

IJCIET

INTERNATIONAL JOURNAL OF CIVIL ENGINEERING AND TECHNOLOGY



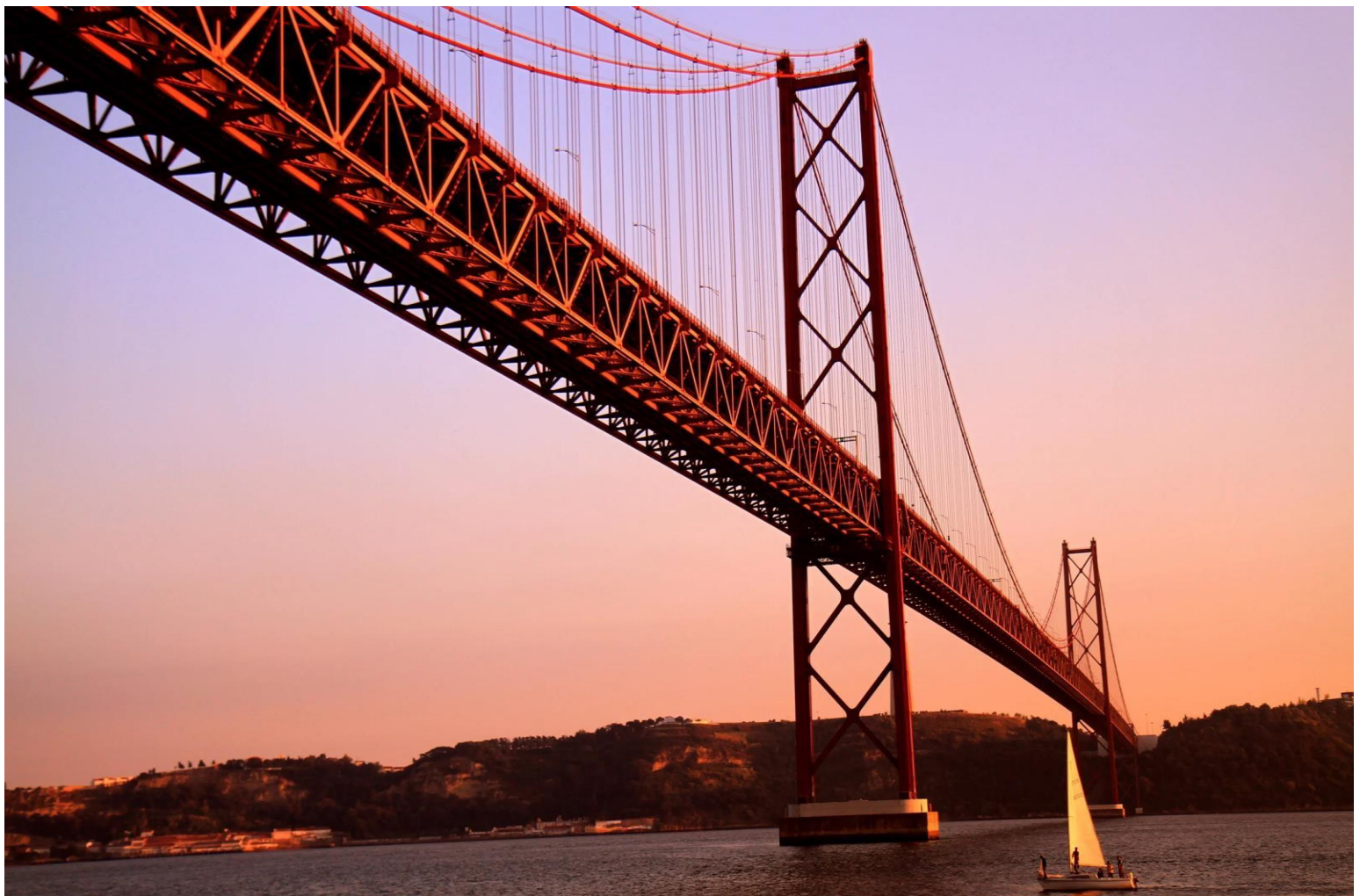
Journal ID: 6971-8185

 Google Scholar

 ResearchGate

ACADEMIA

 Crossref doi



IAEME Publication

Chennai, India

editor@iaeme.com/ iaemedu@gmail.com

<https://iaeme.com/Home/journal/IJCIET>





FEATURE TOKENIZER TRANSFORMER BASED ADVANCED DEEP LEARNING METHOD FOR PREDICTION OF THE BEARING CAPACITY OF BRANCHED PILES ON PEAT

Dinesh Kumar Jamatia*

Ph.D. scholar, Department of Civil Engineering,
National Institute of Technology, Silchar, Assam, India

Ashim Kanti Dey

Professor, Department of Civil Engineering,
National Institute of Technology, Silchar, Assam, India

***Corresponding Author: Dinesh Kumar Jamatia**

ABSTRACT

Peat is regarded as one of the most complex category of soil, in terms of construction, owing to its very soft and highly compressible nature. Thus, in order to overcome shortcomings of peat, deep foundations are most commonly adopted to transfer load through soft deposits to harder strata. However, such techniques may be uneconomical, particularly for low to medium-rise structures, pavements, embankments, and other lightly loaded critical infrastructures. This shortcoming is of particular concern especially in peat deposits, where the highly compressible peat can induce substantial deformation even under relatively light structural loads. Hence, there is a need to develop alternative ground improvement techniques that are cost-effective, practically feasible, and capable of mitigating excessive deformation associated with peat while addressing limitations of

conventional stabilization methods. In this regard, the present study evaluates the bearing capacity of peat reinforced with branched piles. Large-scale experimental studies were conducted to investigate the effectiveness of the proposed technique under controlled loading conditions, while machine learning (ML) techniques were simultaneously incorporated to develop predictive models for estimating bearing capacity and to assess the influence of key soil and pile parameters on the overall foundation response. To predict the bearing capacity of branched piles (BPs) and fine-tune the prediction outcomes, four ML models, namely the Feature Tokenizer Transformer (FTT), Bayesian Regularized Feed Forward Neural Network (BRFFNN), Generalized Regression Neural Network (GRNN), and Multi-Layer Perceptron Regressor (MLPR), were employed. The predictive performance of the four models was comparatively evaluated to identify the most accurate and reliable approach for estimating bearing capacity of branched piles (BPs). Model training was performed using a dataset comprising of 570 experimental observations, each representing five key bearing-related attributes of BPs subjected to static vertical loading. Among the four models, FTT demonstrated the highest predictive accuracy, outperforming the other approaches, as evidenced by the statistical evaluation with a coefficient of determination (R^2) value of 0.9940, approaching unity. The high predictive efficiency of the FTT model was further substantiated through rank and sensitivity analyses, which provided insights into the relative importance and influence of the input parameters on BP bearing capacity.

Keywords: Peat, Compressibility, Machine learning, Feature Tokenizer Transformer.

Cite this Article: Dinesh Kumar Jamatia and Ashim Kanti Dey. (2026). Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat. *International Journal of Civil Engineering and Technology (IJCET)*, 17(5), 1-42. DOI: https://doi.org/10.34218/IJCET_17_05_001

1. INTRODUCTION

Soils with high organic content ($>27\%$) are classified as peat [1, 2] which tends to pose significant challenges from a constructional point of view owing to their high water content, very low bulk density, very low shear strength, and high compressibility.

Such problematic soils have resulted in damage to major lifeline infrastructures not only during their construction phase but also during their service life [3-6]. The presence of a highly compressible organic matrix and weakly bonded soil structure makes these deposits particularly susceptible to large and prolonged settlements even under lightly loaded structural and embankment loads. Moreover, the heterogeneous and fibrous nature of organic soils can result in considerable spatial variability in engineering properties, making conventional site characterization and prediction of ground response challenging. The excessive settlement of infrastructure constructed over such deposits may lead to differential deformation, distortion of pavement surfaces, loss of serviceability, and distress in structural and utility components [3, 7, 8]. In addition, the low bearing capacity and high deformation potential of these soils necessitate careful evaluation of their geotechnical characteristics prior to construction, along with the adoption of appropriate ground-mitigation measures to enhance bearing capacity and control excessive settlement [9, 10]. Conventional ground-improvement techniques, such as soil replacement, excavation and replacement, and grouting, may not be practically feasible at majority of sites, particularly where thick organic deposits, difficult site access, high groundwater levels, or extensive treatment areas are encountered conditions that are commonly associated with peatland environments [1, 3, 4, 6, 11]. Further, the above mentioned techniques also involve substantial material, equipment, labour, and construction time, thereby increasing the overall project cost and potentially causing difficulties in meeting construction schedules. Hence, there is a need to develop a cost-effective, rapid, and technically adequate ground-improvement approach that can effectively enhance the load-bearing performance of peat while minimizing settlement. Such a technique should not only provide measurable improvement in the engineering properties of the treated ground but also be simple to implement, adaptable to varying site conditions, and capable of being executed within the stipulated construction time frame.

In order to address the challenges stated in the above section, related to peatlands, researchers in the past worldwide had adopted numerous approaches in the form of X-section piles, squeezed branch (or) under-reamed piles, belled (or) pedestal piles. The suggested methodologies by researchers tends to modify the pile geometry, enlarge base (or) intermediate portion, tends to change the embedment conditions, provisions of added bulbs in the weak strata (under reamed piles), are mainly applicable for deep strata [12, 13, 14, 15].

These techniques may be theoretically adequate and capable of enhancing the bearing capacity of peat; however, their practical applicability under actual site conditions remains highly questionable. The performance of ground-improvement measures is strongly dependent on the quality of field execution, construction control, and post-treatment verification. Inadequate quality control during mitigation and inadequacy in construction especially at a higher depth ($>8\text{m}$), can result in non-uniform treatment, incomplete ground improvement, and significant deviations between the designed and achieved engineering properties. Several documented case histories have demonstrated that deficiencies in construction quality, inadequate monitoring, and insufficient verification of the treated ground can compromise the effectiveness of otherwise well-established ground-improvement techniques [16, 17]. In particular, field performance of above mentioned piles, can deviate from design expectations when installation parameters, grout properties, treatment geometry, and achieved strength are not adequately controlled and verified during construction [16]. Similarly, the need for systematic monitoring and quality assurance in compaction grouting has been highlighted owing to the difficulty of maintaining consistent treatment quality when injection parameters are controlled manually [18]. Therefore, for peatland environments characterized by high spatial variability, difficult access, and high groundwater conditions, the reliability of an improvement technique should be assessed not only based on its theoretical capacity for strength enhancement but also on its robustness to field execution and quality-control constraints. This highlights the need for a cost-effective and rapidly deployable technique that can provide consistent improvement with relatively simple construction procedures, minimal dependence on specialized equipment, and effective quality control within the stipulated construction timeframe.

Thus, to overcome the limitations associated with specially designed piles, the present study proposes a specially configured pile system that can be installed at relatively shallow depths, referred to as a root pile. Root piles have demonstrated enhanced bearing capacity at shallower embedment depths compared with conventional piles [14, 19, 20]. Their load-transfer mechanism involves contributions from both shaft and end-bearing resistance, while their distinctive geometry enhances pile–soil interaction and improves resistance to compressive, uplift, and lateral loading. The performance of root piles is further governed by several geometric parameters, including root length, number of roots, and embedment depth [12, 21, 22, 23].

Despite these advancements, identifying an optimum root-pile configuration remains challenging because its load-carrying behaviour is governed by the complex and nonlinear interactions among multiple geometric and soil parameters. Conventional analytical and empirical approaches may not adequately capture these intricate relationships, particularly when the influence of individual parameters varies in combination with other factors. In this context, artificial intelligence (AI) and machine-learning (ML) techniques provide an effective alternative for modelling complex nonlinear relationships without requiring any predefined functional relationship between input and output variables [24, 25]. Accordingly, artificial neural network (ANN)-based models have been widely employed for predicting pile capacity, settlement, and load-displacement behaviour [26], following their early application to pile-capacity prediction [27]. Among the various ANN architectures, Bayesian Regularized Feedforward Neural Networks (BRFFNNs) improve model generalization and mitigate overfitting through regularization [28, 29], whereas Generalized Regression Neural Networks (GRNNs) provide rapid nonlinear mapping and are particularly suitable for relatively small datasets [30]. Therefore, the application of AI/ML-based approaches provides a promising framework for evaluating the combined influence of root-pile geometry and soil characteristics and for identifying configurations that can maximize pile performance at relatively shallow embedment depth.

Recent advances in machine learning (ML) have demonstrated its growing potential for predicting pile capacity and identifying the governing parameters influencing pile performance [31, 32, 33, 34]. In parallel, symbolic-regression and evolutionary approaches have enabled the derivation of explicit mathematical correlations from experimental databases, thereby enhancing the interpretability and practical applicability of ML-based predictions [35]. Alongside these developments, transformer-based architectures have recently emerged as promising approaches for structured tabular data, particularly because of their ability to capture complex interactions among input features through self-attention mechanisms. FT-Transformer employs self-attention to model feature interactions and has demonstrated strong performance in tabular regression tasks [36]. These developments provide a basis for evaluating applicability of FT-Transformer alongside established approaches, BRFFNN, GRNN, and MLPR, conventional neural-network models, as well as symbolic-regression techniques, for predicting the bearing capacity of modified pile systems. Another vital aspect of predictive modelling is establishment of dimensionless parameters reflecting pile geometry and soil behaviour.

Dimensionless analysis decreases scale dependency and increases applicability of predictive model for different pile geometry and soil conditions [37]. In this regard, dimensionless input parameters are considered in the form of: number of branches (n), branch inclination ratio (β/α), branch length to frustum diameter ratio (L_b/D_f) and elastic settlement to frustum diameter ratio (s_e/D_f). As far as dependent variable is concerned, bearing capacity is concerned, it is defined as the ratio of load carrying capacity to the undrained cohesion (q/c_u), parameters chosen, as they represent geometrical properties and deformation behavior controlling behavior of BP systems in peat.

The primary objective of this study is to develop reliable and robust ML models for predicting the bearing capacity of branched pile foundations in peat. A comprehensive database of 570 datasets compiled from the literature was used for model development, training, and testing [38]. Four models, namely BRFFNN, GRNN, MLPR, and FT-Transformer, were implemented in Python using Google Colab and evaluated using R , R^2 , MSE, RMSE, MAE, MAPE, VARE, MEDAE, and SE. Symbolic regression was also employed to derive a simplified empirical equation for predicting normalized bearing response, enhancing practical applicability and interpretability of the ML predictions. The novelty lies in the systematic comparison of conventional and advanced ML models, particularly FT-Transformer, for BP capacity prediction in peat.

2. DATA SET AND FEATURES

The present study utilizes a dataset of 570 data points derived from 38 static load tests carried out in the source literature [38] as summarized in Table A1 in the Appendix. A schematic representation of the branched pile (BP) is presented in Fig. 1, while the details of the experimental studies are summarized in Table 1. The dataset was extracted from the load–settlement curves presented graphically in the source study (Fig. 2), following a methodology similar to that adopted in the studies of previous researchers [39-42]. Each marked point on the load–settlement curve was considered an individual data point for predicting bearing response of the BP. The dataset comprises five non-dimensional variables, including four input variables (n , α , (L_b/D_f) , and (s_e/D_f)) and one output variable (q/c_u) as tabulated in Table 2. The branch length (L_b) and settlement (s_e) were normalized by diameter of frustum cone (D_f) to obtain dimensionless parameters.

Loading intensity (q) at each settlement level was calculated by dividing applied load (Q) by cross-sectional area of top frustum. Undrained shear strength (c_u) was obtained from consolidated undrained triaxial tests. These dimensionless parameters were subsequently used as inputs and output for the ML models.

2.1. Correlation Matrix Analysis

Fig. 3 showcases Pearson correlation matrix of four input variables (x_1 to x_4) and output variable (y), providing a clear visualization of their linear associations [43]. The analysis reveals that (x_4) has strongest positive correlation with (y), coefficient of 0.83, indicating a strong linear association between (x_4) and the target response, suggesting that x_4 is a key predictor of bearing capacity and may therefore contribute substantially to the predictive performance of the developed ML models.

In contrast, x_1 , x_2 , and x_3 exhibit very weak correlations with the output variable, with coefficients of 0.03, -0.01 , and 0.02, respectively, indicating negligible individual linear associations with response. Inter-correlations among the input variables are also generally weak, ranging from -0.12 to 0.20. The strongest positive correlation occurs between x_1 and x_3 ($R = 0.20$), while the strongest negative correlation is observed between x_3 and x_4 ($R = -0.12$). Near-zero correlations among most predictor pairs indicate a high degree of independence and, consequently, low multicollinearity within the dataset. Overall, correlation matrix confirms that x_4 has most pronounced linear association with target variable, whereas remaining predictors exhibit minimal individual linear relationships.

2.2. Multicollinearity Analysis

Table 3 clearly showcases variance inflation factor (VIF) values for all predictor variables tends to range from 1.0246 to 1.1324, that are substantially below general threshold value of 5.0, even less than stricter value of 2.00. Thus, the results tends to showcase that multicollinearity among variables are negligible, which is an indicative that each predictor contributes distinct information to the predictive model.

Among all predictors, x_4 exhibited the largest standardized effect in the regression model, consistent with the strong positive correlation observed in the correlation heatmap. These findings demonstrate that the predictor variables are statistically independent and suitable for subsequent machine learning modelling without requiring variable elimination [44-47].

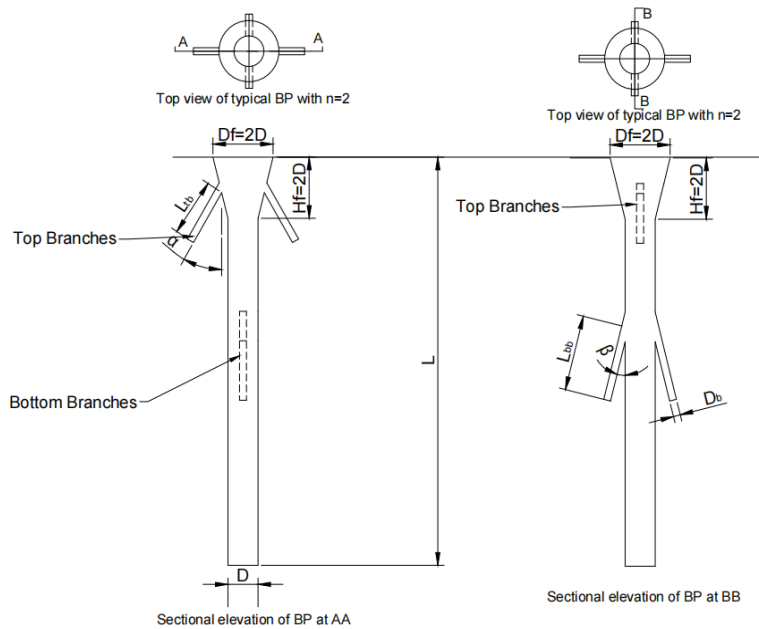
Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

Table 1 Summary of Experimental programme.

Sl. No.	L _b	α	n	Sl. No.	L _b	α	n	Sl. No.	L _b	α	n
1	1.5D	14°	2	13	1.5D	14°	3	25	1.5D	14°	4
2	2D			14	2D			26	2D		
3	2.5D			15	2.5D			27	2.5D		
4	1.5D	30°		16	1.5D	30°		28	1.5D	30°	
5	2D			17	2D			29	2D		
6	2.5D			18	2.5D			30	2.5D		
7	1.5D	45°		19	1.5D	45°		31	1.5D	45°	
8	2D			20	2D			32	2D		
9	2.5D			21	2.5D			33	2.5D		
10	1.5D	60°		22	1.5D	60°		34	1.5D	60°	
11	2D			23	2D			35	2D		
12	2.5D			24	2.5D			36	2.5D		
								37	2.5D	45°	6
								38	2.5D	45°	8

Table 2 Selected features and denoted code

SL. no.	Feature name	Code
1	Number of branches (n)	X ₁
2	The ratio of the Inclination angle of the bottom whorl branches to the top whorl branches (β/α)	X ₂
3	The ratio of the Length of the branches to the Diameter of top frustum cone (L_b/D_f)	X ₃
4	The ratio of the Settlement of Branched pile to the Diameter of top frustum cone (s_e/D_f)	X ₄
5	The ratio of the Load intensity on Branched pile to the undrained cohesion (q/c_u)	X ₅



Where,

D = Diameter of the main pile trunk.

D_f = Diameter of the top frustum equal to $2D$.

H_f = Height of the frustum equal to $2D$.

L_b = Length of the branches, where L_{tb} and L_{bb} are length of the top and bottom branches respectively.

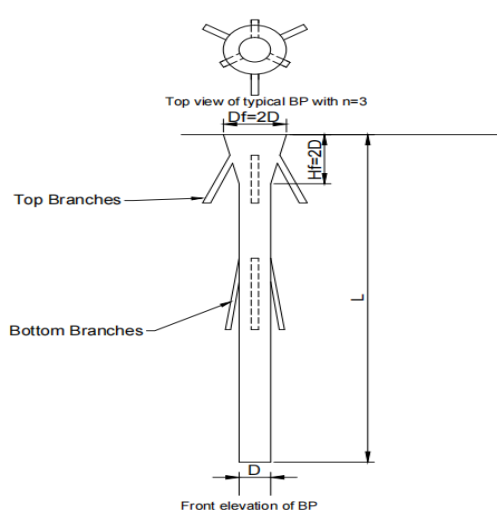
L = Total length of the pile.

α = Angle of inclination of the top branches with vertical.

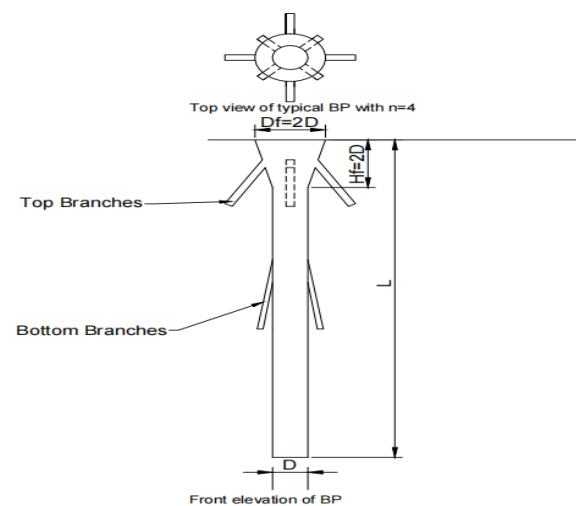
β = Angle of inclination of the bottom branches with vertical.

n = number of branches (eg.: $n=2$ means top branches on the top whorl and 2 branches on the bottom whorl, $n=3$ means branches on the top whorl and 3 branches on the bottom whorl and so on).

(a)



(b)



(c)

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

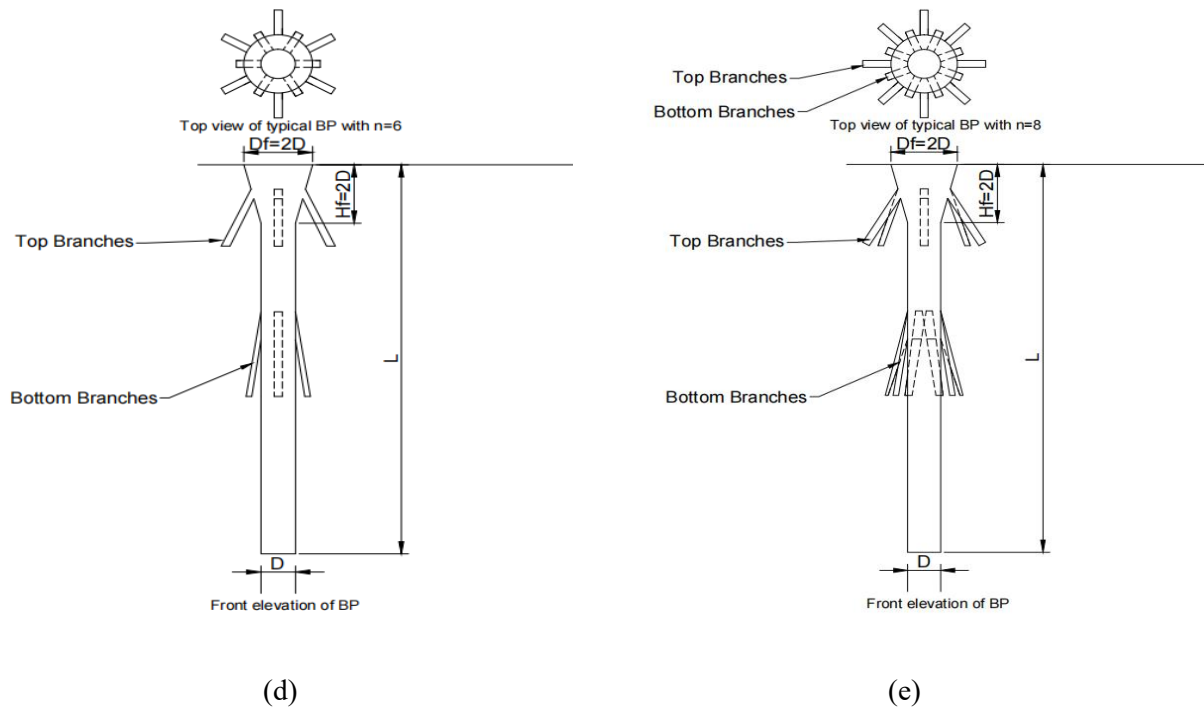


Fig. 1 Schematic diagram of the model branched pile (BP): (a) BP with $n=2$; (b) BP with $n=3$; (c) BP with $n=4$; (d) BP with $n=6$; (e) BP with $n=8$.

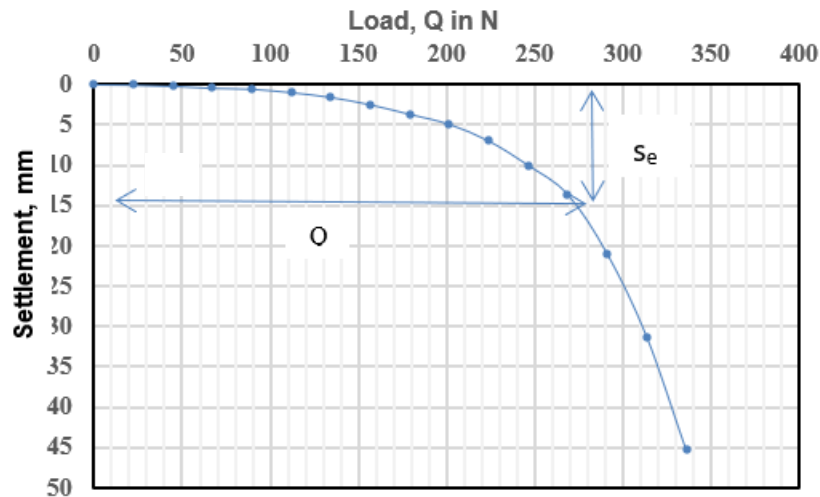


Fig. 2 A typical load-settlement plot for BP with $n=4$.

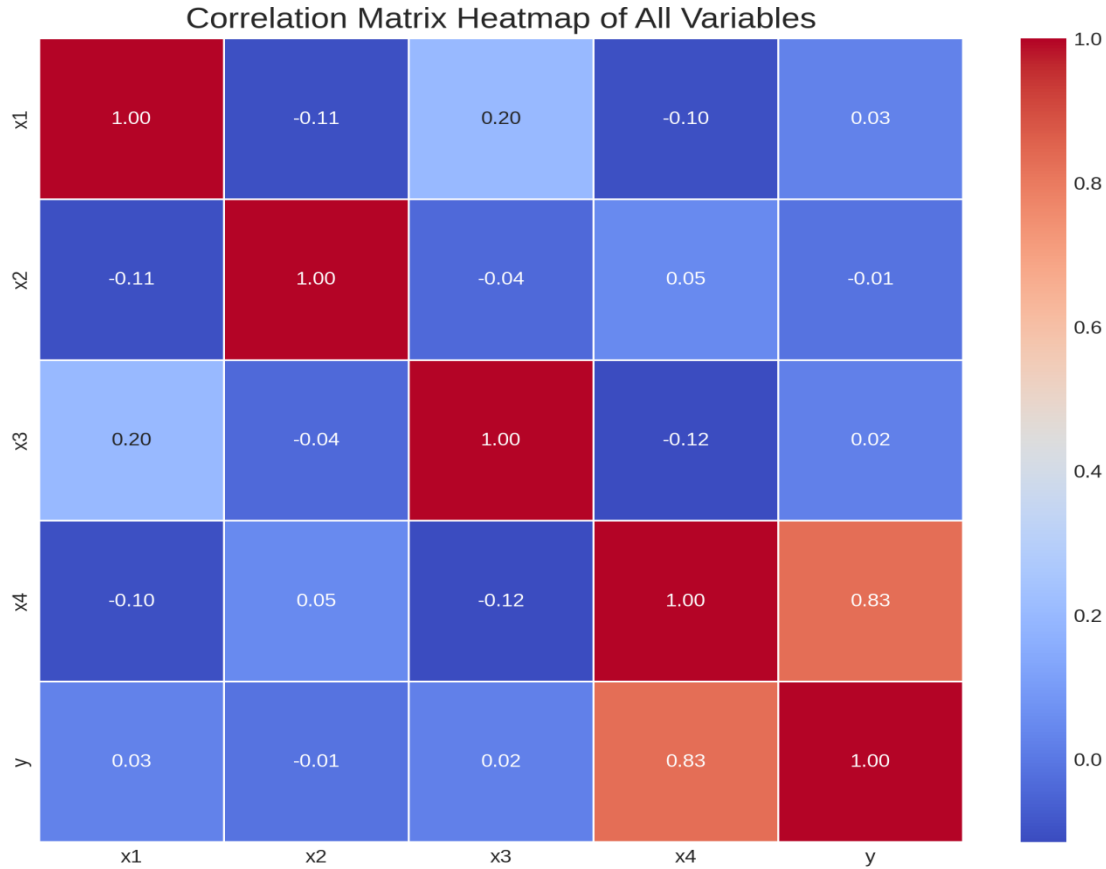


Fig. 3 Pearson correlation matrix heatmap of the variables.

2.3. Sensitivity analysis using Cosine Amplitude Method (CAM)

Focusing on the identification of the key input variables that influences the prediction of the bearing capacity of the branched pile, Cosine Amplitude Method (CAM), a nonlinear approach was applied [41, 48]. The analysis considers x_1 , x_2 , x_3 and x_4 as the input variables for estimating the bearing capacity of the branched pile in peat. The CAM equation is expressed as :

$$X(u, v) = \frac{\sum_{j=1}^m u_j v_j}{\sqrt{\sum_{j=1}^m u_j^2} * \sqrt{\sum_{j=1}^m v_j^2}} \quad (1)$$

$X(u, v)$ = cosine similarity (CAM value) between vectors; u_j and v_j are the components of vectors u and v , respectively; m is the number of elements in the vectors.

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

A CAM value close to 1 (one) points out that the specific input variable is highly relevant in the prediction, whereas a CAM value near to 0 (zero) suggests that the variable has minimum importance.

Figure 4 showcases the sensitivity of the input variables for the prediction of the bearing capacity of branched pile.

Table 3 Multicollinearity levels of the variables

Input features	Coefficients	Std. error	T-stat	P-value	VIF	Levels
Intercepts	4.4837	1.1720	3.8256	0.0001	—	—
x1	0.5922	0.1736	3.4105	0.0007	1.1324	Low
x2	-1.1964	0.7069	-1.6925	0.0912	1.0275	Low
x3	4.3933	1.0560	4.1601	<0.0001	1.1112	Low
x4	34.2124	0.9936	34.4321	<0.0001	1.0246	Low

The CAM sensitivity analysis indicates that all four input variables have a very strong influence on the predicted the bearing capacity, as evidenced by the high CAM values ranging from 0.7559 to 0.8694. However, their relative importance differs. The parameter x_3 (L_b/d_f) exhibits the highest sensitivity value of 0.8694, indicating that the ratio of branch length to frustum diameter is the most influential variable affecting the model output. This implies that the geometrical adjustment of the branch length contributes dominantly to the increase of the load-carrying capacity of the branched pile system. The second most important parameter is x_4 (s_c/d_f) with a CAM value of 0.8502, which emphasizes the strong impact of settlement-related factors on the prediction of bearing capacity. This emphasizes the role of pile-soil deformation features in the load transfer mechanism. Also the parameter x_1 (n) has a large influence with the CAM value of 0.8280. It suggests that increasing the number of branches adds greatly to bearing capacity development by enlarging the effective stress-transfer zone and mobilizing greater soil resistance. The lowest CAM value among all variables is x_2 (β/α) with 0.7559. The branch geometry in terms of length and settlement characteristics has a greater control over the pile performance than the branch inclination, which is still very significant but its impact is relatively smaller than the other parameters.

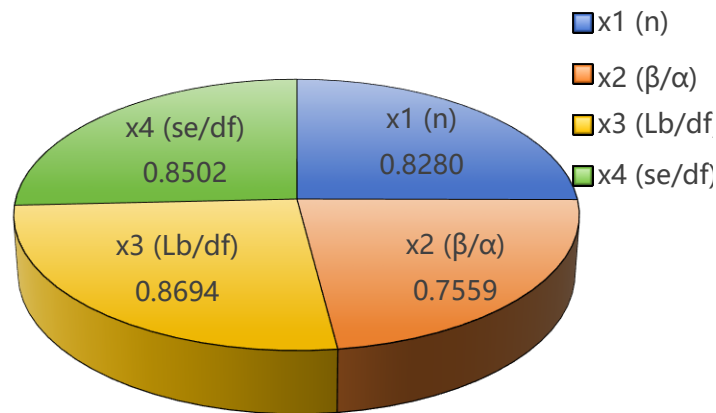


Fig. 4 Graph of the sensitivity analysis

3. BRIEF DESCRIPTION OF DIFFERENT MODELS USED

Artificial intelligence (AI) and machine learning (ML) approaches have been developed as powerful tools to solve complicated nonlinear problems in geotechnical engineering. Traditional empirical and analytical methods frequently have difficulties to deal with highly nonlinear interaction of soil, pile shape and loading circumstances. Machine learning models are an efficient alternative that immediately learn the hidden relations from experimental and field data without the need for explicit assumptions about the underlying physical mechanisms. Thus, ML-based techniques are widely used to estimate the bearing capacity, settlement and performance of foundation systems [25, 49].

In the present study, four machine learning models, namely the Bayesian Regularized Feedforward Neural Network (BRFFNN), Generalized Regression Neural Network (GRNN), Multi-Layer Perceptron (MLP) Regressor, and Feature Tokenizer Transformer (FT-Transformer), were employed to predict the bearing capacity of branched piles in peat soil. The models were chosen to illustrate successive generations of data-driven learning algorithms, from classical artificial neural networks to advanced transformer-based deep learning architectures. A comparison of the models provides useful information on their predictive capability, robustness, and applicability for foundation engineering applications.

3.1. Bayesian Regularized Feedforward Neural Network

The Bayesian Regularized Feedforward Neural Network is a supervised learning model which combines the nonlinear mapping capabilities of the feed forward neural networks with the Bayesian regularization approaches. The network architecture has an input layer, one or more hidden layers and an output layer that are connected by weighted links. Bayesian regularization enhances generalization by altering the objective function to minimize the prediction error and the complexity of the network, which helps to prevent overfitting and increase predictive performance on unknown data [28, 29, 50]. This attribute makes BRFFNN particularly useful for geotechnical engineering data sets that are generally small and extremely nonlinear.

The Bayesian Regularized Feed-Forward Neural Networks (BR-FFNNs) have recently been employed in many civil and geotechnical engineering applications because of their capacity to simulate complex nonlinear interactions with less possibility of overfitting. BR-FFNN used to estimate the bond strength of corroded reinforcing bars and concrete and reported good predictive performance as compared to previously presented soft-computing models [51]. Later, Bayesian-regularised feed-forward neural network was also used to predict active earth pressure on stiff retaining walls with negative wall–soil friction, proving its power to mimic complicated soil–structure interaction behavior [33]. Isogeometric analysis was used with BR-FFNN to estimate the stability of rectangular tunnels in cohesive-frictional soils [52]. They used a dataset of 12,946 numerical examples and obtained very accurate predictions for ultimate surcharge loading. Multiple feed-forward neural-network designs were examined to predict the uniaxial compressive strength of carbonate rocks and found the Bayesian-regularized network to have good prediction accuracy with a test-set (R^2) of ~ 0.967 [53]. Similarly, Bayesian regularization-trained artificial neural network was used for predicting the thermal conductivity of quartz sand and obtained significantly better performance compared to conventional multiple linear regression, further showing the suitability of Bayesian-regularized neural networks for nonlinear geotechnical prediction problems [54].

3.2. Generalized Regression Neural Network (GRNN)

The Generalized Regression Neural Network (GRNN) is a memory-based neural network, built based on the kernel regression theory [30]. The model has four layers, input, pattern, summation and output layers. Traditional neural networks have an iterative training process.

GRNN predicts directly using a smoothing parameter based on recorded training patterns. This allows for quick model construction with good nonlinear approximation capabilities. GRNN has been shown to have good performance in engineering prediction tasks with complex interactions and relatively limited data sets [30]. Recent studies have demonstrated the applicability of general regression neural networks (GRNNs) to a range of nonlinear civil and geotechnical problems, including pile load–displacement behaviour, soil strength, subsurface stratification, frost heave, seismic site characterization and tunnelling-induced settlement [55-60].

3.3. Multi-Layer Perceptron (MLP) Regressor

The Multi-Layer Perceptron (MLP) Regressor is one of the most used artificial neural network models for regression problems. It consists of linked neurons in the input, hidden and output layers. The connection weights are updated iteratively throughout the training phase by using the backpropagation learning technique to minimize the prediction errors [50]. With the nonlinear activation functions in the hidden layers, the MLP describe complex interactions between the input variables and the target outputs. The MLP Regressor is widely used in prediction studies in geotechnical engineering due to its flexibility and robustness [24].

Recent studies have demonstrated the effectiveness of Multilayer Perceptron (MLP) models for predicting complex nonlinear responses in pile and foundation engineering. MLP-based models was applied successfully to estimate pile bearing capacity using combinations of pile geometry, soil properties, SPT information, and field load-test data [61-63]. MLP can effectively predict pile-bearing behaviour under saline and cold-region conditions, where deterioration mechanisms introduce additional complexity [64]. MLP was applied to socketed-pile settlement prediction and was demonstrated that metaheuristic optimization can considerably improve model accuracy [65]. Similarly, optimized MLP models achieved high predictive performance for bearing capacity, confirming the suitability of MLP for strongly nonlinear geotechnical regression problems [66].

3.4. FT-Transformer

The FT-Transformer is a deep learning architecture based on transformers intended for tabular data. The model turns input variables into learnable feature tokens and employs multi-head self-attention methods to record interactions among features [36].

Unlike conventional neural networks that heavily depend on weighted connections, FT-Transformer learns local and global dependencies of the dataset via attention-based representation learning. This characteristic allows the model to accurately describe complex nonlinear relationships and feature interactions that are typically present in engineering datasets. Recent work has shown that transformer-based systems can attain state-of-the-art performance on structured data prediction challenges [36, 67]

Recent studies have demonstrated the growing potential of Feature Tokenizer Transformer (FT-Transformer) models for handling structured engineering datasets characterized by complex and nonlinear feature interactions. An optimized FT-Transformer framework was applied to predict the compressive strength of high-performance concrete [68], while performance in predicting the creep compliance of high-strength concrete using grouped validation was reported [69]. FT-Transformer can be effectively integrated with complementary deep-learning models to improve prediction accuracy, robustness, and uncertainty assessment for concrete strength estimation. In geotechnical and underground engineering applications [70], the capability of FT-Transformer to extract meaningful information from geological, stress, and mining-related variables for roadway roof-stability assessment was also reported [71]. FT-Transformer was employed as a benchmark model for predicting nonlinear tunnel responses under extreme loading conditions, confirming its applicability to complex civil-engineering tabular data [72].

4. PROPOSED METHODOLOGY

The detailed statistics of the dataset used for training and testing are presented in Table 4. All datas were normalized to the range -1 to $+1$ using the min-max normalization technique prior to model training to improve convergence and numerical stability of the neural network models as follows [24, 73]:

$$X_{norm} = 2 \cdot \frac{X - X_{min.}}{X_{max.} - X_{min.}} - 1 \quad (2)$$

Out of the 38 experimental pile tests, 34 pile tests were used for training, and the remaining 4 pile test results were utilized for the testing phase. 15 data points from each load vs. settlement curve of the pile test were extracted for the formation of the dataset.

Accordingly, 510 dataset arising from 34 experimental pile tests is used for training the model, and 60 dataset arising from 4 experimental pile test are used for the testing phase. This data segregation is done to avoid the inflation of the reported performance.

These four models were selected to provide a comparative assessment of the predictive models representing different learning architectures and generalization mechanisms. The MLP Regressor was chosen as a traditional feedforward neural-network benchmark due to its demonstrated ability to model nonlinear interactions in engineering situations. BRFFNN was considered because Bayesian regularization can manage the network complexity and helps to limit the chance of overfitting, thereby boosting the generalization ability of the model. The GRNN was considered as a non-parametric radial-basis-function-based neural model, which may approximate complicated nonlinear interactions with relatively easy training needs. By comparison, FT-Transformer is a more modern attention-based deep-learning architecture tailored for tabular data. It enables the model to capture complicated interactions among individual input characteristics by means of feature tokenization and self-attention methods. Thus, the chosen models cover standard feedforward learning, regularized neural learning, non-parametric regression, and transformer-based deep learning. Their simultaneous examination offers a wider and more rigorous framework for determining the modeling strategy that is most suitable for forecasting rather than limiting the comparison to algorithms having comparable structures. Accordingly, the selected algorithms provide a progression from conventional neural regression to regularized, non-parametric, and attention-based learning, enabling the influence of model architecture on prediction accuracy and generalization to be systematically examined. The performance of the models utilized is evaluated using eight performance metrics, including Correlation Coefficient (R), Coefficient of Determination (R^2), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Variance Accounted For (VARE), and Median Absolute Error (MEDAE). Rank analysis is performed to understand the precision of the predictive models by comparing their results against the experimentally observed data.

Table 4 Detailed statistics dataset.

Variable	count	mean	std	min	25%	50%	75%	max
Training Phase								
x ₁ (n)	510	3.2059	1.2563	2	2	3	4	8
x ₂ (β/α)	510	0.4915	0.2940	0.2333	0.3111	0.3111	0.4667	1.0000
x ₃ (L _b /D _f)	510	1.0221	0.2046	0.7500	0.7500	1.0000	1.2500	1.2500
x ₄ (s _e /D _f)	510	0.1656	0.2088	0.0002	0.0114	0.0610	0.2605	0.7943
x ₅ (q/c _u)	510	15.9492	8.4341	2.0314	8.1255	16.2511	23.2611	30.4707
Testing Phase								
x ₁ (n)	60	3.2500	0.8362	2	2	3	4	4
x ₂ (β/α)	60	0.5028	0.3016	0.2333	0.2333	0.3111	0.4667	1.0000
x ₃ (L _b /D _f)	60	0.9375	0.2090	0.7500	0.7500	0.7500	1.0000	1.2500
x ₄ (s _e /D _f)	60	0.1753	0.2214	0.0008	0.0113	0.0617	0.2759	0.7724
x ₅ (q/c _u)	60	16.1498	8.7010	2.0314	8.1255	16.2511	24.3766	30.4707

5. PARAMETRIC INDICATORS

The performance of the machine learning models were evaluated by various parametric indicators after the prediction of the bearing capacity. The parametric indicators are outlined as below [74, 75]:

Correlation Coefficient (R),

$$R = \frac{\sum_{j=1}^m (\delta_j - \delta_j'')(\delta_j' - \delta_j'')}{\sqrt{\sum_{j=1}^m (\delta_j - \delta_j'')^2 \sum_{j=1}^m (\delta_j' - \delta_j'')^2}} \quad (3)$$

Coefficient of Determination (R^2)

$$R^2 = \frac{\sum_{j=1}^m (\delta_j' - \delta_j'')^2}{\sum_{j=1}^m (\delta_j'' - \delta_j')^2} \quad (4)$$

Mean Squared Error (MSE)

$$MSE = \frac{\sum_{j=1}^m (\delta_j - \delta_j')^2}{m} \quad (5)$$

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{j=1}^m (\delta_j' - \delta_j'')^2}{m}} \quad (6)$$

Mean Absolute Error (MAE)

$$MAE = \frac{\sum_{j=1}^m |\delta_j' - \delta_j''|}{m} \quad (7)$$

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{\sum_{j=1}^m |\delta_j' - \delta_j''|}{\sum_{j=1}^m \delta_j'} * \frac{100}{m} \quad (8)$$

Variance Accounted For (VARE)

$$VARE = \left(1 - \frac{\sum_{j=1}^m (\delta_j' - \delta_j'')^2}{\sum_{j=1}^m \delta_j'^2}\right) \quad (9)$$

Median Absolute Error (MEDAE)

$$MEDAE = \text{median}(\delta_j - \delta_j') \quad (10)$$

Where, δ_j = observed value; δ_j' = predicted value; δ_j'' = mean value; m = number of data.

6. RESULTS AND DISCUSSION

6.1. Bearing capacity prediction

Each model was subjected to hyperparameter optimization for increased robustness and fair contribution of all the models. Hyperparameters of the prediction framework were optimized as shown in Table 5. The performance of the model was assessed using Correlation Coefficient (R), Coefficient of Determination (R^2), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Variance Accounted For (VARE) and Median Absolute Error (MEDAE) and the results are presented in Table 6.

The results indicate that all four machine learning models successfully predicted the bearing capacity with high accuracy. Nevertheless, the FT-Transformer emerged as the most effective model, achieving the highest predictive accuracy ($R = 0.9973$, $R^2 = 0.9940$) and the lowest error

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

indices (RMSE = 0.6701, MAE = 0.5310). The results showcased that machine learning algorithms, particularly FT-Transformer, are very effective for predicting the bearing capacity of branched piles from static load test data. FT-Transformer's superior performance across training and testing phases indicates strong calibration, stability, and generalization. Its consistently low error metrics and high variance accounted for suggest minimal residual deviation and reliable extrapolation to unseen data. BRFFNN and MLP Regressor offer viable alternatives; GRNN shows the least capability; however, GRNN also shows its strong capability to predict the bearing capacity of branched pile. Figure 5. showcases the comparison of a few important parametric indicators for different models during the testing and training phases.

Table 5 Selected hyperparameters.

Model	Hyperparameters	Values
FT-Transformer	input_dim	4
	embed_dim	128
	num_heads	4
	num_transformer_blocks	2
	ff_hidden_dim	256
	dropout_rate	0.1
	learning_rate	0.001
	batch_size	64
	num_epochs	500
	weight_decay	0.00001
	optimizer_name	AdamW
BRFFNN	Learning Rate	0.001
	Epochs	500
	Batch Size	32
	Validation Split	0.2
	L2 Regularization (lambda)	0.01
	Optimizer	Adam
	Activation Functions	ReLU (Hidden), Linear (Output)
	Dropout Rate	0.2
	Early Stopping Patience	50
	Loss Function	Mean Squared Error
MLP Regressor	activation	relu
	alpha	0.0001
	hidden_layer_sizes	(100,)
	learning_rate_init	0.1
	max_iter	500
	solver	sgd
GRNN	Optimal Smoothing Parameter (sigma)	0.01

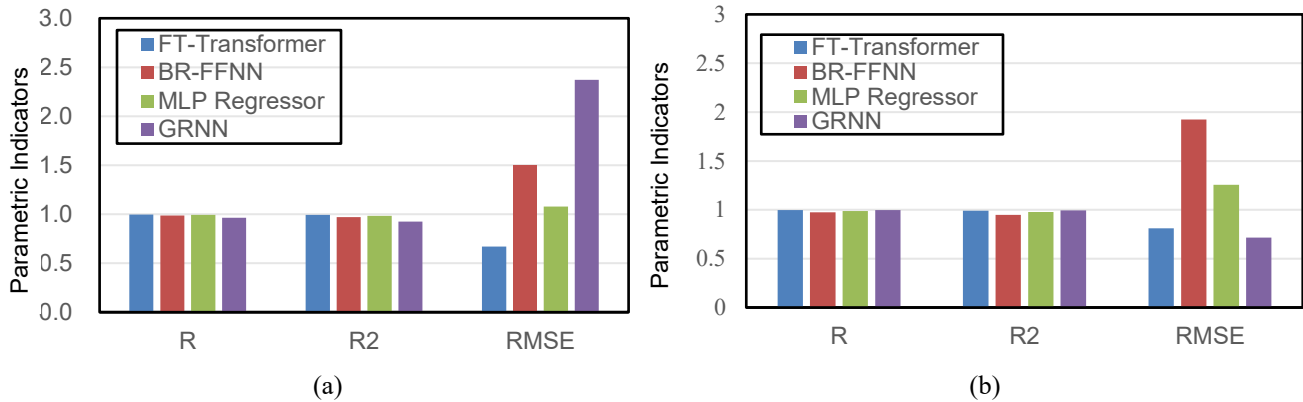


Fig. 5 Comparison of Parametric indicators for different models: (a) Testing phase and (b) Training phase.

Figure 6 showcases the regression plots for the testing phase, emphasizing that FT-Transformer predictions closely align with the regression line, confirming its superior accuracy. Figure 7 also showcases the distinction plots, which illustrate the relative performance differences among models in wave format, further emphasizing FT-Transformer's consistency. The FT-Transformer algorithm is the most reliable and generalizable approach for forecasting the bearing capacity of branched piles on peat soil and has shown superior performance in all evaluation phases. Machine learning models can anticipate difficult geotechnical applications with high accuracy, particularly FT-Transformer.

6.2. Rank Analysis

A basic technique to compare model performance is “Rank Analysis” [41, 76]. For each performance metric, the best-performing model is assigned a score of n ($n=4$ in this study, which is the number of models) and the worst a score of 1. Table 7 evidently shows that FT-Transformer surpassed the other models in obtaining the highest overall rank of 29 in the test set, which is better than all the other algorithms. The lowest rank observed was 13, obtained by the GRNN model in the testing phase.

Table 6 Performance metrics indicators for different ML models.

Model	Phase	R	R ²	MSE	RMS E	MAE	MAPE	VAR E	MEDA E	Time (s)
FT-Transformer	Training	0.9964	0.9907	0.6581	0.8112	0.5822	10.9400	0.9907	0.4238	29.0344
	Testing	0.9973	0.9940	0.4490	0.6701	0.5310	7.6936	0.9940	0.5413	0.0148
BRFFNN	Training	0.9759	0.9478	3.7054	1.9249	1.4535	26.2015	0.9503	1.0941	69.3148
	Testing	0.9883	0.9697	2.2568	1.5023	1.0112	22.9028	0.9724	0.6988	0.4220
MLPRegressor	Training	0.9890	0.9778	1.5792	1.2567	0.8279	16.1230	0.9778	0.4760	0.3167
	Testing	0.9929	0.9844	1.1592	1.0767	0.8217	14.8052	0.9845	0.6612	0.0016
GRNN	Training	0.9964	0.9928	0.5142	0.7171	0.3385	9.0524	0.9928	0.0000	0.0001
	Testing	0.9644	0.9244	5.6310	2.3730	1.2805	12.7401	0.9298	0.7773	0.0011

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

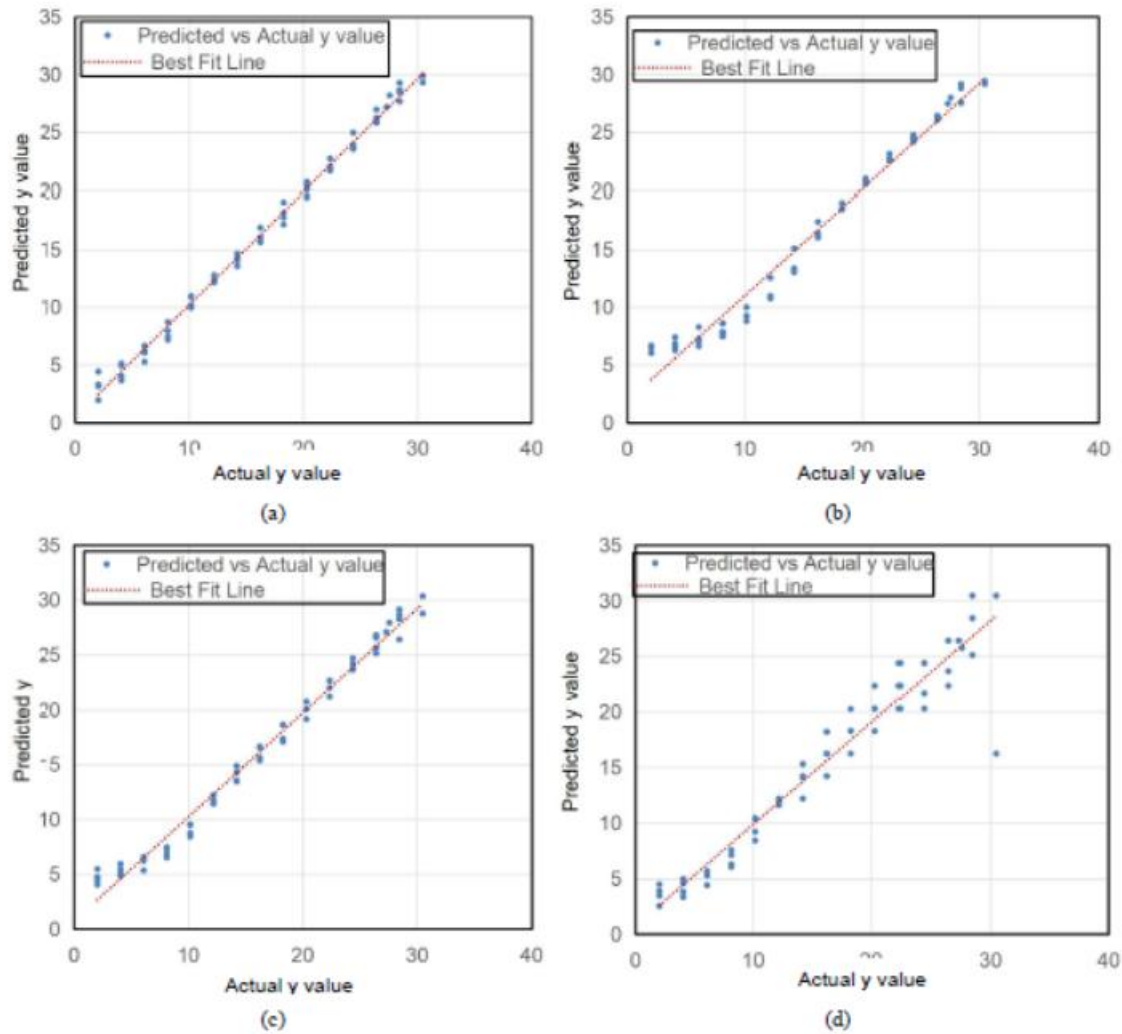


Fig. 6 Regression plots of Actual vs Predicted y for the testing phase: (a) FT-Transformer, (b) BRFFNN, (c) MLP Regressor, and (d) GRNN.

Table 7 Rank analysis result

Model	R	R ²	MSE	RMSE	MAE	MAPE	VARE	MEDAE	Rank
Testing Phase									
FT-Transformer	4	4	4	4	4	4	1	4	29
BR-FFNN	2	2	2	2	2	1	3	2	16
MLPRegressor	3	3	3	3	3	2	2	3	22
GRNN	1	1	1	1	1	3	4	1	13

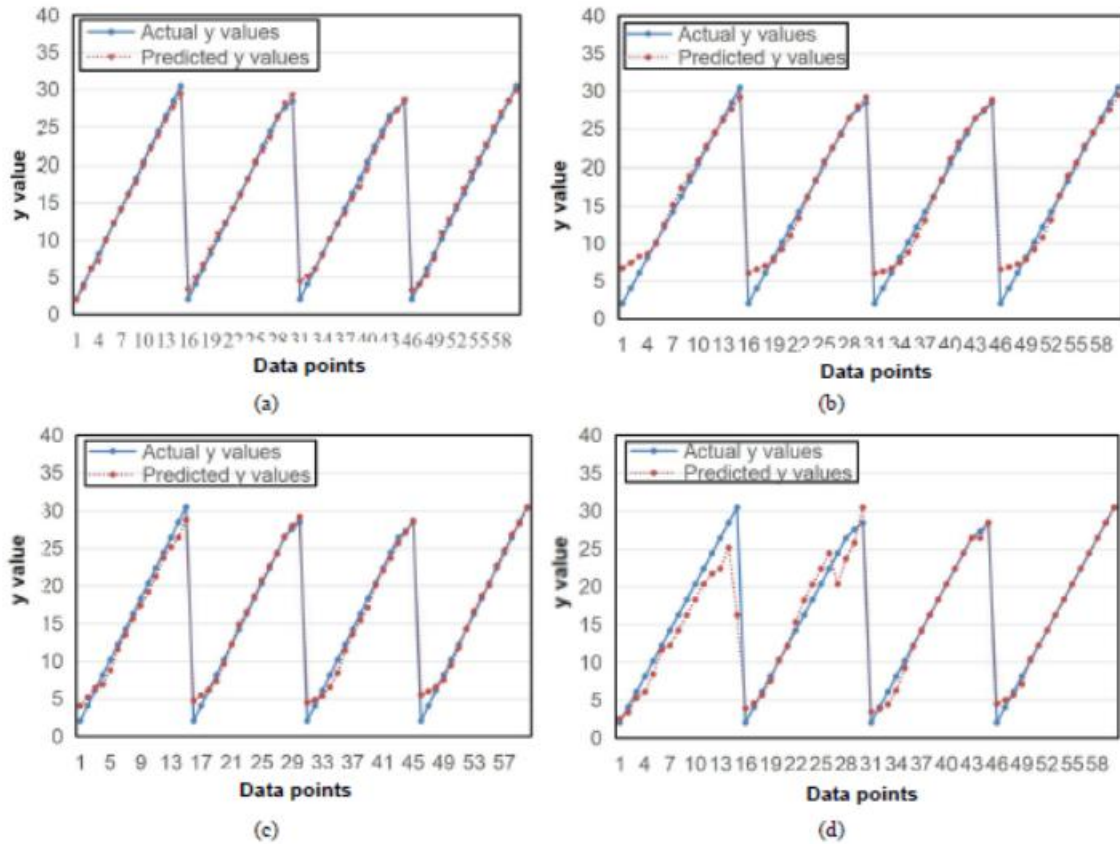


Fig. 7 Distinction plots of Actual vs Predicted y for the testing phase: (a) FT-Transformer, (b) BRFFNN, (c) MLP Regressor, and (d) GRNN.

7. CONCLUSION

Present study tends to propose a soft computing approach for pre-assessment of bearing capacity due to branched piles. Data generated was further analysed via sensitivity analysis. Proposed ML models were evaluated via multiple parameters that finally concluded that FT-Transformer as best-performing model. The salient concluding remarks from present study are stated below:

1. The FT-Transformer exhibited superior predictive performance in estimating the bearing capacity of BP, as evidenced by the statistical indicators, achieving an R^2 of 0.9940, RMSE of 0.6701, MAE of 0.5310, and MAPE of 7.6936. These results demonstrate its high accuracy and reliability compared with the other adopted ML models.

2. FT-Transformer showed significant reductions in error metrics, better than BRFFNN, MLP Regressor, and GRNN on RMSE, MAE, and MAPE. The results demonstrate that the FT-Transformer is robust and reliable for estimating the bearing capacity of a branched pile.
3. The results show that the most significant design parameter controlling the bearing capacity of the branched pile is the branch length ratio (L_b/D_f) with a CAM value of 0.8694, followed by the settlement ratio (s_e/D_f) with a CAM value of 0.8502 and the number of branches (n) with a CAM value of 0.8280, while the branch angle (β/α) with a CAM value of 0.7559 has the least but still significant influence. The results are of major significance for the optimization design of the branched pile. More attention should be paid to the selection of the right length of the branch and the management of settlement behavior to achieve better load-carrying performance.
4. The Predictive accuracy and model reliability can be improved by incorporating more additional experimental datasets to widen the representation of input-parameter variability. This study specifically addresses the branched pile foundation systems on peat.
5. The primary limitation is that the developed model have been trained and validated for branched pile foundation on peat only therefore its predictive performance may not be directly transferable to other types of soils such as sandy, silty, or mixed soil conditions without further calibration.

Author contributions: All authors contributed to the conception and design of the study. Material preparation, data collection, and data analysis were carried out by Dinesh Kumar Jamatia and Ashim Kanti Dey. The first draft of the manuscript was prepared by Dinesh Kumar Jamatia, and both authors reviewed, revised, and provided critical comments on successive versions of the manuscript. Both authors read and approved the final version of the manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability: Relevant supporting data related to the study are shown in Table [A1](#).

Declarations

Conflict of interest The authors declare no competing interest.

REFERENCES

- [1] Paul, A., Hussain, M., Ramu, B. (2021) The physicochemical properties and microstructural characteristics of peat and their correlations: reappraisal. *International Journal of Geotechnical Engineering*, 15(6), 692–703. <https://doi.org/10.1080/19386362.2018.1483099>
- [2] Debnath, R., Saha, R., Haldar, S. (2021) Static and dynamic characterisation of Agartala peat. *Mires and Peat*, 27, 25. <https://doi.org/10.19189/MaP.2020.BG.StA.2132>
- [3] Tunar Özcan, N., Ulusay, R., & Işık, N. S. (2019). Assessment of dynamic site response of the peat deposits at an industrial site (Turkey) and comparison with some seismic design codes. *Bulletin of Engineering Geology and the Environment*, 78(4), 2215–2235. <https://doi.org/10.1007/s10064-018-1285-7>
- [4] Tunar Özcan, N., Ulusay, R., & Işık, N. S. (2020). Geo-engineering characterization and an approach to estimate the in-situ long-term settlement of a peat deposit at an industrial district. *Engineering Geology*, 265, 105329. <https://doi.org/10.1016/j.enggeo.2019.105329>
- [5] Debnath, R., Saha, R., Haldar, S. (2022) Assessment of small strain dynamic soil properties of railway site Agartala, India, by bender element tests. *Arabian Journal of Geosciences*, 15(18), 1500. <https://doi.org/10.1007/s12517-022-10749-4>
- [6] Debnath R, Dey AK, Paul A, Kundu SK (2025) Critical appraisal of geo-engineering properties of tropical peat soils. *Sāadhanā* 50(3). <https://doi.org/10.1007/s12046-025-02884-x>
- [7] Mahmod AA-W, Mohd S, Mohd Masirin MI, Ahmad Tajudin SA, Bakar I, Zainorabidin A, Kifli AZ, Ling JH (2016) Construction of buildings on peat: case studies and lessons learned. *MATEC Web of Conferences* 47:03013. <https://doi.org/10.1051/mateconf/20164703013>
- [8] Marshall C, Large DJ, Athab A, Evers SL, Sowter A, Marsh S, Sjögersten S (2018) Monitoring tropical peat related settlement using ISBAS InSAR, Kuala Lumpur International Airport (KLIA). *Engineering Geology* 244:57–65. <https://doi.org/10.1016/j.enggeo.2018.07.015>
- [9] Singh, M.J., Yao, K., Hou, Y., Liu, S., Yao, Z. and Borana, L., 2025. Deep cement mixing columns for infrastructure development: a review of research progress, challenges, and future research directions. *Geomechanics and Geoengineering*, 20(4), pp.911-936.

- [10] Yu, Xiaoxuan, Yao Hou, Haizuo Zhou, Tao Yao, Yewei Zheng, and Zhenyu Wang. "Probabilistic assessment of seismic settlement of embankment constructed on soft soil by combining simulation of stochastic ground motion and spatial soil variability." *Soil Dynamics and Earthquake Engineering* 202 (2026): 109993.
- [11] Mesri G (2003) Primary compression and secondary compression. *Geotechnical Special Publication* 119:122–166. [https://doi.org/10.1061/40659\(2003\)5](https://doi.org/10.1061/40659(2003)5)
- [12] Ilamparuthi, K., Dickin, E.A., (2001) The influence of soil reinforcement on the uplift behaviour of belled piles embedded in sand. *Geotextiles and Geomembranes*, 19(1), 1–22. [https://doi.org/10.1016/S0266-1144\(00\)00010-8](https://doi.org/10.1016/S0266-1144(00)00010-8)
- [13] Lv, Y., Liu, H., Ding, X., Kong, G., (2012) Field tests on bearing characteristics of X-section pile composite foundation. *Journal of Performance of Constructed Facilities*, 26(2), 180–189. [https://doi.org/10.1061/\(ASCE\)CF.1943-5509.0000247](https://doi.org/10.1061/(ASCE)CF.1943-5509.0000247)
- [14] Zhang, M., Xu, P., Cui, W., & Gao, Y., (2018) Bearing behavior and failure mechanism of squeezed branch piles. *Journal of Rock Mechanics and Geotechnical Engineering*, 10(5), pp.935–946. <https://doi.org/10.1016/j.jrmge.2017.12.010>
- [15] Zhou, J., Huang, X., Zhang, J., Wei, L., Yuan, J., (2021) Experimental investigation of the uplift and lateral bearing capacity of root piles. *Soil Mechanics and Foundation Engineering*, 57(6), 473–479. <https://doi.org/10.1007/s11204-021-09695-2>
- [16] Osborne N, Ng DCC (2009) Jet grouting (JGP) for deep excavation—The importance of quality control. In: *Ground Improvement Technologies and Case Histories (ISGI09)*, Singapore. <https://doi.org/10.3850/GI155>
- [17] Walker SA, Rollo FL (2014) SFMOMA expansion—Multiple uses of ground improvement: a case study. In: *DFI/EFCC International Conference on Piling and Deep Foundations Conference Proceedings*, Stockholm, Sweden, 23 May 2014. Deep Foundations Institute.
- [18] Seo SH, Lee JS, Jung EY, Park SY, Lee HB (2020) Development and reliability verification of quality control system for compaction grouting method. *Journal of the Korean Geosynthetics Society* 19(4):11–20. <https://doi.org/10.12814/jkgss.2020.19.4.011>
- [19] Xiong, L., Li, G., Zhou, Y., He, G., (2023) Experimental and analytical investigation of the bearing capacity of bulbs for squeezed branch piles. *International Journal of Geomechanics*, 23(5), 04023045. <https://doi.org/10.1061/IJGNAI.GMENG-8298>

- [20] Wang, D., Ju, Y., Song, M., Zhao, J., Bai, J., (2023) Uplift bearing capacity of squeezed branch pile group. *Geotechnical and Geological Engineering* 41(1), 283–293. <https://doi.org/10.1007/s10706-022-02279-3>
- [21] Li, X., Dai, G., Gong, W., Zhu, M., (2020) Experimental investigations of capacity response of root piles on combination of O-Cell test and conventional head-down test methods. In: *Proceedings of GeoMEast 2019 International Congress and Exhibition*, Sustainable Civil Infrastructures. Springer, Cham, pp. 18–39. https://doi.org/10.1007/978-3-030-34190-9_3
- [22] Luo, X.-G., Ren, W.-X., Yin, Y.-G., Yu, Y.-C., (2022) An extended BNWF model for root piles incorporating soil–root interactions under lateral loading. *Computers and Geotechnics*, 152, 105039. <https://doi.org/10.1016/j.compgeo.2022.105039>
- [23] Du, P., Fu, Y.J., Liu, X.L., Li, G.F., (2023) Model test study on the load-bearing behaviour of multi-tooth piles under different bearing layers. *KSCE Journal of Civil Engineering*, 27(3), 955–965. <https://doi.org/10.1007/s12205-023-1734-0>
- [24] Haykin, S., (2009) *Neural Networks and Learning Machines*, 3rd ed. Pearson Education, Upper Saddle River, NJ.
- [25] Shahin, M.A., Maier, H.R., Jaksa, M.B., (2009) Data-driven methods in geotechnical engineering. *Canadian Geotechnical Journal* 46(3), 247–272. <https://doi.org/10.1139/T08-121>.
- [26] Momeni, E., Nazir, R., Jahed Armaghani, D., Maizir, H., (2014) Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. *Measurement* 57, 122–131. <https://doi.org/10.1016/j.measurement.2014.08.007>.
- [27] Lee, I.M., Lee, J.H., (1996) Prediction of pile bearing capacity using artificial neural networks. *Computers and Geotechnics* 18(3), 189–200. [https://doi.org/10.1016/0266-352X\(95\)00011-9](https://doi.org/10.1016/0266-352X(95)00011-9).
- [28] MacKay, D.J.C., (1992) Bayesian interpolation. *Neural Computation* 4(3), 415–447. <https://doi.org/10.1162/neco.1992.4.3.415>.
- [29] Foresee, F.D., Hagan, M.T., (1997) Gauss–Newton approximation to Bayesian learning. In: *Proceedings of the 1997 IEEE International Joint Conference on Neural Networks (IJCNN'97)*, Houston, TX, USA, 1930–1935. <https://doi.org/10.1109/IJCNN.1997.614194>.
- [30] Specht, D.F., (1991) A general regression neural network. *IEEE Transactions on Neural Networks* 2(6), 568–576. <https://doi.org/10.1109/72.97934>.

- [31] Singh, A.P., Sharma, A. (2024). Predicting strength of concrete containing waste foundry sand and glass waste using artificial neural network. *Asian Journal of Civil Engineering*, 25, 787–804. <https://doi.org/10.1007/s42107-023-00812-0>
- [32] Dauji, S. (2024). Estimation of fundamental period of concrete frames with infill walls using decision tree. *Asian Journal of Civil Engineering*, 25, 3395–3414. <https://doi.org/10.1007/s42107-024-00986-1>
- [33] Nguyen, T.H., Nguyen, K.V.T., Ho, V.C., Nguyen, D.D. (2024). Efficient hybrid machine learning model for calculating load-bearing capacity of driven piles. *Asian Journal of Civil Engineering*, 25, 883–893. <https://doi.org/10.1007/s42107-023-00818-8>.
- [34] Thapa, I., Kumar, N., Ghani, S., Kumar, S., Gupta, M. (2024). Applications of bentonite in plastic concrete: a comprehensive study on enhancing workability and predicting compressive strength using hybridized AI models. *Asian Journal of Civil Engineering*, 25, 3113–3128. <https://doi.org/10.1007/s42107-023-00966-x>.
- [35] Albostami, A.S., Al-Hamd, R.K.S., Alzabeebee, S., Minto, A., Keawsawasvong, S. (2024). Application of soft computing in predicting the compressive strength of self-compacted concrete containing recyclable aggregate. *Asian Journal of Civil Engineering*, 25, 183–196. <https://doi.org/10.1007/s42107-023-00767-2>.
- [36] Gorishniy, Y., Rubachev, I., Khrulkov, V., Babenko, A., (2021) Revisiting deep learning models for tabular data. *Advances in Neural Information Processing Systems* 34, 18932–18943. <https://proceedings.neurips.cc/paper/2021/hash/9d86d83f925f2149e9edb0ac3b49229c-Abstract.html>
- [37] Butterfield, R., (1999) Dimensional analysis for geotechnical engineers. *Géotechnique* 49(3), 357–366. <https://doi.org/10.1680/geot.1999.49.3.357>
- [38] Jamatia, D.K., Dey, A.K., Debnath, R., (2025) Experimental study on the bearing capacity behaviour of branched piling for stabilisation of constructions on peat: A model investigation. *Mires and Peat* 32, Article 30. <https://doi.org/10.19189/001c.150380>.
- [39] Borthakur, N., & Das, M. (2021). Modelling the capacity of micropiled-raft foundation rested on soft clayey soil using an artificial neural network approach. *International Journal of Geotechnical Engineering*. <https://doi.org/10.1080/19386362.2021.1959138>

- [40] Borthakur, N., Das, M. (2024). Predicting micropile group capacity in soft cohesive soil by artificial neural network. *Indian Geotechnical Journal*. <https://doi.org/10.1007/s40098-024-01058-6>
- [41] Mukherjee A, Borthakur N (2024) Settlement prediction of micropile supported raft using machine learning: modelling and performance evaluation. *J Build Pathol Rehabil* 10:21. <https://doi.org/10.1007/s41024-024-00531-2>
- [42] Mukherjee, A., Borthakur, N. (2025). Prediction of bearing capacity for performance enhancement of micropiled raft foundations using ensemble boosting algorithms. *Geotechnical and Geological Engineering*, 44, 2. <https://doi.org/10.1007/s10706-025-03515-2>
- [43] James, G., Witten, D., Hastie, T., Tibshirani, R., (2021) *An Introduction to Statistical Learning: with Applications in R*, 2nd ed. Springer, New York. <https://doi.org/10.1007/978-1-0716-1418-1>
- [44] Montgomery, D.C., Peck, E.A., Vining, G.G., (2021) *Introduction to Linear Regression Analysis*, 6th ed. John Wiley & Sons, Hoboken, NJ.
- [45] Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., (2005) *Applied Linear Statistical Models*, 5th ed. McGraw-Hill, New York.
- [46] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., (2019) *Multivariate Data Analysis*, 8th ed. Cengage Learning, Hampshire, UK.
- [47] Kim, J.H., (2019) Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology* 72(6), 558–569. <https://doi.org/10.4097/kja.19087>
- [48] Khatti J, Grover KS (2024) Prediction of Uniaxial strength of rocks using relevance vector machine improved with dual kernels and metaheuristic algorithms. *Rock Mech Rock Eng* 57:6227– 6258. <https://doi.org/10.1007/s00603-024-03849-y>
- [49] Moayedi, H., Mosallanezhad, M., Rashid, A.S.A., Muazu, M.A., (2020) A systematic review and meta-analysis of artificial neural network application in geotechnical engineering: Theory and applications. *Neural Computing and Applications* 32(2), 495–518. <https://doi.org/10.1007/s00521-019-04109-9>.
- [50] Rumelhart, D.E., Hinton, G.E., Williams, R.J., (1986) Learning representations by back-propagating errors. *Nature* 323(6088), 533–536. <https://doi.org/10.1038/323533a0>.

- [51] Nguyen, T.H., Nguyen, T., Truong, T.T., Doan, D.T.V., & Tran, D.H. (2023). Corrosion effect on bond behavior between rebar and concrete using Bayesian regularized feed-forward neural network. *Structures*, 51, 1525–1538. <https://doi.org/10.1016/j.istruc.2023.03.128>
- [52] Nguyen-Minh, T., Bui-Ngoc, T., Shiao, J., Nguyen, T., & Nguyen-Thoi, T. (2025). Stability of rectangular tunnels in cohesive-frictional soil under surcharge loading using isogeometric analysis and Bayesian neural networks. *Advances in Engineering Software*, 201, 103861. <https://doi.org/10.1016/j.advengsoft.2024.103861>
- [53] Naal-Pech, J.W., Palemón-Arcos, L. & El Hamzaoui, Y. (2025). Comparative evaluation of feed-forward neural networks for predicting uniaxial compressive strength of Seybaplaya carbonate rock cores. *Applied Sciences*, 15(10), 5609. <https://doi.org/10.3390/app15105609>
- [54] Baghbani, A., Abuel-Naga, H. & Shirkavand, D. (2023). Accurately predicting quartz sand thermal conductivity using machine learning and grey-box AI models. *Geotechnics*, 3(3), 638–660. <https://doi.org/10.3390/geotechnics3030035>
- [55] Ibrahim, F., Alzo'ubi, A. & Odhabi, H. (2023). A generalized regression neural network model to predict CFA piles performance using borehole and static load test data. *Arabian Journal for Science and Engineering*, 48(4), 4403–4419. <https://doi.org/10.1007/s13369-022-06969-1>.
- [56] Xie, J., Huang, J.S., Zeng, C., Huang, S. & Burton, G.J. (2022). A generic framework for geotechnical subsurface modeling with machine learning. *Journal of Rock Mechanics and Geotechnical Engineering*, 14(5), 1366–1379. <https://doi.org/10.1016/j.jrmge.2022.08.001>.
- [57] NourEldin, A.A. (2023). Prediction of yield shear strength of saturated sandy soils using artificial neural networks. *HBRC Journal*, 19(1), 199–213. <https://doi.org/10.1080/16874048.2023.2252720>.
- [58] Li, T., Shu, J. & Chang, D. (2024). Time-series prediction of settlement deformation in shallow buried tunnels based on EMD-SSA-GRNN model. *Scientific Reports*, 14, 324. <https://doi.org/10.1038/s41598-024-51165-w>.
- [59] Cai, H., Wang, C., Ma, Z., Meng, F., Lin, Z., Ren, J. & Li, S. (2024). Predicting frost heave in soil-water systems using the generalized regression neural network optimized with particle swarm optimization algorithm. *Cold Regions Science and Technology*, 226, 104291. <https://doi.org/10.1016/j.coldregions.2024.104291>.

- [60] Darzi, A.G., Sadeghi, H. & Moezzi, S.M.M. (2026). Comparison of GRNN-MC and RF models for predicting soil hydrogeological and geotechnical profile using borehole data. *Scientific Reports*, 16, 19301. <https://doi.org/10.1038/s41598-026-50568-1>.
- [61] Ozturk, B., & Iskander, M. (2026). Lessons learned from training an MLP model using layered feature sets for estimating pile bearing capacity. *Machine Learning and Data Science in Geotechnics*, 2(1), 122–134. <https://doi.org/10.1108/MLAG-12-2025-0022>.
- [62] Jesus, J.N., São Mateus, M.S.C., & Gádea, A.S.M. (2026). Neural model for calculating the bearing capacity of driven and bored piles. *REM – International Engineering Journal*, 79(2), e250034. <https://doi.org/10.1590/0370-44672025790034>.
- [63] Shen, Y. (2024). Optimized systems of multi-layer perceptron predictive model for estimating pile-bearing capacity. *Journal of Engineering and Applied Science*, 71, 52. <https://doi.org/10.1186/s44147-024-00386-x>.
- [64] Deng, Y., Zhang, K., Feng, Z., Zhang, W., Zou, X., & Zhao, H. (2024). Machine learning based prediction model for the pile bearing capacity of saline soils in cold regions. *Structures*, 59, 105735. <https://doi.org/10.1016/j.istruc.2023.105735>
- [65] Wang, L., Jiang, Z., Liang, Z., & Liu, J. (2024). Prediction of socketed pile settlement based on a hybrid form of multilayer perceptron via meta-heuristic algorithms. *Multiscale and Multidisciplinary Modeling, Experiments and Design*, 7, 711–726. <https://doi.org/10.1007/s41939-023-00228-y>.
- [66] Cai, J., Zhou, J., Li, M., & Chen, S. (2025). Using meta-heuristic algorithms with multi-layer perceptron for prediction of ultimate bearing capacity. *Smart Science*, 13(3), 361–379. <https://doi.org/10.1080/23080477.2024.2387929>.
- [67] La, H., & Nguyen, T. (2026). Symbolic distillation of a fully tokenized transformer for settlement prediction in pre-bored grouted planted nodular piles in stratified soils. *Knowledge-Based Systems*, 334, 115186. <https://doi.org/10.1016/j.knosys.2025.115186>
- [68] Kong, L. & Wei, Y. (2026). Prediction model for compressive strength of high-performance concrete based on Feature Tokenizer Transformer and integrated optimization mechanisms. *Construction and Building Materials*, 533, 146809. <https://doi.org/10.1016/j.conbuildmat.2026.146809>

- [69] Wang, X., Xue, Y., Xue, R., Wang, Z., Wu, D., Ye, L., Zhang, Y., et al. (2026). Prediction of creep compliance in high strength concrete using a Feature Tokenizer Transformer with grouped validation. *Discover Applied Sciences*. <https://doi.org/10.1007/s42452-026-09280-y>
- [70] Dong, S. & Zhang, Z. (2025). Hybrid deep learning with conformal prediction for recycled aggregate self-compacting concrete strength prediction. *Buildings*, 15(24), 4419. <https://doi.org/10.3390/buildings15244419>.
- [71] Wang, X., Zhu, Y., Wang, P., Luo, Y., Wei, M. & Wu, Y. (2025). Research on FT-transformer-based intelligent discrimination model for roof stability classification of retreating coal roadways. *Results in Engineering*, 28, 107115. <https://doi.org/10.1016/j.rineng.2025.107115>
- [72] [72] Cheng, R., Chen, W., Hao, H. & Li, Q. (2024). Prediction of BLEVE-induced response of road tunnel using Transformer network with modified self-attention (SAMT). *Engineering Structures*, 314, 118415. <https://doi.org/10.1016/j.engstruct.2024.118415>
- [73] Masters, T., (1993) Practical Neural Network Recipes in C++. *Academic Press*, San Diego. <https://doi.org/10.1680/geot.1999.49.3.357>.
- [74] Amjad Raja MN, Abbas Jaffar ST, Bardhan A, Shukla SK (2023) Predicting and validating the load-settlement behavior of large scale geosynthetic-reinforced soil abutments using hybrid intelligent modeling. *J Rock Mech Geotech Eng* 15(3):773–788. <https://doi.org/10.1016/j.jrmge.2022.04.012>
- [75] Raja MNA, Abdoun T, El-Sekelly W (2024) Exploring the potential of machine learning in stochastic reliability modelling for reinforced soil foundations. *Buildings* 14:954. <https://doi.org/10.3390/buildings14040954>
- [76] Khatti J, Grover KS (2023) Assessment of fine-grained soil compaction parameters using advanced soft computing techniques. *Arab J Geosci* 16:208. <https://doi.org/10.1007/s12517-023-11268-6>

Appendix-I

Table A1. Data set acquired from the experiment of Branched Pile in dimensionless form.

Training Dataset

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c_u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c_u)
1	3	1	0.75	0.0032	2.0314	256	3	0.3111	1.25	0.0005	2.0314
2	3	1	0.75	0.0077	4.0628	257	3	0.3111	1.25	0.0018	4.0628
3	3	1	0.75	0.0139	6.0941	258	3	0.3111	1.25	0.0031	6.0941
4	3	1	0.75	0.0187	8.1255	259	3	0.3111	1.25	0.0048	8.1255
5	3	1	0.75	0.0277	10.1569	260	3	0.3111	1.25	0.0108	10.1569
6	3	1	0.75	0.0448	12.1883	261	3	0.3111	1.25	0.0163	12.1883
7	3	1	0.75	0.0708	14.2197	262	3	0.3111	1.25	0.0271	14.2197
8	3	1	0.75	0.1157	16.2511	263	3	0.3111	1.25	0.0384	16.2511
9	3	1	0.75	0.1907	18.2824	264	3	0.3111	1.25	0.0592	18.2824
10	3	1	0.75	0.2683	20.3138	265	3	0.3111	1.25	0.0884	20.3138
11	3	1	0.75	0.3333	21.6741	266	3	0.3111	1.25	0.1272	22.3452
12	3	1	0.75	0.3743	22.3452	267	3	0.3111	1.25	0.1977	24.3766
13	3	1	0.75	0.5313	24.3766	268	3	0.3111	1.25	0.2933	26.4080
14	3	1	0.75	0.5833	25.1202	269	3	0.3111	1.25	0.4573	28.4393
15	3	1	0.75	0.6890	26.4080	270	3	0.3111	1.25	0.7627	30.4707
16	3	1	1	0.0015	2.0314	271	3	0.2333	0.75	0.0022	2.0314
17	3	1	1	0.0058	4.0628	272	3	0.2333	0.75	0.0043	4.0628
18	3	1	1	0.0112	6.0941	273	3	0.2333	0.75	0.0082	6.0941
19	3	1	1	0.0133	8.1255	274	3	0.2333	0.75	0.0131	8.1255
20	3	1	1	0.0214	10.1569	275	3	0.2333	0.75	0.0228	10.1569
21	3	1	1	0.0368	12.1883	276	3	0.2333	0.75	0.0374	12.1883
22	3	1	1	0.0538	14.2197	277	3	0.2333	0.75	0.0618	14.2197
23	3	1	1	0.0813	16.2511	278	3	0.2333	0.75	0.0930	16.2511
24	3	1	1	0.1129	18.2824	279	3	0.2333	0.75	0.1491	18.2824
25	3	1	1	0.1669	20.3138	280	3	0.2333	0.75	0.1980	20.3138
26	3	1	1	0.2390	22.3452	281	3	0.2333	0.75	0.2500	21.4927
27	3	1	1	0.3306	24.3766	282	3	0.2333	0.75	0.2934	22.3452
28	3	1	1	0.4443	26.4080	283	3	0.2333	0.75	0.4094	24.3766
29	3	1	1	0.5833	27.7501	284	3	0.2333	0.75	0.5536	26.4080

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)
30	3	1	1	0.6867	28.4393	285	3	0.2333	0.75	0.7374	28.4393
31	3	1	1.25	0.0010	2.0314	286	3	0.2333	1	0.0009	2.0314
32	3	1	1.25	0.0037	4.0628	287	3	0.2333	1	0.0035	4.0628
33	3	1	1.25	0.0078	6.0941	288	3	0.2333	1	0.0057	6.0941
34	3	1	1.25	0.0103	8.1255	289	3	0.2333	1	0.0091	8.1255
35	3	1	1.25	0.0168	10.1569	290	3	0.2333	1	0.0188	10.1569
36	3	1	1.25	0.0274	12.1883	291	3	0.2333	1	0.0287	12.1883
37	3	1	1.25	0.0428	14.2197	292	3	0.2333	1	0.0411	14.2197
38	3	1	1.25	0.0634	16.2511	293	3	0.2333	1	0.0596	16.2511
39	3	1	1.25	0.0846	18.2824	294	3	0.2333	1	0.0913	18.2824
40	3	1	1.25	0.1276	20.3138	295	3	0.2333	1	0.1353	20.3138
41	3	1	1.25	0.1777	22.3452	296	3	0.2333	1	0.1877	22.3452
42	3	1	1.25	0.2583	24.3766	297	3	0.2333	1	0.2663	24.3766
43	3	1	1.25	0.3558	26.4080	298	3	0.2333	1	0.3333	25.6643
44	3	1	1.25	0.5211	28.4393	299	3	0.2333	1	0.3924	26.4080
45	3	1	1.25	0.7943	30.4707	300	3	0.2333	1	0.7026	28.4393
46	2	1	0.75	0.0059	2.0314	301	3	0.2333	1.25	0.0005	2.0314
47	2	1	0.75	0.0110	4.0628	302	3	0.2333	1.25	0.0025	4.0628
48	2	1	0.75	0.0163	6.0941	303	3	0.2333	1.25	0.0051	6.0941
49	2	1	0.75	0.0243	8.1255	304	3	0.2333	1.25	0.0077	8.1255
50	2	1	0.75	0.0419	10.1569	305	3	0.2333	1.25	0.0130	10.1569
51	2	1	0.75	0.0728	12.1883	306	3	0.2333	1.25	0.0185	12.1883
52	2	1	0.75	0.0833	12.6961	307	3	0.2333	1.25	0.0286	14.2197
53	2	1	0.75	0.1239	14.2197	308	3	0.2333	1.25	0.0443	16.2511
54	2	1	0.75	0.1891	16.2511	309	3	0.2333	1.25	0.0663	18.2824
55	2	1	0.75	0.2612	18.2824	310	3	0.2333	1.25	0.0943	20.3138
56	2	1	0.75	0.3472	20.3138	311	3	0.2333	1.25	0.1426	22.3452
57	2	1	0.75	0.4167	21.8555	312	3	0.2333	1.25	0.2074	24.3766
58	2	1	0.75	0.4463	22.3452	313	3	0.2333	1.25	0.3108	26.4080
59	2	1	0.75	0.5000	23.1251	314	3	0.2333	1.25	0.4796	28.4393
60	2	1	0.75	0.5880	24.3766	315	3	0.2333	1.25	0.7500	30.4707

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c_u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c_u)
61	2	1	1	0.0043	2.0314	316	2	0.2333	0.75	0.0039	2.0314
62	2	1	1	0.0088	4.0628	317	2	0.2333	0.75	0.0071	4.0628
63	2	1	1	0.0163	6.0941	318	2	0.2333	0.75	0.0108	6.0941
64	2	1	1	0.0209	8.1255	319	2	0.2333	0.75	0.0186	8.1255
65	2	1	1	0.0311	10.1569	320	2	0.2333	0.75	0.0306	10.1569
66	2	1	1	0.0528	12.1883	321	2	0.2333	0.75	0.0486	12.1883
67	2	1	1	0.0844	14.2197	322	2	0.2333	0.75	0.0804	14.2197
68	2	1	1	0.1385	16.2511	323	2	0.2333	0.75	0.1254	16.2511
69	2	1	1	0.2088	18.2824	324	2	0.2333	0.75	0.1689	18.2824
70	2	1	1	0.2888	20.3138	325	2	0.2333	0.75	0.2543	20.3138
71	2	1	1	0.3333	21.3114	326	2	0.2333	0.75	0.3559	22.3452
72	2	1	1	0.3872	22.3452	327	2	0.2333	0.75	0.4167	23.3065
73	2	1	1	0.5214	24.3766	328	2	0.2333	0.75	0.4903	24.3766
74	2	1	1	0.5833	25.2109	329	2	0.2333	0.75	0.5833	25.6643
75	2	1	1	0.6646	26.4080	330	2	0.2333	0.75	0.6371	26.4080
76	2	1	1.25	0.0026	2.0314	331	2	0.2333	1	0.0027	2.0314
77	2	1	1.25	0.0067	4.0628	332	2	0.2333	1	0.0049	4.0628
78	2	1	1.25	0.0127	6.0941	333	2	0.2333	1	0.0093	6.0941
79	2	1	1.25	0.0163	8.1255	334	2	0.2333	1	0.0141	8.1255
80	2	1	1.25	0.0243	10.1569	335	2	0.2333	1	0.0266	10.1569
81	2	1	1.25	0.0390	12.1883	336	2	0.2333	1	0.0469	12.1883
82	2	1	1.25	0.0604	14.2197	337	2	0.2333	1	0.0685	14.2197
83	2	1	1.25	0.0962	16.2511	338	2	0.2333	1	0.1055	16.2511
84	2	1	1.25	0.1392	18.2824	339	2	0.2333	1	0.1578	18.2824
85	2	1	1.25	0.1986	20.3138	340	2	0.2333	1	0.2238	20.3138
86	2	1	1.25	0.2759	22.3452	341	2	0.2333	1	0.3128	22.3452
87	2	1	1.25	0.3333	23.5785	342	2	0.2333	1	0.4466	24.3766
88	2	1	1.25	0.3726	24.3766	343	2	0.2333	1	0.5000	25.0295
89	2	1	1.25	0.4980	26.4080	344	2	0.2333	1	0.5833	25.9364
90	2	1	1.25	0.7000	28.4393	345	2	0.2333	1	0.6276	26.4080
91	4	0.4667	0.75	0.0011	2.0314	346	2	0.2333	1.25	0.0019	2.0314

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)
92	4	0.4667	0.75	0.0040	4.0628	347	2	0.2333	1.25	0.0042	4.0628
93	4	0.4667	0.75	0.0061	6.0941	348	2	0.2333	1.25	0.0077	6.0941
94	4	0.4667	0.75	0.0107	8.1255	349	2	0.2333	1.25	0.0123	8.1255
95	4	0.4667	0.75	0.0190	10.1569	350	2	0.2333	1.25	0.0204	10.1569
96	4	0.4667	0.75	0.0294	12.1883	351	2	0.2333	1.25	0.0308	12.1883
97	4	0.4667	0.75	0.0436	14.2197	352	2	0.2333	1.25	0.0458	14.2197
98	4	0.4667	0.75	0.0626	16.2511	353	2	0.2333	1.25	0.0668	16.2511
99	4	0.4667	0.75	0.0958	18.2824	354	2	0.2333	1.25	0.0984	18.2824
100	4	0.4667	0.75	0.1468	20.3138	355	2	0.2333	1.25	0.1519	20.3138
101	4	0.4667	0.75	0.2225	22.3452	356	2	0.2333	1.25	0.2241	22.3452
102	4	0.4667	0.75	0.3135	24.3766	357	2	0.2333	1.25	0.3154	24.3766
103	4	0.4667	0.75	0.4267	26.4080	358	2	0.2333	1.25	0.4296	26.4080
104	4	0.4667	0.75	0.5564	28.4393	359	2	0.2333	1.25	0.5833	27.8408
105	4	0.4667	0.75	0.7719	30.4707	360	2	0.2333	1.25	0.6640	28.4393
106	4	0.4667	1	0.0005	2.0314	361	6	0.3111	1.25	0.0005	2.0314
107	4	0.4667	1	0.0024	4.0628	362	6	0.3111	1.25	0.0014	4.0628
108	4	0.4667	1	0.0045	6.0941	363	6	0.3111	1.25	0.0028	6.0941
109	4	0.4667	1	0.0071	8.1255	364	6	0.3111	1.25	0.0043	8.1255
110	4	0.4667	1	0.0119	10.1569	365	6	0.3111	1.25	0.0077	10.1569
111	4	0.4667	1	0.0174	12.1883	366	6	0.3111	1.25	0.0127	12.1883
112	4	0.4667	1	0.0276	14.2197	367	6	0.3111	1.25	0.0224	14.2197
113	4	0.4667	1	0.0444	16.2511	368	6	0.3111	1.25	0.0333	16.2511
114	4	0.4667	1	0.0608	18.2824	369	6	0.3111	1.25	0.0479	18.2824
115	4	0.4667	1	0.0813	20.3138	370	6	0.3111	1.25	0.0584	20.3138
116	4	0.4667	1	0.1279	22.3452	371	6	0.3111	1.25	0.0843	22.3452
117	4	0.4667	1	0.2076	24.3766	372	6	0.3111	1.25	0.1397	24.3766
118	4	0.4667	1	0.3124	26.4080	373	6	0.3111	1.25	0.2070	26.4080
119	4	0.4667	1	0.4607	28.4393	374	6	0.3111	1.25	0.3779	28.4393
120	4	0.4667	1	0.7556	30.4707	375	6	0.3111	1.25	0.7633	30.4707
121	4	0.4667	1.25	0.0003	2.0314	376	8	0.3111	1.25	0.0008	2.0314
122	4	0.4667	1.25	0.0015	4.0628	377	8	0.3111	1.25	0.0024	4.0628

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c_u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c_u)
123	4	0.4667	1.25	0.0023	6.0941	378	8	0.3111	1.25	0.0038	6.0941
124	4	0.4667	1.25	0.0039	8.1255	379	8	0.3111	1.25	0.0059	8.1255
125	4	0.4667	1.25	0.0069	10.1569	380	8	0.3111	1.25	0.0100	10.1569
126	4	0.4667	1.25	0.0103	12.1883	381	8	0.3111	1.25	0.0149	12.1883
127	4	0.4667	1.25	0.0154	14.2197	382	8	0.3111	1.25	0.0256	14.2197
128	4	0.4667	1.25	0.0251	16.2511	383	8	0.3111	1.25	0.0373	16.2511
129	4	0.4667	1.25	0.0409	18.2824	384	8	0.3111	1.25	0.0668	18.2824
130	4	0.4667	1.25	0.0606	20.3138	385	8	0.3111	1.25	0.0952	20.3138
131	4	0.4667	1.25	0.0931	22.3452	386	8	0.3111	1.25	0.1263	22.3452
132	4	0.4667	1.25	0.1317	24.3766	387	8	0.3111	1.25	0.1843	24.3766
133	4	0.4667	1.25	0.2095	26.4080	388	8	0.3111	1.25	0.3128	26.4080
134	4	0.4667	1.25	0.3130	28.4393	389	8	0.3111	1.25	0.4973	28.4393
135	4	0.4667	1.25	0.4626	30.4707	390	8	0.3111	1.25	0.7621	30.4707
136	2	0.4667	0.75	0.0048	2.0314	391	4	1	1	0.0008	2.0314
137	2	0.4667	0.75	0.0079	4.0628	392	4	1	1	0.0031	4.0628
138	2	0.4667	0.75	0.0128	6.0941	393	4	1	1	0.0073	6.0941
139	2	0.4667	0.75	0.0209	8.1255	394	4	1	1	0.0092	8.1255
140	2	0.4667	0.75	0.0336	10.1569	395	4	1	1	0.0162	10.1569
141	2	0.4667	0.75	0.0536	12.1883	396	4	1	1	0.0261	12.1883
142	2	0.4667	0.75	0.0904	14.2197	397	4	1	1	0.0413	14.2197
143	2	0.4667	0.75	0.1363	16.2511	398	4	1	1	0.0611	16.2511
144	2	0.4667	0.75	0.1989	18.2824	399	4	1	1	0.0817	18.2824
145	2	0.4667	0.75	0.2759	20.3138	400	4	1	1	0.1163	20.3138
146	2	0.4667	0.75	0.3333	21.5834	401	4	1	1	0.1666	22.3452
147	2	0.4667	0.75	0.3749	22.3452	402	4	1	1	0.2268	24.3766
148	2	0.4667	0.75	0.5086	24.3766	403	4	1	1	0.3495	26.4080
149	2	0.4667	0.75	0.5833	25.3923	404	4	1	1	0.5209	28.4393
150	2	0.4667	0.75	0.6612	26.4080	405	4	1	1	0.7556	30.4707
151	2	0.4667	1	0.0030	2.0314	406	4	1	1.25	0.0005	2.0314
152	2	0.4667	1	0.0059	4.0628	407	4	1	1.25	0.0021	4.0628
153	2	0.4667	1	0.0112	6.0941	408	4	1	1.25	0.0043	6.0941

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)
154	2	0.4667	1	0.0171	8.1255	409	4	1	1.25	0.0058	8.1255
155	2	0.4667	1	0.0284	10.1569	410	4	1	1.25	0.0092	10.1569
156	2	0.4667	1	0.0504	12.1883	411	4	1	1.25	0.0128	12.1883
157	2	0.4667	1	0.0768	14.2197	412	4	1	1.25	0.0194	14.2197
158	2	0.4667	1	0.1200	16.2511	413	4	1	1.25	0.0334	16.2511
159	2	0.4667	1	0.1817	18.2824	414	4	1	1.25	0.0521	18.2824
160	2	0.4667	1	0.2555	20.3138	415	4	1	1.25	0.0787	20.3138
161	2	0.4667	1	0.3363	22.3452	416	4	1	1.25	0.1185	22.3452
162	2	0.4667	1	0.4167	23.6692	417	4	1	1.25	0.1653	24.3766
163	2	0.4667	1	0.4591	24.3766	418	4	1	1.25	0.2419	26.4080
164	2	0.4667	1	0.5833	25.7550	419	4	1	1.25	0.3369	28.4393
165	2	0.4667	1	0.6442	26.4080	420	4	1	1.25	0.5095	30.4707
166	2	0.4667	1.25	0.0022	2.0314	421	3	0.4667	0.75	0.0024	2.0314
167	2	0.4667	1.25	0.0047	4.0628	422	3	0.4667	0.75	0.0053	4.0628
168	2	0.4667	1.25	0.0092	6.0941	423	3	0.4667	0.75	0.0103	6.0941
169	2	0.4667	1.25	0.0138	8.1255	424	3	0.4667	0.75	0.0161	8.1255
170	2	0.4667	1.25	0.0219	10.1569	425	3	0.4667	0.75	0.0253	10.1569
171	2	0.4667	1.25	0.0334	12.1883	426	3	0.4667	0.75	0.0403	12.1883
172	2	0.4667	1.25	0.0482	14.2197	427	3	0.4667	0.75	0.0646	14.2197
173	2	0.4667	1.25	0.0719	16.2511	428	3	0.4667	0.75	0.0962	16.2511
174	2	0.4667	1.25	0.1037	18.2824	429	3	0.4667	0.75	0.1562	18.2824
175	2	0.4667	1.25	0.1644	20.3138	430	3	0.4667	0.75	0.2243	20.3138
176	2	0.4667	1.25	0.2386	22.3452	431	3	0.4667	0.75	0.3309	22.3452
177	2	0.4667	1.25	0.3277	24.3766	432	3	0.4667	0.75	0.4386	24.3766
178	2	0.4667	1.25	0.4395	26.4080	433	3	0.4667	0.75	0.5000	25.3016
179	2	0.4667	1.25	0.5833	27.8408	434	3	0.4667	0.75	0.5945	26.4080
180	2	0.4667	1.25	0.6610	28.4393	435	3	0.4667	0.75	0.7750	28.4393
181	4	0.3111	0.75	0.0007	2.0314	436	3	0.4667	1.25	0.0007	2.0314
182	4	0.3111	0.75	0.0023	4.0628	437	3	0.4667	1.25	0.0030	4.0628
183	4	0.3111	0.75	0.0038	6.0941	438	3	0.4667	1.25	0.0059	6.0941
184	4	0.3111	0.75	0.0059	8.1255	439	3	0.4667	1.25	0.0089	8.1255

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c_u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c_u)
185	4	0.3111	0.75	0.0144	10.1569	440	3	0.4667	1.25	0.0144	10.1569
186	4	0.3111	0.75	0.0240	12.1883	441	3	0.4667	1.25	0.0208	12.1883
187	4	0.3111	0.75	0.0373	14.2197	442	3	0.4667	1.25	0.0307	14.2197
188	4	0.3111	0.75	0.0555	16.2511	443	3	0.4667	1.25	0.0480	16.2511
189	4	0.3111	0.75	0.0800	18.2824	444	3	0.4667	1.25	0.0694	18.2824
190	4	0.3111	0.75	0.1083	20.3138	445	3	0.4667	1.25	0.0979	20.3138
191	4	0.3111	0.75	0.1722	22.3452	446	3	0.4667	1.25	0.1471	22.3452
192	4	0.3111	0.75	0.2571	24.3766	447	3	0.4667	1.25	0.2295	24.3766
193	4	0.3111	0.75	0.3553	26.4080	448	3	0.4667	1.25	0.3311	26.4080
194	4	0.3111	0.75	0.4934	28.4393	449	3	0.4667	1.25	0.4906	28.4393
195	4	0.3111	0.75	0.7555	30.4707	450	3	0.4667	1.25	0.7654	30.4707
196	4	0.3111	1	0.0003	2.0314	451	2	0.3111	0.75	0.0033	2.0314
197	4	0.3111	1	0.0013	4.0628	452	2	0.3111	0.75	0.0059	4.0628
198	4	0.3111	1	0.0023	6.0941	453	2	0.3111	0.75	0.0086	6.0941
199	4	0.3111	1	0.0039	8.1255	454	2	0.3111	0.75	0.0148	8.1255
200	4	0.3111	1	0.0078	10.1569	455	2	0.3111	0.75	0.0238	10.1569
201	4	0.3111	1	0.0136	12.1883	456	2	0.3111	0.75	0.0393	12.1883
202	4	0.3111	1	0.0245	14.2197	457	2	0.3111	0.75	0.0660	14.2197
203	4	0.3111	1	0.0350	16.2511	458	2	0.3111	0.75	0.1079	16.2511
204	4	0.3111	1	0.0478	18.2824	459	2	0.3111	0.75	0.1642	18.2824
205	4	0.3111	1	0.0729	20.3138	460	2	0.3111	0.75	0.2331	20.3138
206	4	0.3111	1	0.1118	22.3452	461	2	0.3111	0.75	0.3305	22.3452
207	4	0.3111	1	0.1704	24.3766	462	2	0.3111	0.75	0.4167	23.5785
208	4	0.3111	1	0.2577	26.4080	463	2	0.3111	0.75	0.4733	24.3766
209	4	0.3111	1	0.4088	28.4393	464	2	0.3111	0.75	0.5833	25.8457
210	4	0.3111	1	0.7391	30.4707	465	2	0.3111	0.75	0.6278	26.4080
211	4	0.3111	1.25	0.0002	2.0314	466	2	0.3111	1	0.0025	2.0314
212	4	0.3111	1.25	0.0005	4.0628	467	2	0.3111	1	0.0040	4.0628
213	4	0.3111	1.25	0.0013	6.0941	468	2	0.3111	1	0.0068	6.0941
214	4	0.3111	1.25	0.0027	8.1255	469	2	0.3111	1	0.0122	8.1255
215	4	0.3111	1.25	0.0053	10.1569	470	2	0.3111	1	0.0218	10.1569

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)
216	4	0.3111	1.25	0.0084	12.1883	471	2	0.3111	1	0.0334	12.1883
217	4	0.3111	1.25	0.0126	14.2197	472	2	0.3111	1	0.0610	14.2197
218	4	0.3111	1.25	0.0208	16.2511	473	2	0.3111	1	0.0955	16.2511
219	4	0.3111	1.25	0.0336	18.2824	474	2	0.3111	1	0.1343	18.2824
220	4	0.3111	1.25	0.0501	20.3138	475	2	0.3111	1	0.1667	19.4070
221	4	0.3111	1.25	0.0728	22.3452	476	2	0.3111	1	0.1978	20.3138
222	4	0.3111	1.25	0.1093	24.3766	477	2	0.3111	1	0.2759	22.3452
223	4	0.3111	1.25	0.1708	26.4080	478	2	0.3111	1	0.4061	24.3766
224	4	0.3111	1.25	0.2487	28.4393	479	2	0.3111	1	0.5000	25.3923
225	4	0.3111	1.25	0.4095	30.4707	480	2	0.3111	1	0.6083	26.4080
226	3	0.3111	0.75	0.0021	2.0314	481	4	0.2333	1	0.0003	2.0314
227	3	0.3111	0.75	0.0031	4.0628	482	4	0.2333	1	0.0018	4.0628
228	3	0.3111	0.75	0.0058	6.0941	483	4	0.2333	1	0.0024	6.0941
229	3	0.3111	0.75	0.0108	8.1255	484	4	0.2333	1	0.0053	8.1255
230	3	0.3111	0.75	0.0195	10.1569	485	4	0.2333	1	0.0107	10.1569
231	3	0.3111	0.75	0.0303	12.1883	486	4	0.2333	1	0.0164	12.1883
232	3	0.3111	0.75	0.0524	14.2197	487	4	0.2333	1	0.0173	14.2197
233	3	0.3111	0.75	0.0847	16.2511	488	4	0.2333	1	0.0410	16.2511
234	3	0.3111	0.75	0.1258	18.2824	489	4	0.2333	1	0.0538	18.2824
235	3	0.3111	0.75	0.1813	20.3138	490	4	0.2333	1	0.0778	20.3138
236	3	0.3111	0.75	0.2613	22.3452	491	4	0.2333	1	0.1248	22.3452
237	3	0.3111	0.75	0.3537	24.3766	492	4	0.2333	1	0.1909	24.3766
238	3	0.3111	0.75	0.4167	25.3923	493	4	0.2333	1	0.2870	26.4080
239	3	0.3111	0.75	0.4902	26.4080	494	4	0.2333	1	0.4463	28.4393
240	3	0.3111	0.75	0.6696	28.4393	495	4	0.2333	1	0.7478	30.4707
241	3	0.3111	1	0.0008	2.0314	496	4	0.2333	1.25	0.0002	2.0314
242	3	0.3111	1	0.0024	4.0628	497	4	0.2333	1.25	0.0005	4.0628
243	3	0.3111	1	0.0040	6.0941	498	4	0.2333	1.25	0.0020	6.0941
244	3	0.3111	1	0.0063	8.1255	499	4	0.2333	1.25	0.0035	8.1255
245	3	0.3111	1	0.0148	10.1569	500	4	0.2333	1.25	0.0057	10.1569
246	3	0.3111	1	0.0244	12.1883	501	4	0.2333	1.25	0.0092	12.1883

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)
247	3	0.3111	1	0.0377	14.2197	502	4	0.2333	1.25	0.0136	14.2197
248	3	0.3111	1	0.0559	16.2511	503	4	0.2333	1.25	0.0236	16.2511
249	3	0.3111	1	0.0808	18.2824	504	4	0.2333	1.25	0.0390	18.2824
250	3	0.3111	1	0.1093	20.3138	505	4	0.2333	1.25	0.0576	20.3138
251	3	0.3111	1	0.1731	22.3452	506	4	0.2333	1.25	0.0857	22.3452
252	3	0.3111	1	0.2575	24.3766	507	4	0.2333	1.25	0.1207	24.3766
253	3	0.3111	1	0.3558	26.4080	508	4	0.2333	1.25	0.1851	26.4080
254	3	0.3111	1	0.4946	28.4393	509	4	0.2333	1.25	0.2793	28.4393
255	3	0.3111	1	0.7758	30.4707	510	4	0.2333	1.25	0.4298	30.4707

Testing dataset

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)
1	4	1	0.75	0.0013	2.0314	31	2	0.3111	1.25	0.0016	2.0314
2	4	1	0.75	0.0058	4.0628	32	2	0.3111	1.25	0.0031	4.0628
3	4	1	0.75	0.0113	6.0941	33	2	0.3111	1.25	0.0051	6.0941
4	4	1	0.75	0.0133	8.1255	34	2	0.3111	1.25	0.0098	8.1255
5	4	1	0.75	0.0218	10.1569	35	2	0.3111	1.25	0.0173	10.1569
6	4	1	0.75	0.0371	12.1883	36	2	0.3111	1.25	0.0295	12.1883
7	4	1	0.75	0.0539	14.2197	37	2	0.3111	1.25	0.0411	14.2197
8	4	1	0.75	0.0814	16.2511	38	2	0.3111	1.25	0.0617	16.2511
9	4	1	0.75	0.1133	18.2824	39	2	0.3111	1.25	0.0901	18.2824
10	4	1	0.75	0.1669	20.3138	40	2	0.3111	1.25	0.1410	20.3138
11	4	1	0.75	0.2393	22.3452	41	2	0.3111	1.25	0.2103	22.3452
12	4	1	0.75	0.3308	24.3766	42	2	0.3111	1.25	0.2889	24.3766
13	4	1	0.75	0.4443	26.4080	43	2	0.3111	1.25	0.4053	26.4080
14	4	1	0.75	0.5727	28.4393	44	2	0.3111	1.25	0.5000	27.2967
15	4	1	0.75	0.7724	30.4707	45	2	0.3111	1.25	0.6637	28.4393
16	3	0.4667	1	0.0012	2.0314	46	4	0.2333	0.75	0.0008	2.0314
17	3	0.4667	1	0.0040	4.0628	47	4	0.2333	0.75	0.0028	4.0628

Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat

Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)	Sl. No.	x_1 (n)	x_2 (β/α)	x_3 (L_b/d_f)	x_4 (s_e/d_f)	y (q/c _u)
18	3	0.4667	1	0.0070	6.0941	48	4	0.2333	0.75	0.0052	6.0941
19	3	0.4667	1	0.0112	8.1255	49	4	0.2333	0.75	0.0088	8.1255
20	3	0.4667	1	0.0201	10.1569	50	4	0.2333	0.75	0.0168	10.1569
21	3	0.4667	1	0.0303	12.1883	51	4	0.2333	0.75	0.0258	12.1883
22	3	0.4667	1	0.0443	14.2197	52	4	0.2333	0.75	0.0394	14.2197
23	3	0.4667	1	0.0627	16.2511	53	4	0.2333	0.75	0.0596	16.2511
24	3	0.4667	1	0.0965	18.2824	54	4	0.2333	0.75	0.0919	18.2824
25	3	0.4667	1	0.1478	20.3138	55	4	0.2333	0.75	0.1252	20.3138
26	3	0.4667	1	0.2007	22.3452	56	4	0.2333	0.75	0.1961	22.3452
27	3	0.4667	1	0.2759	24.3766	57	4	0.2333	0.75	0.2855	24.3766
28	3	0.4667	1	0.4276	26.4080	58	4	0.2333	0.75	0.3963	26.4080
29	3	0.4667	1	0.5833	27.5688	59	4	0.2333	0.75	0.5242	28.4393
30	3	0.4667	1	0.7278	28.4393	60	4	0.2333	0.75	0.7662	30.4707

Citation: Dinesh Kumar Jamatia and Ashim Kanti Dey. (2026). Feature Tokenizer Transformer Based Advanced Deep Learning Method for Prediction of the Bearing Capacity of Branched Piles on Peat. International Journal of Civil Engineering and Technology (IJCET), 17(5), 1-42.

Abstract Link: https://iaeme.com/Home/article_id/IJCET_17_05_001

Article Link: https://iaeme.com/MasterAdmin/Journal_uploads/IJCET/VOLUME_17_ISSUE_5/IJCET_17_05_001.pdf

Copyright: © 2026 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Creative Commons license: Creative Commons license: CC BY 4.0



✉ editor@iaeme.com