DNN optimisation. Khan et al. /17/ employs gene expression programming (GEP) to derive an interpretable explicit formula for axial pile capacity and uses SHAP for model interpretation. Khatti et al. /18/ tested 33 variants of relevance vector machines (RVMs) optimised via genetic and particle swarm algorithms for modelling time-dependent capacity in cohesive soils.

A growing trend is integrating geotechnical knowledge into ML models. Xiong et al. /30/ develops a physics-informed ML framework incorporating soil-pile interaction features to enhance model reliability. Al-Atroush /16/ uses PINNs to embed Meyerhof's equations into the ML model. These methods show promise in improving the generalisation and trustworthiness of ML models in engineering practice.

Igoe et al. /19/ train ML models on synthetic data from finite element simulations of screw piles, with Gaussian Process Regression outperforming traditional theoretical models. Seo et al. /20/ develops ML models for precast piles using data that capture disturbance and construction effects, further expanding ML's applicability beyond traditional empirical datasets.

Thottoth et al. /22/ investigates the ultimate compression capacity of under-reamed bored piles in both clay and sand strata. They train ANN and RF models on a dataset of instrumented under-reamed pile load tests and report prediction performances with R^2 exceeding 0.80 - a positive result given the complex geometry and load-transfer mechanism of under-reamed piles. Helical piles, which are screw-like piles often used in soft or expansive soils, have also been studied via ML. Shuman et al. /23/ compiles a database of full-scale load tests on large-diameter helical piles and supplements it with numerical simulation data (about 3,600 finite element analyses covering varied soil properties and pile configurations) to broaden the range of conditions. Using this augmented dataset, they develop ANN, decision tree, and RF models to predict pile settlement under service loads. The best results come from ensemble methods, with the RF, achieving an $R^2 \sim 0.96$ for the largest pile diameters. This study not only highlights the adaptability of ML models to helical pile design but also demonstrates a practical approach to overcome data scarcity by incorporating synthetic data.

Multiple studies target specific pile types or engineering problems. Palalane and Dithinde /21/ focus on bored piles in sandy soils. Cao et al. /24/ predict the stiffness and residual bearing capacity of bridge pile groups using dynamic test data. Onyelowe et al. /25/ develop ML models for piles socketed in rock, using SHAP analysis to highlight critical features.

This paper presents the application of ML algorithms for predicting the bearing capacity and settlement of piles. Four prediction approaches are conducted, based on input parameters divided into three groups, according to how they are obtained. These groups reflect the conventional analytical procedures typically employed in pile design. These approaches are: (i) LAB, which applies parameters obtained in the laboratory; (ii) CPT, based on cone resistance parameters from the CPT test; (iii) SPT, that relies on the number of blows required for a specific cone penetration, as determined by the SPT test, and (iv) SUMA, which integrates all parameters obtained through previously described methods.

RESEARCH METHODOLOGY

The methodological framework adopted in this study combines systematic data preprocessing, multi-model training, and cross-validation to identify the best performing machine learning models for predicting pile bearing capacity and settlement. The entire workflow is implemented in Python using the scikit-learn, XGBoost, and CatBoost libraries. Figure 1 illustrates the step-by-step methodological flow adopted in this research. The whole process consists of four different stages: data acquisition and structuring, application of ML algorithms, evaluation of results; and finally interpretation of ML models.

The first stage involves collecting, cleaning, and structuring the geotechnical data used for model training. The database incorporates results from field investigations and laboratory testing, and as previously described, is organised into four prediction approaches (LAB, CPT, SPT & SUMA). Each database entry corresponds to a single pile, with soil parameters averaged along the shaft in proportion to layer thickness. This standardised representation ensures consistent feature scaling and reduces data sparsity across the models.

In the second stage, eleven supervised learning algorithms are applied to predict two target variables: the bearing capacity (Y_1) and the settlement (Y_2) of piles.

The force (Q) and settlement (s) values used for model training and validation are obtained from static and dynamic load tests performed on piles. The predicted force corresponds to the bearing capacity that the pile must satisfy under actual service conditions, without the application of safety factors (i.e., serviceability limit state). The predicted settlement is then computed based on this force.

In most cases, the ultimate limit state is not reached during static or dynamic pile load testing. The reliability of extrapolation methods for ultimate bearing capacity depends on the achieved settlement level: the closer this level is to the ultimate state, the greater the reliability. When constructing a structural model, it is essential to know the actual settlement under working load conditions.

The models cover four major ML families: a) Linear regression models: Linear (LR), Ridge (RR), Lasso (LsR) and Elastic Net (ER) regression models; b) Tree-based models: Decision Tree (DT) and Random Forest (RF); c) Kernel- and distance-based models: Support Vector Regressor (SVM) and K-Nearest Neighbours (KNN); d) Ensemble and boosting models: Gradient Boosting (GB), XGBoost (XGB), and CatBoost (CB).

All models are implemented through a pipeline that integrates preprocessing and regression into a single workflow. This ensures that scaling, encoding, and training are consistently applied within each cross-validation fold, thus avoiding data leakage. Hyperparameter optimisation is performed using GridSearchCV, where parameter combinations are evaluated based on the negative mean squared error (MSE) criterion. The tuning process identified the most suitable configurations for each algorithm and dataset variant.

The third stage focuses on evaluating and comparing the predictive performance of all models. To ensure robustness, a five-fold cross-validation (KFold) procedure is employed allowing every sample to serve as both training and validation data across iterations. Three widely used regression metrics are computed for each model: coefficient of determination (R^2), mean squared error (MSE) and mean absolute error (MAE). Although MSE and MAE are commonly applied in such analyses, this study reports comparative algorithm performance exclusively through the R^2 metric.

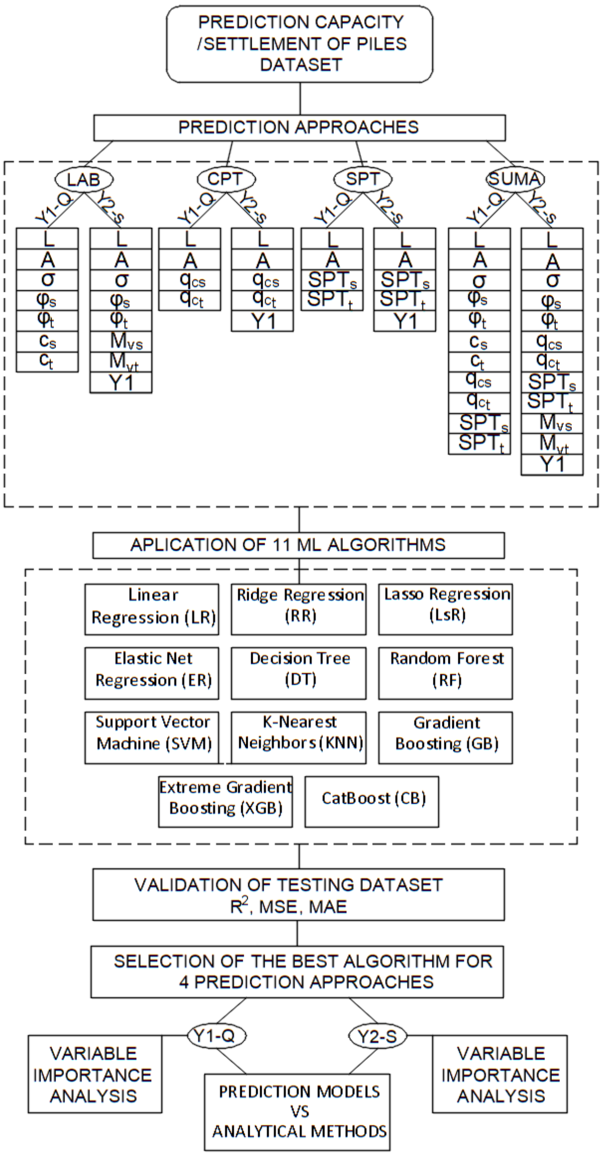


Figure 1. Methodological flowchart of this study.

The fourth stage of the methodological framework focuses on the interpretation of the trained machine learning models using Shapley Additive Explanations (SHAP). SHAP is employed as the sole interpretability technique in this study to ensure a transparent understanding of how each input feature influences the predicted pile bearing capacity and settlement. After identifying the best-performing algorithm based on the highest mean R^2 value, the model is retrained on the entire dataset using the same preprocessing workflow applied during the training phase. The SHAP explainer is then used to compute Shapley values for each observation and each input variable, quantifying their contribution, both in magnitude and direction to the model's output, relative to the global mean prediction. A SHAP summary plot is generated to visualise the overall importance and effect of individual features. This analysis provides global interpretability, by ranking the most influential parameters across the entire dataset. By integrating SHAP analysis as the final step, the fourth stage ensures that the machine learning results are not only accurate, but also explainable and aligned with established geotechnical principles, reinforcing the engineering credibility of the proposed approach.

DATA SUMMARY AND PREPARATION

For the purposes of this research, a database is compiled consisting of 170 samples (i.e., piles). All piles underwent load testing to confirm the calculation models of piles for construction of various structures within the territory of Belgrade. The pile execution technologies included bored piles, Franki piles, CFA piles, SDP piles, and driven piles.

Prior to testing each of the 170 piles, a full geotechnical investigation is performed, including geotechnical boreholes with sampling and laboratory analysis, as well as conducting field cone (CPT) and standard (SPT) penetration tests. Based on the obtained data, the stratigraphy along each pile and the corresponding layer depths are determined. All parameters along the pile shaft are averaged proportionally to the thicknesses of the corresponding soil layers.

As a result, four prediction approaches are established (LAB, CPT, SPT, and SUMA) as shown in Table 1.

Table 1. Summary of prediction approaches and input parameters.

Approach	Input parameters	Source of parameters
LAB	ϕ_s (friction angle at shaft) ϕ_t (friction angle at tip) c_s (cohesion at shaft)* c_t (cohesion at tip)* σ_v (effective overburden pressure at tip) M_{vs} (compressibility modulus at shaft)** M_{vt} (compressibility modulus at tip)** L (pile length) A (cross-sectional area)	Laboratory tests on field samples
CPT	q_{cs} (cone resistance at shaft) q_{ct} (cone resistance at tip) L, A	Cone Penetration Test (CPT)
SPT	SPT_s (number of SPT blows at shaft) SPT_t (number of SPT blows at tip) L, A	Standard Penetration Test (SPT)
SUMA	LAB, CPT, and SPT models combined	Combination of all three sources

* Only for bearing capacity prediction
** Only for settlement prediction

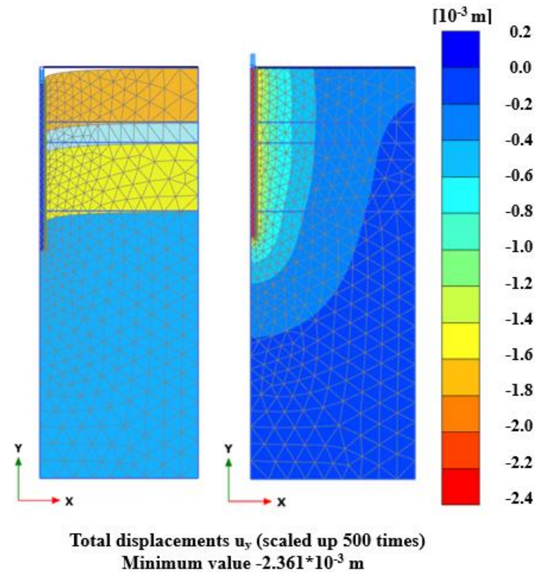
For the prediction of settlement, in addition to the aforementioned input parameters, the measured load from the load tests also serves as an input parameter for all four approaches. In the LAB and SUMA approaches, the compressibility modulus is added as an input parameter, as it directly governs the deformation response relevant for settlement calculations. On the other hand, cohesion is excluded since it represents apparent, and generally unreliable, strength parameterisation, and is not relevant for predicting settlement. The shear strength of the soil, defined by Mohr-Coulomb's law, is implemented in bearing capacity calculations and therefore retains as input parameters for predicting pile bearing capacity.

Using these input parameters, pile bearing capacity and settlement predictions are carried out by applying eleven machine learning algorithms according to the measured load (Y_1) and settlement (Y_2) data obtained from the pile load tests. Table 2 summarises the resulting database.

Table 2. Summary of dataset.

	Count	Mean	SD	Min	Max
L (m)	170	15.4	4.3	5.0	25.2
A (m^2)	170	0.4	0.2	0.1	1.8
σ (kN/m^2)	170	254.9	95.8	72.0	509.0
ϕ_s (°)	170	25.9	6.1	13.0	51.9
ϕ_t (°)	170	29.2	7.7	16.0	55.0
c_s (kN/m^2)	170	14.0	34.8	1.0	385.3
c_t (kN/m^2)	170	20.2	56.6	0.0	500.0
M_{vs} (kN/m^2)	170	32950.4	130197.5	5464.3	1000000.0
M_{vt} (kN/m^2)	170	42297.1	129457.5	4000.0	1000000.0
q_{cs} (MPa)	170	12.0	12.3	3.1	90.0
q_{ct} (MPa)	170	20.1	14.2	5.0	90.0
SPT_s	170	15.3	8.6	4.4	60.0
SPT_t	170	23.9	10.0	7.0	60.0
Y_1-Q (kN)	170	2553.7	2066.6	500.0	9894.0
Y_2-s (mm)	170	6.4	4.0	0.7	18.5

To justify the averaging procedure, four piles from the database are analysed in PLAXIS 2D /34/. For each single pile, two axisymmetric finite-element models are created: (i) with the real stratigraphy and (ii) with the averaged-shaft stratigraphy (Fig. 2). Both models use 15-noded triangular elements to create the mesh. Soil behaviour is represented by the Mohr-Coulomb model, with drained conditions applied in the analysis, while the pile is modelled as linear elastic. Pile-soil interaction is modelled using interface elements (R_{int} = 0.8). Calculations are carried out in three phases. The initial phase involves activating the self-weight of the soil. In the second phase, installation effects are taken into account, while in the final phase the pile load is applied.



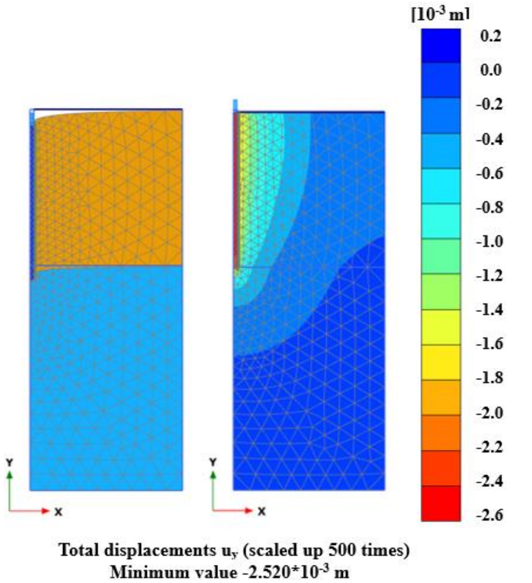


Figure 2. Results of pile settlement under axial load for Pile 1 (∅600, L = 12.3 m).

Standard boundary conditions are applied: vertical boundaries are fixed in the horizontal direction, the bottom boundary is fully fixed, and the top boundary is left free. Figure 2 shows one example of pile modelling in Plaxis 2D with real and average stratigraphy while Table 3 represents the used parameters for soil model in the first and second case.

Table 3. Geotechnical profile for Pile 1 (∅600, L = 12.3 m) and Pile 2 (∅800, L = 12.3 m). H - layer thickness, γ - soil unit weight, ϕ - friction angle, c - cohesion, M_v - compressibility modulus for presented layer.

Layer	H (m)	γ (kN/m^3)	ϕ (°)	c (kPa)	M_v (kN/m^2)
embankment	4	19	16	2	2000
clay	1.5	19	15	10	5000
silt	5	19	20	2	10000
sandy-gravel	1.8	19	30	0	18000
Avg shaft	12.3	19	19.6	2.7	7959

Under the action of load Q on the pile of length L and diameter D, the calculated settlement obtained from the model with real stratigraphy (U_{real}) and from the model with averaged stratigraphy (U_{avg}) is derived. The comparison of settlements in Table 4 is presented through the difference between the calculated settlements U_{real} and U_{avg}, expressed as ΔU, as well as through the comparison of calculated settlements (U_{real} and U_{avg}) with measured settlements U_{slt} obtained from the SLT test.

The validation study demonstrates that thickness-weighted shaft averaging introduces negligible bias in pile settlement predictions. Across four case studies re-analysed in PLAXIS 2D, the averaged-shaft models reproduce settlements within 0.01-0.36 mm (≤ 9 % relative difference) of the full stratigraphy models, indicating that global pile deformation under service-level loads is only marginally influenced by accurate layering along the shaft provided the toe layer is modelled explicitly. Comparisons with static load tests further confirm that both approaches yield errors << mm, consistent with the dispersion observed in repeated field tests. Although the percentage of difference between the calculated settlement and measured settlement for pile 3 appears large, this discrepancy is of limited practical significance due to the very small absolute magnitude of settlements involved. These results justify the use of averaged-shaft parameters in dataset preparation, as they effectively reduce dimensionality and noise without compromising the accuracy required for machine-learning applications.

Table 4. Pile settlements under axial load (Plaxis 2D).

Pile	1	2	3	4
D (mm)	600	800	800	600
L (m)	12.3	12.3	15.5	20.0
Q (kN)	850	1300	1480	1800
UReal (mm)	2.36	4.02	0.60	4.36
UAvg (mm)	2.52	4.38	0.57	4.37
ΔU (mm)/(%)	0.16/6.8	0.36/9.0	0.03/-5.0	0.01/0.2
U SLT (mm)	2.73	3.49	1.12	4.29
UReal vs. USLT (mm)/(%)	-0.37/ 13.7	+0.53/ +15.1	-0.52/ -46.4	+0.07/ +1.6
UAvg vs. USLT (mm)/(%)	-0.21/ -7.8	+0.89/ +25.4	-0.55 / -49.1	+0.08/ +1.9

ML RESULTS AND COMPARISON

Comparison of eleven ML algorithms on the pile bearing capacity and settlement estimation and the importance of influencing variables are presented in this section. Dataset is partitioned into training and testing dataset by cross validation process. To compare the performances of OML algorithms trained on the training dataset, R^2, mean squared error, and absolute squared error are generally applied to the testing dataset as a validation procedure for the executed process. Results comparing the predicted and measured values for bearing capacity and settlement are presented by R^2 metric parameter for the best-performing algorithm among the eleven optimised ML algorithms.

Results of the testing dataset

Table 5 presents the R^2 values for each ML algorithm across the four prediction approaches used to estimate pile bearing capacity, whereas Table 6 provides corresponding results for predicting pile settlement. These results refer to the test dataset, which contains 20 % of the total samples from the database.

Table 5. R^2 for bearing capacity (Y_1) prediction.

Y_1	Approach			
ML Model	*LAB*	*CPT*	*SPT*	*SUMA*
LR	0.338	0.478	0.640	0.375
RR	0.553	0.486	0.662	0.658
LsR	0.472	0.483	0.652	0.531
ER	0.554	0.492	0.662	0.666
DT	0.701	0.816	0.906	0.791
RF	0.785	0.863	0.911	0.831
SVM	0.314	0.208	0.496	0.478
KNN	0.777	0.818	0.874	0.842
GB	0.818	0.875	0.890	0.826
XGB	0.815	0.876	0.900	0.828
CB	0.834	0.869	0.921	0.875

Across all evaluated machine-learning algorithms, ensemble and boosting models and Tree-based models consistently demonstrate superior performance compared to linear and kernel-based algorithms. Models such as Gradient Boosting, XGBoost, Random Forest, and particularly CatBoost are better suited for modelling the nonlinear relationships present in geotechnical datasets. On the contrary, Support Vector Machine and linear models (LR, RR, LsR, ER) show limited prediction capability, reflecting their reduced flexibility in modelling complex soil-structure interactions.

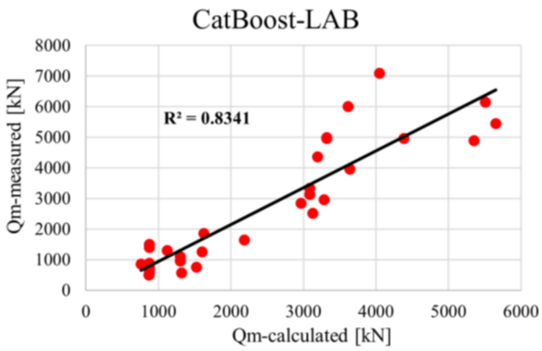
Table 6. R^2 for settlement (Y_2) prediction.

Y_2	Approach			
ML Model	LAB	CPT	SPT	SUMA
LR	0.452	0.000	0.051	0.587
RR	0.464	0.057	0.094	0.705
LsR	0.429	0.042	0.097	0.713
ER	0.453	0.056	0.105	0.709
DT	0.534	0.373	0.430	0.546
RF	0.657	0.637	0.528	0.714
SVM	0.694	0.578	0.452	0.767
KNN	0.677	0.571	0.551	0.627
GB	0.713	0.642	0.551	0.740
XGB	0.694	0.632	0.572	0.747
CB	0.751	0.691	0.570	0.761

For bearing-capacity prediction, field-based approaches (CPT and SPT) generally provide more reliable results than laboratory-only inputs, reflecting the advantage of measurements obtained in undisturbed soil conditions. CatBoost again delivers the strongest performance, with SPT achieving the highest R^2 of 0.921 and the remaining approaches showing competitive values: LAB (0.834), CPT (0.869), and SUMA (0.875). All four prediction approaches fall within a relatively narrow accuracy range of 0.83-0.92, indicating that bearing-capacity prediction is generally more robust across approaches compared to settlement. These findings clearly show that field tests remain the most reliable basis for bearing capacity evaluation.

For predicting pile settlement, CatBoost achieves the most reliable results across all prediction approaches, most notably in LAB (0.751) and SUMA (0.761). These two approaches outperform CPT and SPT (both below 0.7), primarily because they include the compressibility modulus (M_v), a key stiffness parameter directly measured in oedometer tests. SVM also performs competitively in SUMA (0.767), while Random Forest and Gradient Boosting show solid performance, particularly within LAB and SUMA subsets. Linear models, especially Linear Regression, exhibit the weakest results, as illustrated by the CPT configuration where LR yields an R^2 of 0.000, confirming that linear models cannot adequately capture settlement behaviour of geomaterials.

Figures 3 and 4 graphically show regression plots of the measured and predicted (calculated) values of bearing capacity (Q_m) and pile settlement (s), respectively.



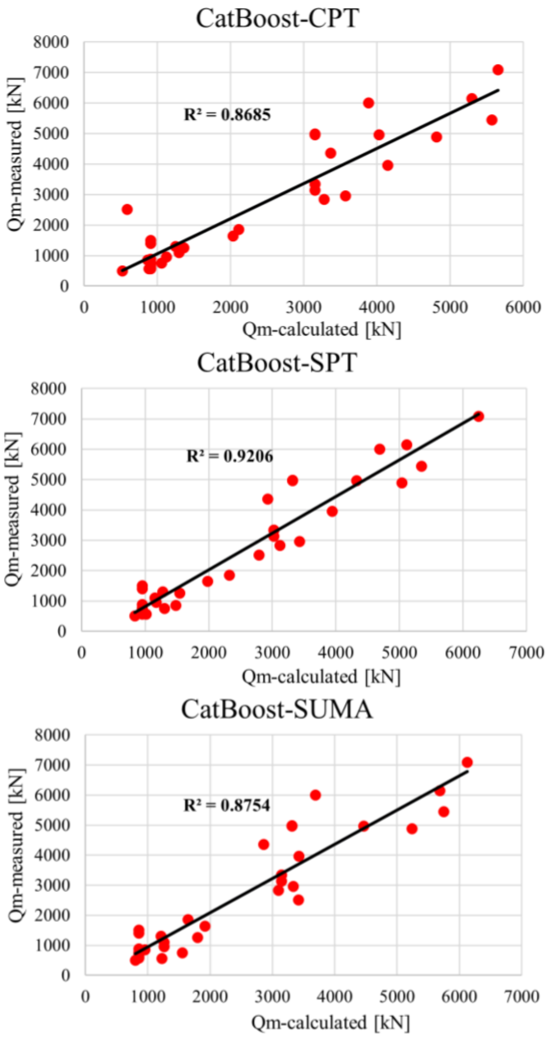
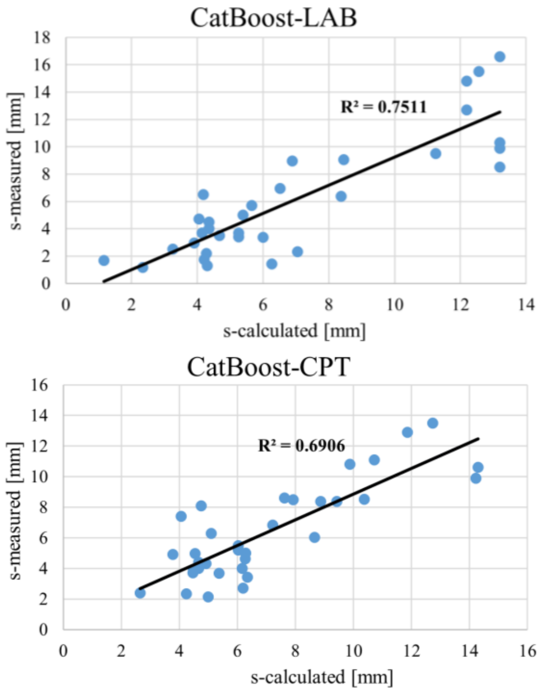


Figure 3. Regression plots for bearing capacity prediction using CatBoost algorithm.



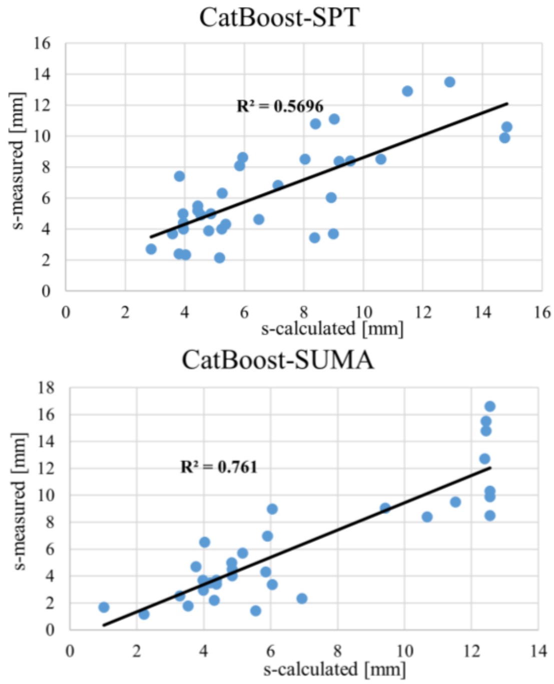


Figure 4. Regression plots for settlement prediction using CatBoost algorithm.

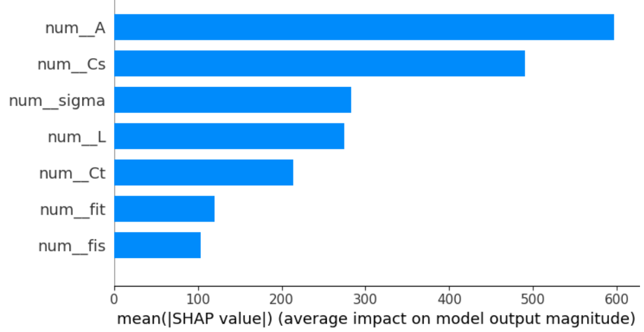
The all of four prediction approaches are presented for one random test set of CatBoost algorithm.

Variable importance analysis

In this section, the SHAP-based feature-importance results are presented to identify which input parameters govern the predicted pile bearing capacity and settlement. The bar plots summarise the contribution of each variable based on the CatBoost, selected in the previous stage as the best-performing algorithm. These plots enable comparison across prediction approaches and support the evaluation of the geotechnical relevance of identified feature rankings. Although the SHAP analysis offers a clear ranking of feature contributions, some results need to be interpreted from a geotechnical perspective. SHAP explains how the model distributes predictive importance within the dataset, not how the soil actually behaves.

Figure 5 presents the variable importance scores for bearing capacity prediction across the four prediction approaches, while Fig. 6 summarises the corresponding results for settlement prediction.

LAB approach



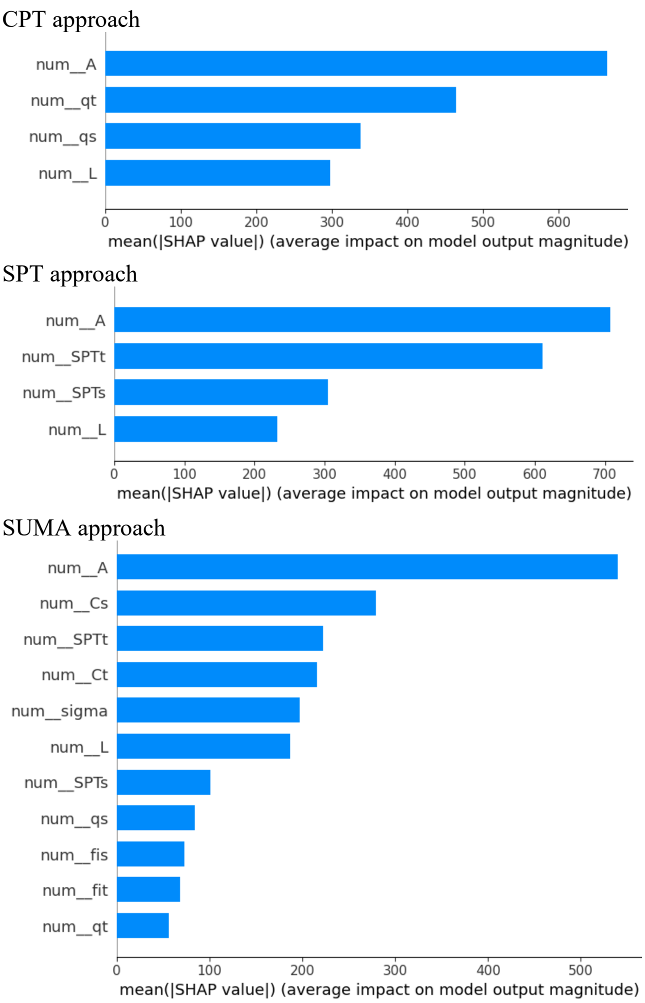
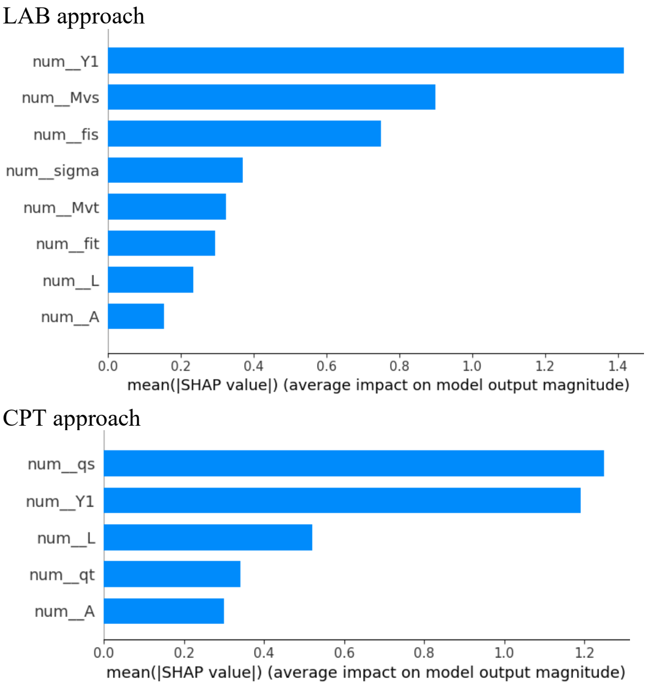


Figure 5. Results of Variable Importance for prediction of bearing capacity using CatBoost algorithm.



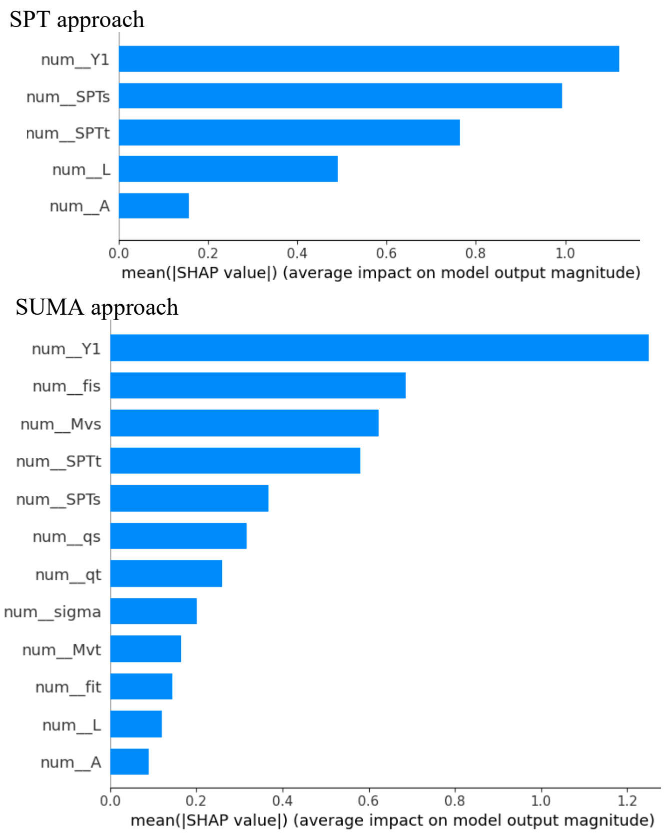


Figure 6. Results of Variable Importance for prediction of settlement using CatBoost algorithm.

For bearing capacity, the cross-sectional area is consistently identified as the dominant variable, which is expected since both tip and shaft resistance increase with pile size. For field-based approaches (CPT and SPT), the results align with geotechnical practice, with penetration resistance and geometry, contributing at similar levels to bearing-capacity prediction. For laboratory-based approaches (LAB and SUMA), cohesion (c) appears with higher importance than expected, while the friction angle (ϕ) shows a relatively low contribution, which does not fully correspond to soil mechanics understanding. One would expect a more pronounced influence of friction angle. The main reason for their reduced importance here is the way the soil data are processed. By averaging soil parameters along the pile shaft, much of the variation between layers is smoothed out. This smoothing may mask the effect of friction angle in the model, despite its practical importance. In addition, when geometric parameters explain a large part of the variation, the contribution of strength-related parameters can appear smaller than expected.

For settlement prediction, the input load (Y_1) and oedometer modulus (M_{vs}) emerge as key parameters in the LAB and SUMA approaches, which is expected because settlement is primarily governed by stiffness, and M_{vs} directly controls the compressibility of the soil matrix. The friction angle (ϕ), however, shows a higher relative importance than anticipated. Although ϕ is not part of traditional settlement formulations, its contribution reflects the indirect link between friction angle and stiffness in granular and mixed soils. Higher values of ϕ generally correspond to higher relative density, higher small-strain stiffness, and reduced compressibility under service stress levels. In the CPT and SPT approaches, the applied load again appears as the dominant factor, supported by penetration-based parameters that capture deformation characteristics relevant to settlement prediction.

It is noteworthy that the remaining variables also exhibit non-negligible importance, confirming that all included features contribute meaningfully to the prediction process and reflect parameters commonly used in engineering practice. Overall, the SHAP analysis captures several key geotechnical relationships: geometry governing bearing capacity, stiffness controlling settlement, and field-test parameters providing stronger predictive signals than laboratory-derived inputs.

COMPARISON BETWEEN ANALYTICAL METHODS AND ML BASED PREDICTIONS

In routine engineering, analytical calculations of pile bearing capacity and settlement are used for design purposes. Most often, the pile capacity is calculated in several ways, and an assessment of the capacity is made based on the obtained results. The confirmation of specific capacities is carried out by testing the piles using static or dynamic methods. The most commonly used analytical approaches include the Meyerhof method, which relies on laboratory-derived soil parameters, and LPC-CPT method, based on cone penetration test results. In practice, piles are rarely calculated based on SPT test results, and the literature recommends various approaches for determining pile capacity using SPT data (Décourt /31/; Shooshpasha et al. /26/; Dai /27/).

For the purposes of this research, the capacities of all 170 piles are determined using the Meayerhof method, LPC-CPT method, and Decourt method based on the results of SPT tests. For most of the tested piles, design calculations are carried out using the Serbian national standard that is in use before Eurocode 7. The allowable pile capacity according to the Meyerhof method is calculated using mobilised shear-strength parameters, where the internal friction angle is reduced as $\tan \phi_m = \tan \phi/1.5$ and cohesion as $c_m = c/2.5$. For the LPC-CPT method, safety factors of 2 for shaft resistance and 3 for toe resistance are applied. The resulting capacities represent the maximal values allowable under service (working) load conditions.

Eurocode 7 has since been adopted as the official standard for pile design (SRPS EN 1997-1:2017 /32/; SRPS EN 1997-1/NA:2020 /33/). It is based on the application of partial factors in determining the design capacity of piles, requiring that the design action E_d is less than design resistance R_d. Here, E_d denotes the factored actions relevant to the Ultimate Limit State. The calculation in this study is also carried out using Eurocode 7, according to the Meayerhof method and LPC-CPT method. However, since the predicted values of bearing capacity relate to pile bearing capacities close to the serviceability limit state, the design capacity values (R_d) are reduced by a factor of 1.3875, assuming 75 % of the permanent load with a factor of $\Upsilon_g = 1.35$ and 25 % of variable load with a factor of $\Upsilon_q = 1.5$. In this way, a comparison between the pile bearing capacity, according to former regulations, and the design capacity of piles, according to the actual regulations is ensured.

For all the piles in the database, the predicted settlement is calculated according to Poulos method, i.e., the axial stiffness of the pile is determined, based on the linear relationship between the load and the settlement of the pile. This procedure involves determining the influencing factors based on the geometric data of the pile, the elastic modulus of the concrete of the pile, the elastic modulus and the Poisson's ratio of the soil in which the pile is installed.

Table 7 presents the effectiveness of each analytical method using the R^2 parameter by comparing their results for each sample from the database with measurements from static and dynamic load tests. The calculated pile bearing capacities obtained using analytical methods based on the former national regulations are denoted as R and calculated pile bearing capacities according to the Eurocode 7 are denoted as R'. The values of R' are reduced by the factor 1.3875 to ensure their direct comparability with values of R, as previously explained. The labels LAB, CPT, and SPT refer to the Meyerhof method, the LPC-CPT calculation procedure, and Decourt method, respectively, while k_p denotes the calculated axial pile stiffness according to the Poulos method.

The R^2 values for pile bearing capacities are less than 0.5 which indicates a significant deviation between the measured and calculated bearing capacities. The analytical calculation according to the LPC-CPT method stands out as the more reliable method, with $R^2 = 0.455$, compared to calculation methods based on the application of soil parameters determined in the laboratory and the results of SPT. Due to the comparison of measured and calculated axial stiffness of the piles, a coefficient of $R^2 = 0.437$ is obtained.

Table 7. Representation of effectiveness of analytical methods compared with measured values.

Analytical methods	R^2
R-LAB	0.363
R-CPT	0.455
R-SPT	0.382
R'-LAB	0.259
R'-CPT	0.451
k_p	0.437

Presented results indicate a significant contribution of applying machine learning algorithms in predicting pile bearing capacity and settlement. Statistical analysis has confirmed a high level of reliability. A common practice is that the values of calculated capacities are overestimated, i.e., large safety factors are provided, which is often confirmed through pile testing at a stage when most piles have already been installed. The application of machine learning algorithms based on the known results of already tested piles can help in optimising foundation solutions and determining more realistic values for pile capacity and settlement. Additionally, the use of artificial intelligence in the field of pile foundation design can reduce the scope of site investigation and the amount of testing needed to validate design models.

CONCLUSIONS

This study presents a comprehensive comparison of 11 machine learning (ML) algorithms, including LR, RR, LsR, ER, DT, RF, SVM, KNN, GB, XGB, and CB, to identify the most effective one for estimating the bearing capacity and settlement of piles. The analysis utilises a dataset consisting of 170 cases for the ML algorithms. Four prediction approaches are formed depending on the origin of the input parameters. Performance is presented using R^2 metric parameter with cross-validation employed as the validation method. Several key conclusions can be drawn from the results.

- The most efficient algorithm identified is CatBoost. Algorithms such as KNN, GB, XGB, RF, and DT also demonstrate a high level of efficiency. Linear regression models are unsuitable because the problem exhibits a high degree of nonlinearity.
- The best prediction approach for predicting bearing capacity is SPT ($R^2 = 0.921$), followed by CPT ($R^2 = 0.869$), then SUMA ($R^2 = 0.875$), and LAB ($R^2 = 0.834$). For settlement predictions, the most efficient prediction approaches are SUMA ($R^2 = 0.761$ for CatBoost and $R^2 = 0.767$ for SVM algorithm) and LAB ($R^2 = 0.751$).
- Results of variable importance analysis indicate that the cross-sectional area of the pile is the most influential variable for predicting pile bearing capacity. For estimating pile settlement, the most important variables are measured load Y_1 (LAB, SUMA and SPT) and q_s (CPT).
- A comparison of analytical calculation methods and measured values for pile bearing capacity and settlement demonstrates a clear advantage in applying machine learning algorithms due to their significantly higher prediction accuracy compared to traditional empirical calculation methods.
- Main limitations of this study are the approximation of stratigraphy through parameter averaging and the relatively small database. Expanding the dataset would further improve the reliability and applicability of the proposed ML algorithms.
- Machine learning offers significant potential for predicting pile behaviour prior to field testing and for optimising testing procedures. Challenges in assessing pile bearing capacity across varying soil conditions remain central to geotechnical engineering, with design solutions from the Geotechnical Report closely tied to field load testing and structural design requirements. In this regard, ML provides an effective framework by integrating insights from previous static (SLT) and dynamic (DLT) load tests, strengthened by the computational robustness of advanced mathematical models.

ACKNOWLEDGEMENT

This paper is financially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia within the Project 451-03-136/2025-03/200012. This support is gratefully acknowledged.

REFERENCES

1. Souza, T. de J., Querelli, A., Cruz, F.V.A. de S., Trejo-Noreña, P.C. (2021), *Use of dynamic load test to obtain the pile capacity - the Brazilian experience*, DYNA, 88(217): 169-177. doi: 10.15446/dyna.v88n217.93416
2. Cuajao, V.C.S., Dela Cruz, O.G., Tabaroei, A. (2024), *Comparison of static and dynamic load testing: A review*, GEOMATE J, 27(121): 119-127. https://geomatejournal.com/geomate/article/view/4726
3. Das, S.K., Basudhar, P.K. (2006), *Undrained lateral load capacity of piles in clay using artificial neural network*, Comput. Geotech. 33(8): 454-459. doi: 10.1016/j.compgeo.2006.08.006
4. Lee, I.-M., Lee, J.-H. (1996), *Prediction of pile bearing capacity using artificial neural networks*, Comput. Geotech. 18(3): 189-200. doi: 10.1016/0266-352X(95)00027-8
5. Goh, A.T.C. (1995), *Empirical design in geotechnics using neural networks*, Géotechnique, 45(4): 709-714. doi: 10.1680/geot.1995.45.4.709
6. 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*, Geotech. Geol. Eng. 36: 2057-2076. doi: 10.1007/s10706-018-0445-7
7. Armaghani, D.J., Asteris, P.G., Fatemi, S.A., et al. (2020), *On the use of neuro-swarm system to forecast the pile settlement*, Appl. Sci. 10(6): 1904. doi: 10.3390/app10061904
8. Pal, M., Deswal, S. (2008), *Modeling pile capacity using support vector machines and generalized regression neural network*, J Geotech. Geoenv. Eng. 134(7): 1021-1024. doi: 10.1061/(ASCE)1090-0241(2008)134:7(1021)
9. Samui, P. (2011), *Prediction of pile bearing capacity using support vector machine*, Int. J Geotech. Eng. 5(1): 95-102. doi: 10.3328/IJGE.2011.05.01.95-102
10. Kardani, N., Zhou, A., Nazem, M., Shen, S.-L. (2020), *Estimation of bearing capacity of piles in cohesionless soil using optimised machine learning approaches*, Geotech. Geol. Eng. 38: 2271-2291. doi: 10.1007/s10706-019-01085-8
11. Pham, T.A., Tran, V.Q., Vu, H.-L.T., Ly, H.-B. (2020), *Design deep neural network architecture using a genetic algorithm for estimation of pile bearing capacity*, PLoS ONE, 15(12): e0242537. doi: 10.1371/journal.pone.0243030
12. Benbouras, M.A., Petrișor, A.-I., Zedira, H., et al. (2021), *Forecasting the bearing capacity of the driven piles using advanced machine-learning techniques*, Appl. Sci. 11(22): 10908. doi: 10.3390/app112210908
13. Amjad, M., Ahmad, I., Ahmad, M., et al. (2022), *Prediction of pile bearing capacity using XGBoost algorithm: Modeling and performance evaluation*, Appl. Sci. 12(4): 2126. doi: 10.3390/app12042126
14. Arbi, S.J., ur Rehman, Z., Hassan, W., et al. (2025), *Optimized machine learning-based enhanced modeling of pile bearing capacity in layered soils using random and grid search techniques*, Earth Sci. Inform. 18: 332. doi: 10.1007/s12145-025-01784-2
15. Kumar, M., Kumar, D.R., Khatti, J., et al. (2024), *Prediction of bearing capacity of pile foundation using deep learning approaches*, Front. Struct. Civ. Eng. 18(4): 870-886. doi: 10.1007/s11709-024-1085-z
16. Al-Atroush, M.E. (2025), *A deep learning physics-informed neural network (PINN) for predicting drilled shaft axial capacity*, Appl. Comput. Geosci. 26: 100246. doi: 10.1016/j.acags.2025.100246
17. Khan, A., Khan, M., Khan, W.A., et al. (2025), *Predicting pile bearing capacity using gene expression programming with SHapley Additive exPlanation interpretation*, Discov. Civ. Eng. 2: 58. doi: 10.1007/s44290-025-00215-x
18. Khatti, J., Khanmohammadi, M., Fissha, Y. (2024), *Prediction of time-dependent bearing capacity of concrete pile in cohesive soil using optimized relevance vector machine and long short-term memory models*, Sci. Rep. 14: 32047. doi: 10.1038/s41598-024-83784-8
19. Igoe, D., Zahedi P., Soltani-Jigheh, H. (2024), *Predicting the compression capacity of screw piles in sand using machine learning trained on finite element analysis*, Geotech. 4(3): 807-823. doi: 10.3390/geotechnics4030042

20. Seo, S., Kim, G., Park, J.-B., et al. (2025), *Leveraging data-driven machine learning techniques to enhance bearing capacity estimation in prebored and precast piles*, Expert Syst. Appl. 285: 128070. doi: 10.1016/j.eswa.2025.128070
21. Palalane, L., Dithinde, M. (2024), *The use of machine learning approach to predict pile capacity in non-cohesive soils*, J Civ. Eng. Urb. 14(3S): 311-317. doi: 10.54203/jceu.2024.34
22. Thottoth, S.R., Das, P.P., Khatri, V.N. (2024), *Prediction of compression capacity of under-reamed piles in sand and clay*, Multisc. Multidisc. Model. Exp. Des. 7: 2289-2305. doi: 10.1007/s41939-023-00331-0
23. Shuman, N.M., Khan, M.S., Amini, F. (2024), *Efficient machine learning model for settlement prediction of large diameter helical pile in c-ϕ soil*, AI Civ. Eng. 3: 10. doi: 10.1007/s43503-024-00028-4
24. Cao, Y., Ni, J., Chen, J., Geng, Y. (2025), *Rapid evaluation method to vertical bearing capacity of pile group foundation based on machine learning*, Sensors, 25(4): 1214. doi: 10.3390/s25041214
25. Onyelowe, K.C., Hanandeh, S., Kamchoom, V., et al. (2025), *Developing advanced datadriven framework to predict the bearing capacity of piles on rock*, Sci. Rep. 15: 11051. doi: 10.1038/s41598-025-96186-1
26. Shooshpasha, I., Hasanzadeh, A., Ghalesari, A.T. (2013), *Prediction of the axial bearing capacity of piles by SPT-based and numerical design methods*, Int. J GEOMATE, 4(2): 560-564. doi: 10.21660/2013.8.2118
27. Dai, N.V. (2014), *Determination of bearing capacity of bored pile using SPT number N and undrained shear strength Su*, VNUHCM J Sci. Technol. Devel. 17(3): 86-93. doi: 10.32508/stdj.v17i3.1479
28. Likins, G., Rausche, F. (2004), *Correlation of CAPWAP with static load tests*, In: Proc. of the 7th Int. Conf. on the Application of Stresswave Theory to Piles, Selangor, Malaysia, 2004, pp.153-165.
29. Galib, M.M.H., Billah, M.M., Mofiz, S.A. (2025), *Prediction of bearing capacity of pile using support vector regression and CatBoost regression*, In: AIP Conf. Proc. 3262(1): 020022. doi: 10.1063/5.0247333
30. Xiong, J., Pei, T., Qiu, T. (2023), *A machine learning-based method with integrated physics knowledge for predicting bearing capacity of pile foundations*, Chapter in: E. Rathje, B.M. Montoya, M.H. Wayne (Eds.): Foundations, Retaining Structures, and Geosynthetics, Geo-Congress, Los Angeles, California, 2023, Geo-Institute of ASCE, pp.175-184. doi: 10.1061/9780784484685.018
31. Décourt, L. (1982), *Prediction of the bearing capacity of piles based exclusively on N values of the SPT*, Chapter in: Penetration Testing (volume 1), 1st Ed., Routledge, Taylor & Francis Group, 1982, pp.29-34. doi: 10.1201/9780203743959-4
32. SRPS EN 1997-1:2017, Eurocode 7: Geotechnical design - Part 1: General rules (Institute for Standardization of Serbia).
33. SRPS EN 1997-1/NA:2020, Eurocode 7: Geotechnical design - Part 1: General rules - National Annex (Institute for Standardization of Serbia).
34. PLAXIS 2D license of Institute IMS, type Perpetual, license number 30000000427638 https://www.bentley.com/software/plaxis-2d

© 2026 The Author. Structural Integrity and Life, Published by DIVK (The Society for Structural Integrity and Life 'Prof. Dr Stojan Sedmak') (http://divk.inovacionicentar.rs/ivk/home.html). This is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License