Engineered Science
DOI: https://dx.doi.org/10.30919/es1082

Performance Evaluation and Triangle Diagram of Deep Learning Models for Embedment Depth Prediction in Cantilever Sheet Piles

Pradeep T,[1] Divesh Ranjan Kumar,[2] Nitish Kumar,[3] Warit Wipulanusat,[2,*] Suraparb Keawsawasvong[4] and Jirapon Sunkpho[5]

Abstract

Sheet piles are essential for maintaining the stability and retention of soil in various applications, including railway and highway embankments, offshore structures, post-excavation sites, and slope stabilization projects. The required depth of sheet piles is contingent upon factors such as soil characteristics, groundwater conditions, and the employed construction method. This study focused on predicting the embedment depth of cantilever sheet pile walls in cohesive soil with a cohesionless soil backfill. Artificial intelligence (AI) techniques, specifically deep neural networks (DNNs), recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and bidirectional long short-term memory (Bi-LSTM) networks, are applied for this purpose. Performance evaluation is conducted through rank analysis, performance parameter determination, and comparison of the actual versus predicted curves, accompanied by an error plot. A triangle diagram is introduced as a graphical representation to assess the performance of different datasets or models. External validation was conducted to evaluate the generalizability of the model. All the proposed models meet these criteria, with the DNN model demonstrating superior performance. This comprehensive evaluation demonstrated the effectiveness and robustness of the DNN model as a practical tool for predicting the embedment depth of sheet piles.

Keywords: Embedded depth; Deep learning; Triangle diagram; Sheet pile.
Received: 16 November 2023; Revised: 09 December 2023; Accepted: 02 January 2024.
Article type: Research article.

1. Introduction

Sheet piles are widely used in constructing railway and highway embankments due to their ability to provide robust support to the soil and prevent soil displacement or slippage. Sheet piles are also commonly used to construct offshore structures such as sea walls, jetties, and breakwaters to protect against wave and tidal erosion. In addition, sheet piles are utilized in soil post-excavation procedures to effectively retain the soil and mitigate the risk of collapse during construction activities. These structures can stabilize slopes in hilly terrains, where their purpose is to prevent soil erosion and landslides. In summary, sheet piles are a versatile and efficient solution for retaining soil in various construction projects.

Calculating the depth of a sheet pile involves many factors, including soil properties, groundwater conditions, structural characteristics of the sheet pile, and construction techniques. Several conventional methods, including analytical methods, numerical methods, and empirical methods, are employed to calculate the depth of sheet piles.[1] Analytical methods entail the use of mathematical equations to analyze the behavior of the sheet pile and the soil in which it is installed. Numerical methods, such as finite element analysis, use computational models to simulate the behavior of sheet piles and soil.[2,3] Empirical methods rely on historical project data to predict sheet pile behavior. Several factors affect the required depth of a sheet pile including the soil type, groundwater conditions,

[1] *Department of Civil Engineering, National Institute of Technology, Andhra Pradesh 534101, India.*
[2] *Research Unit in Data Science and Digital Transformation, Department of Civil Engineering, Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathum Thani 12120, Thailand.*
[3] *Department of Civil Engineering, National Institute of Technology, Patna 800005, India.*
[4] *Research Unit in Sciences and Innovative Technologies for Civil Engineering Infrastructures, Department of Civil Engineering, Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathum Thani 12120, Thailand.*
[5] *College of Innovation, Thammasat University, Bangkok 10200, Thailand.*
**Email:* wwarit@engr.tu.ac.th (W. Wipulanusat)

sheet pile material, and structural properties.[4] For example, cohesive soils may require deeper sheet pile penetration to ensure stability, whereas granular soils may permit shallower depths.

Geotechnical engineering pertains to studying the behavior and properties of naturally occurring earth materials, such as soils, rocks, and groundwater, which exhibit variations in composition and characteristics. Uncertainties may arise during the design and construction of structures built on or within the ground.[5] Accurately predicting the depth of sheet piles can pose challenges due to the complexity of the site-specific conditions and environmental factors that can impact their installation.[6] Prior research conducted by.[7] presented the results of a numerical analysis of the behavior of sheet-pile walls when subjected to diverse soil and loading conditions. The findings indicate that the magnitude of the lateral deflection of the wall is dependent primarily on the height of the wall and the properties of the soil. The author additionally examined the impact of factors such as the depth of wall embedment, the stiffness of the wall, and the loading conditions on the behavior of the wall. Gajan[8] presented the issue of designing retaining structures in cohesionless soils and the associated challenges. The paper subsequently describes the different types of retaining structures commonly used and explains the factors that influence their design, including soil properties, wall geometry, and loading conditions. Additionally, the previous study conducted by GuhaRay & Baidya[9] provides a comprehensive overview of cantilever sheet pile walls and emphasizes the importance of reliability analysis in geotechnical engineering. Subsequently, the authors outline their methodology, which involves using the finite-element method to simulate the behavior of the wall and surrounding soil and applying reliability theory to estimate the probability of failure. The paper also includes a detailed description of the different soil types and wall configurations used in the analysis, which provides useful context for interpreting the results.

The empirical approaches suggested by Terzaghi[10] and Rankine[11] are the most widely used classical approaches for calculating the lateral earth pressure and determining the embedment depth of a sheet pile. These involve determining the depth at which the sheet pile needs to be driven into the soil to achieve stability and resist the applied loads. Terzaghi considered the sheet pile as a vertical cantilever retaining wall, and the embedment depth was calculated based on the equilibrium between the lateral earth pressure and the resisting forces of the sheet pile. Rankine's earth pressure theory, on the other hand, is a classical method used in geotechnical engineering to estimate lateral earth pressure acting on retaining structures, such as retaining walls and sheet pile walls. This theory was developed by William John Macquorn Rankine in the 19th century. Rankine proposed two basic conditions for analyzing earth pressure: the active earth pressure and the passive earth pressure. Rankine's theory, while a simplified approach, provides reasonable estimates under certain conditions.

In a recent study, Patil *et al*.[12] applied the finite element method (FEM) to model a sheet pile–soil system and analyzed the results using different parameters, including displacement, bending moment, and shear force. Finally, they conducted a comparative analysis between their results and the analytical solution commonly used in sheet pile design. The results indicated that the performance of sheet piles in dense soil surpasses that in loose soil, and the soil parameters significantly affect the behavior of the sheet piles. Debnath and Pal[13] discussed the application of cantilever sheet pile walls for retaining soil, emphasizing stability and flexibility. Numerous studies on sheet pile behavior, anchorage, and soil conditions have been published, validating their proposed numerical model. This research addresses gaps in the literature, by exploring the impact of wall and anchor stiffness, surcharge loads, and foundation soils on wall deflections. The parametric study revealed significant reductions in deflections with increased wall stiffness and the influence of foundation soil density. The study's findings provide valuable insights for deep earth excavations at urban sites with surrounding structures. Chen *et al.*[14] explored the effectiveness of combining anchored sheet pile quays with deep cement mixing (DCM) columns to enhance the stability of soft soils. Through centrifuge tests and numerical analyses, they demonstrated the shielding effects of DCM columns on lateral earth pressures, bending moments, and deflections. The findings highlight the potential of DCM columns for reducing excavation-induced movements and improving the economic efficiency of anchored sheet pile quays in soft soil conditions. However, the study acknowledges the need for further considerations, including installation effects and groundwater levels, in practical applications. These studies contribute to a broader comprehensive understanding of sheet pile behavior by considering diverse soil conditions and innovative techniques. The collective insights can guide practitioners and engineers in optimizing sheet pile design for specific contexts, ultimately improving the efficiency and stability of deep earth excavations and retaining structures in various geological settings. Employing machine learning techniques can streamline the analysis by providing an efficient prediction of the embedment depth in sheet pile problems. This approach proves advantageous because it necessitates significantly reduced costs in terms of human effort and time, offering a more practical and resource-efficient solution.

In recent years, artificial intelligence (AI) techniques have garnered growing interest across various engineering disciplines.[15,16] This surge in interest can be primarily attributed to the substantial progress in deep learning algorithms and computing power, which have enabled the development of sophisticated AI models capable of performing complex tasks.[17] AI has many potential engineering advantages, including increased efficiency, improved accuracy, and reduced costs.[18] As a result, the application of AI techniques is anticipated to further expand in

prevalence across a wide range of engineering disciplines in the foreseeable future.[19] Deep neural networks (DNNs), recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and bidirectional long short-term memory (Bi-LSTM) networks are all types of deep learning models that can be used for these prediction tasks.

DNNs offer numerous advantages, such as high accuracy, adaptability and generalization, parallel processing and scalability, and unstructured data handling. These advantageous characteristics have rendered DNNs highly potent and versatile tools across diverse domains.[20,21] RNNs possess numerous advantages, including sequential information processing, temporal dependency handling, flexibility in input and output, time series prediction, and long short-term memory; additionally, RNNs are especially suitable for handling sequential data processing.[22–24] LSTM networks present several advantages over conventional RNNs when dealing with long-range dependencies and memory retention in sequential data. Some advantages of LSTM networks include long-term dependency handling, memory retention, gating mechanisms, and improvements over standard RNNs.[22,25,26] Bi-LSTM networks present a range of advantages compared to conventional LSTM networks by considering both past and future contexts when processing sequential data, including bidirectional information flow, better long-range dependency handling, robust feature extraction, and reduced information loss.[27–29] Many scholars opt for these predictive models due to their capacity to deliver exceptional performance across a wide range of tasks.

Nonetheless, very few studies employ deep learning to predict or classify problems in sheet pile walls. This study, therefore, conducts a comprehensive investigation into the predicted embedment depth of a cantilever sheet pile wall embedded into cohesive soil with a backfill of cohesionless soil. The following section of this study is structured as follows: The details regarding the sheet piles and relevant data are presented. The methodology and theoretical background of DNN, RNN, LSTM, and Bi-LSTM models are discussed. The obtained results are analyzed and discussed. A summary of the findings and a conclusion are provided.

2. Research significance

This research aims to contribute to the progression of the field of geotechnical engineering by introducing a novel approach for accurately predicting the required sheet pile depths. The integration of advanced AI techniques can optimize the design and construction processes, ultimately resulting in improved stability, cost savings, and increased efficiency in a wide range of construction projects involving sheet piles. The present research holds significance because it has the potential to transform sheet pile depth predictions using advanced artificial intelligence (AI) techniques. By addressing critical gaps in the existing methods. The study contributes to the safety, cost-efficiency, and sustainability of construction practices across construction projects. The findings of this study can pave the way for increased adoption of AI within the realm of geotechnical engineering and stimulate further exploration of AI applications in the broader field of engineering.

2.1 Details of the sheet pile and data

This research thoroughly examined a sheet pile wall embedded within a clay substrate and subsequently backfilled with sand, as depicted in Fig. 1. Following prior investigations, it has been established that the coefficient of variation *(COV)* in clay soils generally varies from 0.1 to 0.5. Concurrently, the *COV* for sandy soils typically exhibits a spectrum within the range of 0.05 to 0.15.[30–33] The 200 data points generated based on the *COV* for the internal friction angle (ϕ) of sand and cohesion (c) of clay are 0.1 and 0.5, respectively. In the context of generating data, the *COV* can be used to control the dispersion or spread of the generated values. By specifying a desired *COV*, a researcher can generate data that has a similar relative variability as an existing dataset or a specific target distribution. Therefore, the unit weight of sand (γ_s) is taken to be between 16.2 and 16.8 kN/m^3, and its internal friction angle (ϕ) is between 25° and 40°.[34] The unit weight of clay (γ_c) varies between 17.2 and 21.8 kN/m^3,[35] and its cohesion (c) varies between 48 and 96 kN/m^2.

Eq. (1) provides the embedment depth (D) for a cantilever sheet pile retained in cohesionless soil while embedded in a cohesive layer.[36,37].

$$D^2(4c - P_a) - 2DA_a - \frac{A_a(12cz_a + A_a)}{2c + P_p} = 0 \qquad (1)$$

where,

A_a = Area of active earth pressure.
P_a = Active earth pressure.
P_p = Passive earth pressure.
z_a = Distance of active earth pressure act from dredge level.

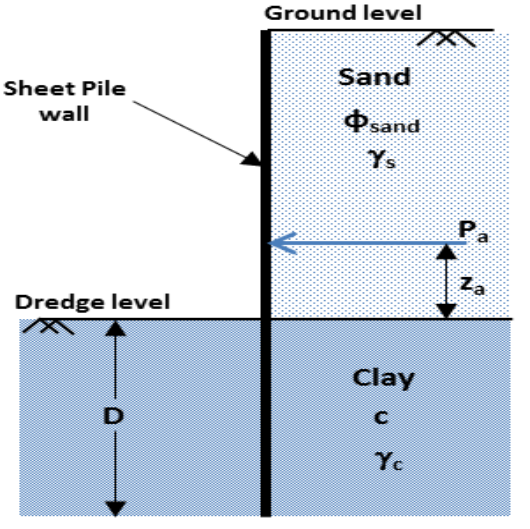


Fig. 1 Sheet pile wall embedded in soil.

2.2 Data preprocessing and statistical description

Table 1 presents the obtained dataset's statistical information before it was subjected to the AI-based analysis process. The

correlation coefficient is a statistical metric ranging between –1 and +1, where –1 indicates a perfect negative correlation, 0 represents no correlation, and +1 signifies a perfect positive correlation. A negative correlation means that as one variable increases, the other variable decreases, while a positive correlation indicates that as one variable increases, the other variable also increases.[38] This negative correlation enhances the balance between depth and cohesion. The correlation matrix is shown in Fig. 2.

Data normalization is an essential preprocessing step performed on the input data, enhancing model performance and stability throughout the learning process. Accurately describing the significance of data normalization in machine learning is highly relevant. This data normalization technique involves scaling the main dataset to a range between 0 and 1, referred to as min–max scaling or min–max normalization, respectively, using Eq. (2).[39–41]

$$X_{normalized} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (2)$$

where X is the original value of a data point in the feature, X_{min} is the minimum value of that feature in the entire dataset, X_{max} is the maximum value of that feature in the entire dataset, and $X_{normalized}$ is the normalized value of the data point.

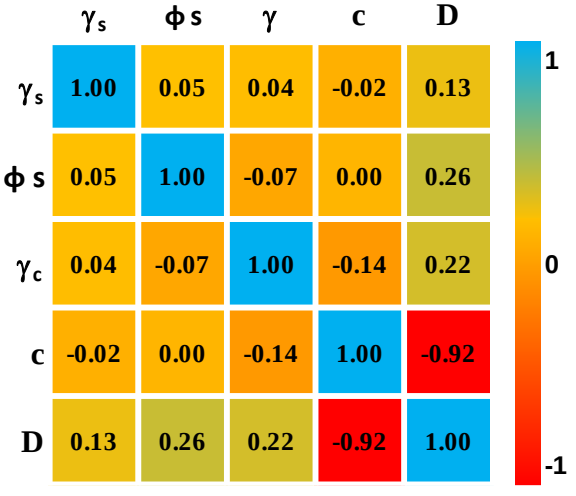


Fig. 2 Correlation matrix of input and output data.

Table 1. Statistical description of dataset.

Statistical Parameter	(γ_s) kN/m^3	(φ)	(γ_c) kN/m^3	(c) kN/m^2	(D) m
Minimum	15.04	25	17.01	48	1.39
1st Quartile	15.7	29	18.34	60	1.71
Median	16.06	34	19.66	71	1.95
Mean	16.09	33.14	19.57	72.15	2.03
3rd Quartile	16.5	37.25	20.9	85	2.3
Maximum	16.98	40	21.86	96	3.07
Standard error	0.03	0.25	0.1	1.4	0.03
Standard deviation	0.5	4.82	1.38	14.46	0.4
Variance	0.25	23.24	1.92	209.34	0.16
Skewness	-0.1	-0.057	-0.1	0.03	0.68
Kurtosis	-0.95	-1.31	-1.23	-1.26	-0.33

3. Theoretical description of the proposed models and research process

The theoretical underpinnings of the DNN, RNN, LSTM, and Bi-LSTM models used are presented in the following subsections.

3.1 Deep Neural Network (DNN)

A DNN is a type of ANN designed to model complex patterns and relationships in data. DNNs comprise multiple layers of interconnected artificial neurons, or nodes, working in tandem to process and transform the input data. In the context of DNNs, the term "deep" refers to the depth of the network, signifying that it encompasses multiple hidden layers between the input and output layers. Within a DNN, each layer of nodes receives input from the preceding layer and transmits its output to the subsequent layer. This hierarchical structure enables DNNs to acquire knowledge and depict progressively abstract features and concepts as they delve deeper into the network. Training a DNN entails feeding a large dataset and iteratively adjusting the connections between its nodes using the backpropagation algorithm. During training, the DNN learns to automatically extract meaningful features from the data and subsequently generates accurate predictions or classifications. The success and popularity of DNNs in recent years can be attributed to several factors, including the ample availability of labeled data, along with advancements in computational power and deep learning algorithms.

3.2 Recurrent Neural Network (RNN)

RNNs excel in tasks that necessitate the consideration of input data order and the dependence of current input on previous inputs. In contrast to feedforward neural networks that process data only in a forward direction without any feedback loops, RNNs have connections that allow information to flow in loops. This recurrent structure enables RNNs to retain memory from past inputs and use it to influence the current output. The basic building block of an RNN is the recurrent unit, which is typically a simple neuron with an additional feedback connection. The most commonly used type of recurrent unit is the LSTM unit, which helps alleviate the vanishing gradient problem that can occur when training RNNs. Another popular variant is the gated recurrent unit (GRU). RNNs have demonstrated their efficacy in various applications, including natural language processing (NLP), speech recognition, machine translation, sentiment analysis, and music generation. They are especially effective in tasks that involve sequential data, such as predicting the next word in a sentence or generating captions for images. However, RNN models have several limitations. Capturing long-range dependencies in sequences is a challenge due to the vanishing gradient problem. This issue arises when the gradients, which are essential for training, exponentially decrease over time.

3.3 Long Short-Term Memory (LSTM)

The LSTM possesses a structure that can retain long-term

memory while also being able to discard unnecessary information from the training data. It is a form of deep learning tailored explicitly for modeling sequential data. The DNN model comprises a singular tangent layer, which is responsible for the model's repeating component. In contrast, the LSTM model consists of four distinct layers that are organized in the manner of a repeating chain. It is an illustration of a single LSTM cell. Within the framework of the LSTM network, this cell is replicated several times. Every LSTM cell has its own unique input signal (x), input gate signal (i), recurrent signal (h), and forget gate signal (f). Additionally, every LSTM cell generates its own output gate signal (o). The LSTM block typically has a traditional structure that is composed of three layers: the input layer, the forget layer, and the output layer. The decision to ignore the information based on the input (X_t) and output (h_{t-1}) is the first stage in the LSTM design. This procedure is carried out with the help of the forget layer (f), and sigmoid activation is utilized for the purpose of serving as the activation function as presented in Eq. (3).

$$f_t = \sigma(W_{f,x} \times X_t + W_{f,h} \times h_{t-1} + b_f) \quad (3)$$

However, the most recent information is found in the cell's input layer in the second stage. The input gate layer (i_t) is updated in this instance by the sigmoid function. Using the Eqs. (4) and (5), the final output data (h_t) are calculated.

$$O_t = \sigma(W_{o,x} \times X_t + W_{o,h} \times h_{t-1} + b_o) \quad (4)$$

$$h_t = O_t \times \tanh(C_t) \quad (5)$$

In the case of the LSTM model, the sigmoid activation function (σ) and the hyperbolic tangent activation function ($tanh$) are commonly used. The training procedure repeats this process over multiple iterations or epochs, gradually refining the model's parameters to improve its performance. As the training progresses, the model learns to better approximate the target output, and the difference between the training data and the LSTM's output is minimized. This minimization is achieved by learning the weighted term W and bias term b.

3.4 Bidirectional Long Short-term Memory (Bi-LSTM)

The Bi-LSTM model is a variant of the LSTM, a type of RNN. In conventional LSTM, the hidden states are solely impacted by the previous time steps. However, Bi-LSTM extends this idea by introducing an additional set of hidden states that are also influenced by the subsequent time steps. This allows the Bi-LSTM model to capture contextual information from the past and future. The Bi-LSTM architecture involves processing the input sequence in two directions: forward, from beginning to end; and backward, from end to beginning. This is achieved by having two separate LSTM layers dedicated to each direction. Each LSTM layer has its own set of weights and biases. In the context of a Bi-LSTM model, the values of y_t, weights, and the conditional probability of states can be computed both before (a^+) and after (a^-) input as presented in Eq. (6).

$$y_t = g\left(W_y[a^{+t}, a^{-t}]\right) + b_y \quad (6)$$

During each time step, the forward LSTM layer receives the input sequence in the original order, whereas the backward LSTM layer receives the input sequence in reverse order. The output of the forward LSTM layer at each time step is connected to the corresponding time step in the backward LSTM layer, and vice versa. This enables the information to flow bidirectionally through the network. Generally, the outputs of the forward and backward LSTM layers are concatenated or combined to produce the final output of the Bi-LSTM model. This combined output contains information from both the previous and subsequent time steps, allowing the model to make predictions based on a more comprehensive understanding of the input sequence. The LSTM and Bi-LSTM-based models are trained using optimizer adaptive moment estimation (ADAM).

3.5 Research process

The present study employed the 200 data points discussed in section 3 to develop advanced regression models to predict the embedment depth (D) of cantilever sheet piles retained in cohesionless or cohesive soil. After visualizing the dataset, the importance of each input parameter was examined using sensitivity analysis, and the correlation between these parameters was determined. The whole dataset was randomly divided into two parts, *i.e.*, training and testing. The training part contains 70% of the whole dataset, equating to 140 data points, while the testing part contains the remaining 30% of the dataset, equivalent to 60 data points. This study involved the development of four different regression models to predict the embedment depth (D) of cantilever sheet piles. Each machine learning model possesses special features and capabilities; the best model was obtained by optimizing the hyperparameters of the respective machine learning models. Finally, a comparative analysis was performed among all the proposed models, validating their simplicity, efficacy, and strong generalizability for the prediction of the embedment depth (D) of cantilever sheet piles. The research methodologies employed in the current study are presented in Fig. 3.

3.6 Performance parameters

The utilization of performance parameters facilitates the assessment of the accuracy and reliability of predictions by comparing actual data with predicted data. The following are several commonly used statistical performance parameters for comparing actual and predicted data. The mathematical equations for the performance parameters, such as the root mean squared error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), performance index (PI), variance account factor (VAF), Willmott's index of agreement (*WI*), Nash Sutcliffe efficiency (*NS*), and weighted mean absolute percentage error (WMAPE), are given below in Eqs. (7) to (14).[42–44]

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(a_i - p_i)^2} \quad (7)$$

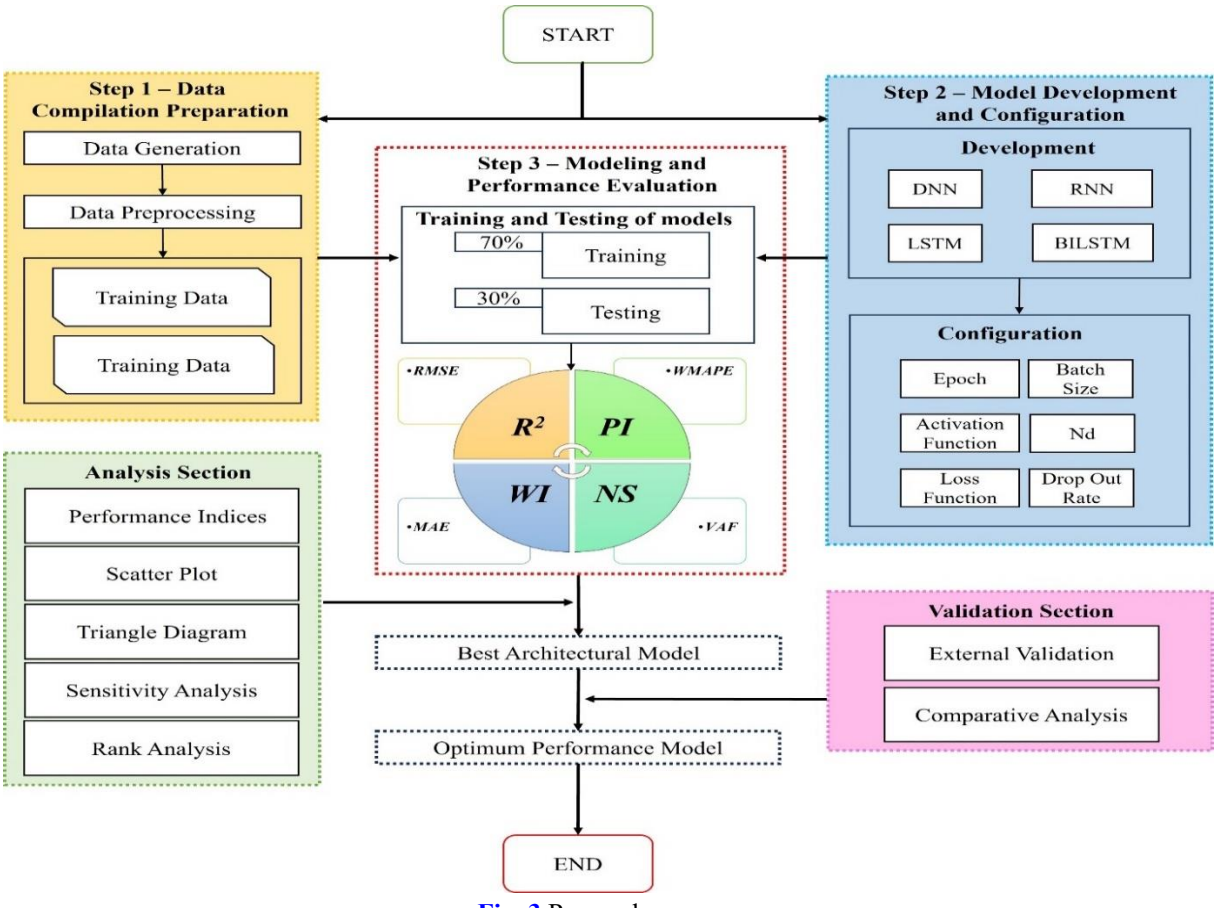


Fig. 3 Research process.

$$VAF = \left(1 - \frac{var(a_i - p_i)}{var(a_i)}\right) \times 100 \quad (8)$$

$$R^2 = \frac{\sum_{i=1}^{N}(a_i - a_{mean})^2 - \sum_{i=1}^{N}(a_i - p_i)^2}{\sum_{i=1}^{N}(a_i - a_{mean})^2} \quad (9)$$

$$PI = adj.R^2 + 0.01VAF - RMSE \quad (10)$$

$$WMAPE = \frac{\sum_{i=1}^{n}\left|\frac{a_i - p_i}{a_i}\right| \times a_i}{\sum_{i=1}^{n} A_i} \quad (11)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|(p_i - a_i)| \quad (12)$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{n}(a_i - p_i)^2}{\sum_{i=1}^{n}\{|p_i - a_{mean}| + |a_i - a_{mean}|\}^2}\right] \quad (13)$$

$$NS = 1 - \frac{\sum_{i=1}^{n}(a_i - p_i)^2}{\sum_{i=1}^{n}(a_i - a_{mean})^2} \quad (14)$$

where d is the average value of the actual data, p_i is the predicted ith value, a_i is the observed ith value, SD is the standard deviation, and N is the total number of data. These statistical metrics provide different perspectives on the accuracy and reliability of the predicted data compared to the actual data. Utilizing multiple metrics is imperative for obtaining a comprehensive understanding of model performance.

3.7 Sensitivity analysis of the parameters affecting the embedment depth (D) of the cantilever sheet pile

Sensitivity analysis is a crucial technique used to determine how variations in input parameters affect the output parameter of a model. It is useful for determining which model input parameters have the most significant impact on the output parameter, allowing for excluding the insignificant parameters. By reducing the input space, both the complexity of the model and the computational time required for training can be decreased.

A sensitivity analysis was conducted for each model in both the training (TR) and testing (TS) phases and the results are shown in Figs. 4 and 5 respectively. The percentage values in Table 2 appear to represent the performance of the models under different input parameter configurations. For example, in the DNN model, during the training phase (TR), when the input parameters vary, the model achieves a sensitivity score of 81% when γ_{sand} is varied, 81% when φ sand is varied, 82% when γ_{clay} is varied, and 51% when c is varied. In the testing phase (TS), the same model (DNN) shows slightly different sensitivities, suggesting how well the model generalizes to new data. For example, it achieves 78% sensitivity when φ sand is varied, 80% when γ_{clay} is varied, 78% when c is varied, and 46% when γ_{sand} is varied. The same analysis is repeated for the RNN, LSTM, and Bi-LSTM models in both the training and testing phases. The main step in this sensitivity analysis is to understand how the model's performance changes with variations in the input parameters. This approach helps identify which input parameters have the most significant impact on the model's performance. In this case, it seems that the γ_{clay} parameter has a relatively strong impact on the model's sensitivity, as it consistently has higher sensitivity values across different models and phases.

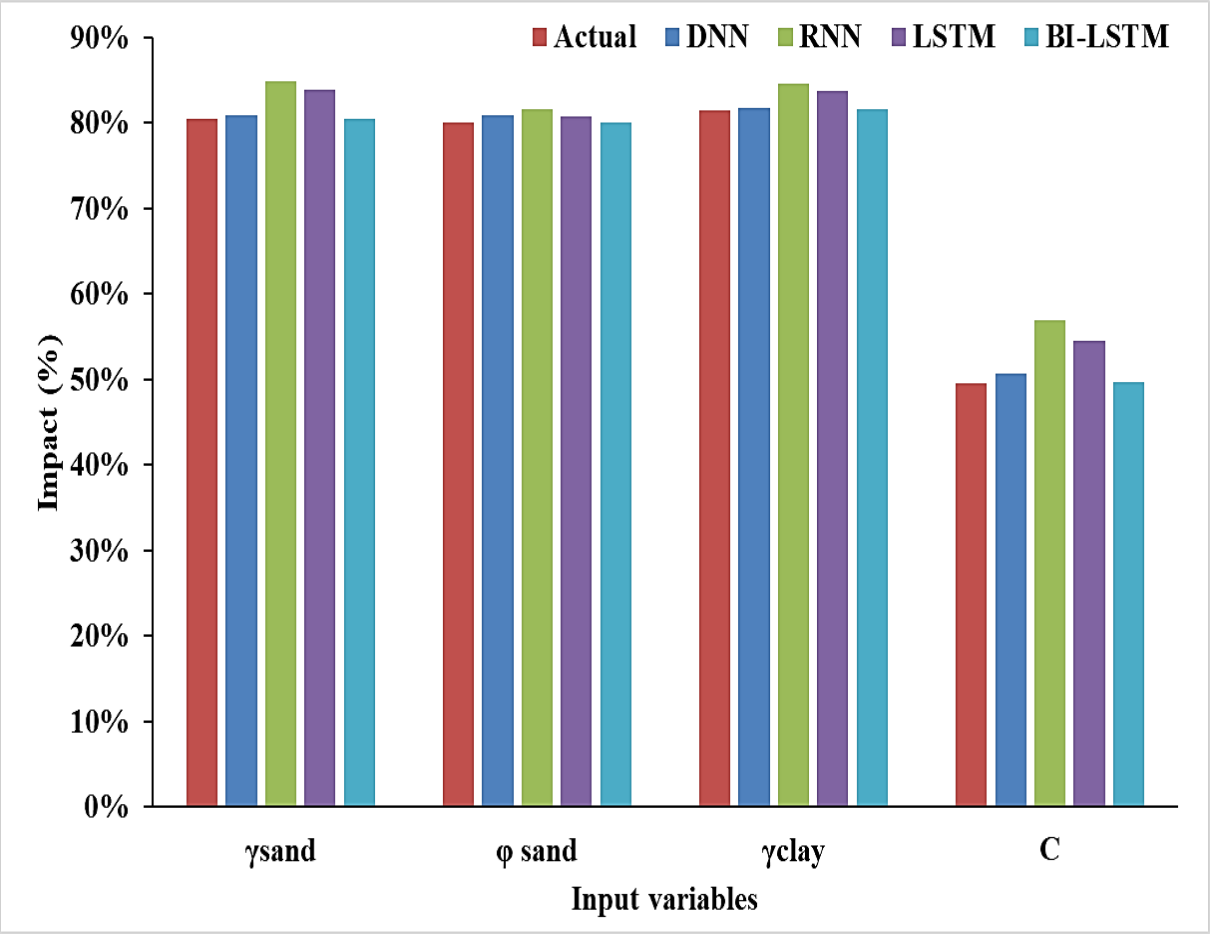


Fig. 4 Strength Parameters for all models in training phase.

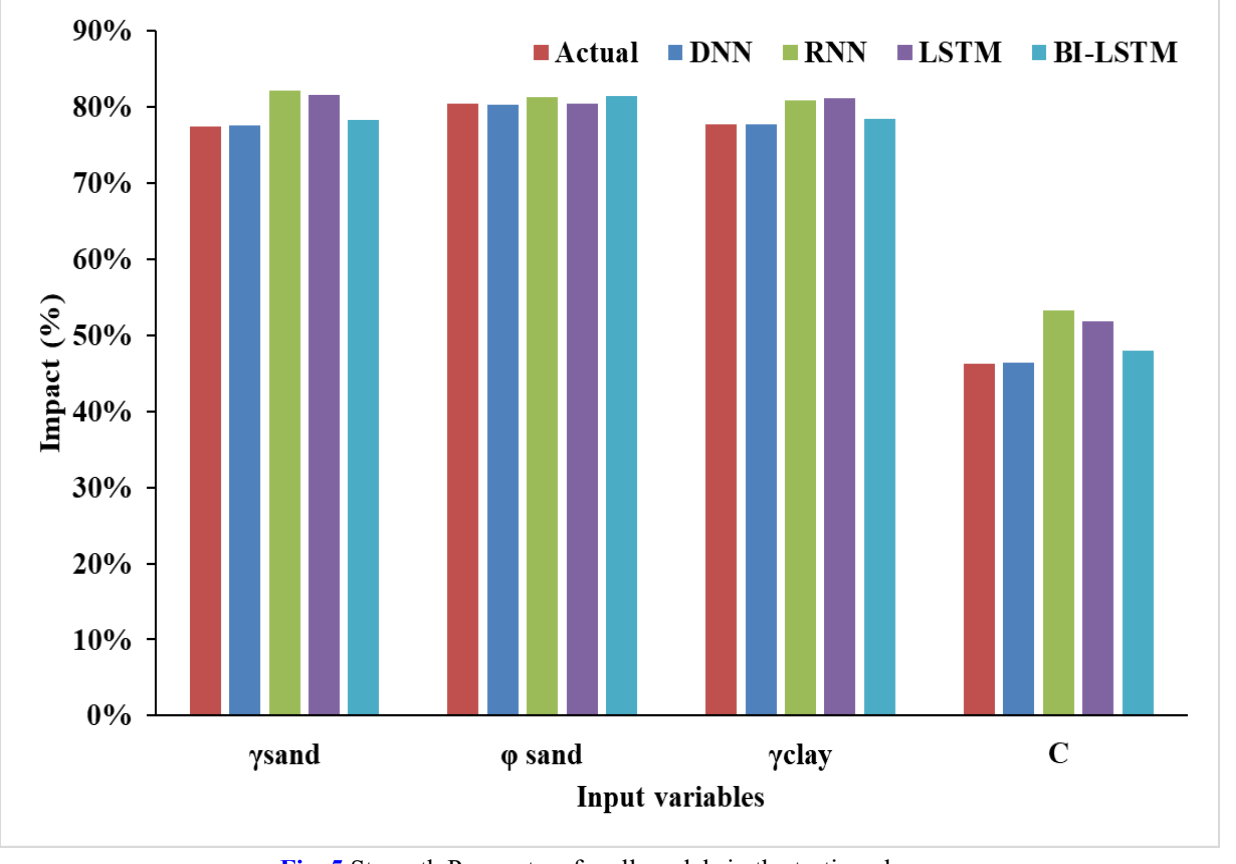


Fig. 5 Strength Parameters for all models in the testing phase.

Table 2. Results of the sensitivity analysis.

Models	Phase	γ_{sand}	φ sand	γ_{clay}	c
Actual	TR	80%	80%	81%	50%
	TS	77%	80%	78%	46%
DNN	TR	81%	81%	82%	51%
	TS	78%	80%	78%	46%
RNN	TR	85%	82%	85%	57%
	TS	82%	81%	81%	53%
LSTM	TR	84%	81%	84%	54%
	TS	82%	80%	81%	52%
BI-LSTM	TR	80%	80%	82%	50%
	TS	78%	81%	78%	48%

4. Results and discussion

4.1 Hyperparameter optimization of the proposed models

In this subsection, the development of the proposed machine learning model and hyperparameter optimization are discussed in detail. It should be noted that the proposed ML models were developed in Python, and numpy and pandas are the two basic libraries used for data preprocessing. Several important libraries, such as keras and TensorFlow, are also used to run the model and compute the desired output. Several important hyperparameters, such as the number of hidden layers, the maximum number of iterations/epochs, batch size, number of dense layers, and dropout rate, were chosen and tuned properly during the model construction. Table 3 represents the optimum hyperparameters that reflect the best model configuration obtained by the trial-and-error method. For all the proposed models, the rectified linear activation function (ReLU) was used, and the performance of the set parameters was calculated using the mean absolute error (MAE). To optimize the objective function, the Adam optimizer is used instead of the classical stochastic gradient descent procedure. The other necessary parameters were set to the default standard values in the original study during the construction of the models.

Table 3. Optimum hyperparameters for DNN, RNN, LSTM, and BI-LSTM.

Hyperparameters	RNN	DNN	LSTM	BILSTM
Number of hidden layers	2	2	2	2
Batch size	50	50	50	50
Number of Dense layers	64	32	64	64
Number of epochs	500	500	500	500
Dropout rate	0.3	0.2	0.3	0.3

4.2 Performance analysis

Table 4 provides the performance parameter values obtained for each model metric for each model, along with their training and testing ranks. After considering the Bi-LSTM, LSTM, and RNN models, the DNN model demonstrated the highest R^2 value of 0.996, an NS of 0.99, a VAF of 99.5, a PI of 1.975, and a WI of 0.999. Table 4 presents the metrics used to assess the performance of each model during both the training (TR) and testing (TS) phases. Each performance parameter is assigned a score ranging from 1 to 4, with 4 representing the highest possible score. The DNN model attained the highest overall score of 64 and secured the top position, indicating its superior performance compared to the other models based on these specific metrics. In contrast, the RNN model achieved a total score of 16, placing it in the fourth position, suggesting that it was the least proficient among the compared models.

Compared to other prediction models, the DNN model emerged as the most accurate and reliable model, attaining the highest ranking with the highest R^2 value in both stages. The Bi-LSTMs, LSTM, and RNN followed in the second, third, and fourth ranks, respectively, in both stages. It is worth noting that all models exhibiting a coefficient of determination exceeding 0.93 demonstrate a substantial degree of confidence.

Table 4. Performance parameter values.

Models	Phase	R^2	WMAPE	NS	RMSE	VAF	PI	WI	MAE	Total Score	Rank
DNN	TR	0.996	0.031	0.995	0.017	99.591	1.975	0.999	0.012		
	Score	4	4	4	4	4	4	4	4		
	TS	0.996	0.032	0.995	0.016	99.548	1.975	0.999	0.012	64	1st
	Score	4	4	4	4	4	4	4	4		
RNN	TR	0.936	0.141	0.895	0.075	90.033	1.759	0.966	0.055		
	Score	1	1	1	1	1	1	1	1		
	TS	0.958	0.136	0.917	0.069	92.216	1.809	0.974	0.050	16	4th
	Score	1	1	1	1	1	1	1	1		
LSTM	TR	0.972	0.095	0.954	0.050	95.456	1.875	0.986	0.037		
	Score	2	2	2	2	2	2	2	2		
	TS	0.975	0.105	0.952	0.052	95.309	1.874	0.986	0.038	32	3rd
	Score	2	2	2	2	2	2	2	2		
BI-LSTM	TR	0.990	0.033	0.989	0.024	98.955	1.955	0.997	0.013		
	Score	3	3	3	3	3	3	3	3		
	TS	0.995	0.037	0.994	0.019	99.422	1.970	0.998	0.014	48	2nd
	Score	3	3	3	3	3	3	3	3		

within the scope of this investigation, it is evident that both the RNN and LSTM models exhibit robust performance, with R^2 values of 0.936 and 0.972, respectively, for the training dataset. Nevertheless, real breakthroughs occur when deploying DNN and Bi-LSTM models, which significantly enhance the overall performance when compared to that of the individual algorithms. The R^2 values for the RNN and BI-LSTM models demonstrate high-performance levels, with values of 0.996 and 0.990 for the training dataset and 0.996 and 0.995 for the testing dataset. Fig. 6 visually depicts the comparison between the predicted and actual values for both the training and testing datasets. Notably, deep learning techniques, particularly the DNN and Bi-LSTM models, exhibit exceptional accuracy in terms of their predictive capabilities. Furthermore, Fig. 6 comprehensively examines the errors associated with predictions generated by the DNN and Bi-LSTM models. This visualization clearly demonstrated the superior predictive accuracy of the deep learning models, as evidenced by the minimal errors observed in the majority of the data points within both the training and testing datasets.

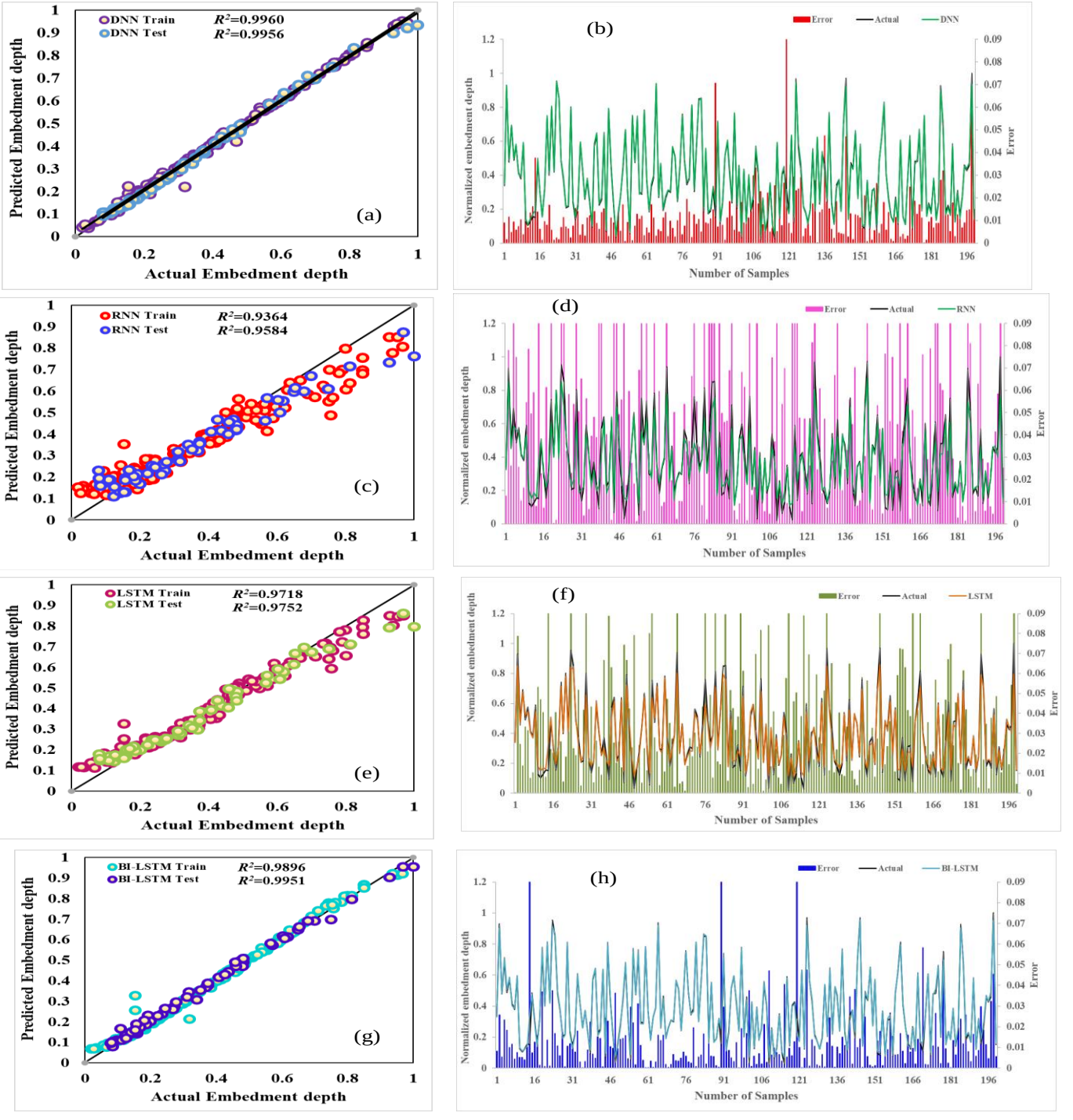


Fig. 6 a) Performance of DNN model, b) Error and line plot of DNN model, c) Performance of RNN model, d) Error and line plot of RNN model, e) Performance of LSTM model, f) Error and line plot of LSTM model, g) Performance of BI-LSTM model, h) Error and line plot of BI-LSTM model.

4.3 Triangle diagram

A triangle diagram is a graphical representation used to evaluate and compare the performance of various datasets or models. This unique visualization tool is particularly effective for assessing the accuracy of predictions generated by models within disciplines such as statistics, machine learning, and data science. Triangle diagrams serve to comprehensively assess and compare datasets or models against a reference dataset. The reference dataset typically represents a baseline or ground truth that the other datasets or models aim to replicate or predict. The diagram is especially valuable when dealing with regression or prediction tasks, where the goal is to understand how well different datasets or models align with the reference data regarding various performance metrics.

4.4 Geometric representation

The triangle diagram itself is depicted as an equilateral triangle, with three distinct axes originating from its vertices. Each axis corresponds to a different performance metric and is inclined at a specific angle, reflecting its unique significance in the assessment process as presented in Fig. 7.

4.5 Components of a triangle diagram

i. Horizontal axis (R-squared or coefficient of determination)

This axis, positioned at the bottom of the diagram, signifies the R-squared (R^2) or coefficient of determination. R^2 measures the overall difference between the datasets or models under evaluation and the reference dataset. This approach provides insights into the goodness of fit and the proportion of variance explained by the models. The scale on this axis ranges from 0, indicating no correlation with the reference dataset, to 1, representing a perfect fit.

ii. Acute angle axis (standard deviation ratio)

The acute angle axis is an extension of the obtuse angle axis and is used to represent the ratio of the standard deviation of the compared datasets or models to the standard deviation of the reference dataset. This axis assesses how well the variability in the compared datasets or models matches the reference dataset. A value of 1 on this axis indicates that the variability of the compared dataset is identical to that observed in the reference dataset.

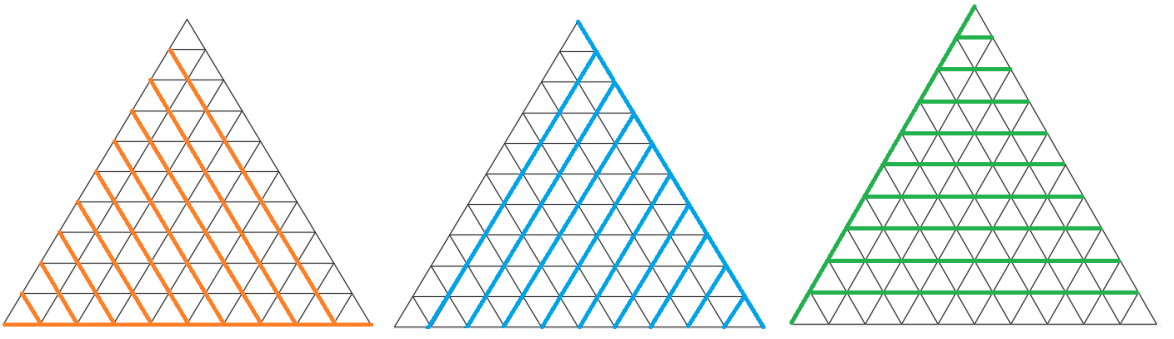
iii. Obtuse angle axis (root-mean-square error: RMSE):

The obtuse angle axis represents the root-mean-square error (RMSE) between the compared datasets or models and the reference dataset. The RMSE quantifies the magnitude of errors in predictions, thereby providing insight into the proximity of these predictions to the actual values. A lower RMSE is indicative of superior model accuracy and precision.

In a triangle diagram, each dataset or model is represented by a line extending from the center of the triangle toward one of its vertices. The angle of each line is precisely 120 degrees concerning the corresponding axis. The position of the line on each axis reflects the value of the respective performance metric for the dataset or model. Researchers and analysts can visually assess and compare the quality of different datasets or models by examining the positioning of their respective lines within the triangle. Datasets or models that perform well across all metrics will have lines that converge toward the center of the triangle, while poorer performers will have lines closer to the triangle's vertices. By simultaneously considering all three metrics, researchers can obtain holistic insight into the degree of alignment between their datasets or models and the reference dataset.

This triangle diagram is a valuable tool for assessing the performance of datasets or models in a holistic manner, taking into account multiple performance metrics. This geometric representation provides a clear visual depiction of the quality of predictions or comparisons, aiding researchers and analysts in making informed decisions and optimizations in various scientific and analytical contexts. Researchers plot datasets or models as points or lines within the triangle diagram. These lines represent the performances of the datasets or models in terms of the selected metrics. When these lines intersect or overlap, they form an area referred to as the "error area." This area represents the region where the datasets or models are evaluated relative to each other. The bottom right corner of the triangle diagram is commonly referred to as the "reference data area." This area represents the ideal or reference dataset or model used for comparing other datasets or models, serving as a benchmark for performance assessment.

Ideally, datasets or models should be positioned close to



(a) Horizontal axis (b) Obtuse angle axis (c) Acute angle axis

Fig. 7 Triangle diagram axis.

the reference data area in terms of correlation. This indicates that there is a robust linear relationship between the dataset/model and the reference data. The standard deviation should also be similar to that of the reference data. Similar standard deviations suggest that the dataset and model exhibit similar levels of variability. A low RMSE, near zero, is considered favorable because it signifies that the predictions of the dataset or model closely align with the reference data.

Deviations from the reference dataset are easily observable in the diagram. When the area representing the datasets or models moves away from the reference area, it indicates differences in performance. Researchers can assess the strengths and weaknesses of datasets or models by analyzing the positions of their lines in a triangle diagram. For example, when a dataset or model exhibits a strong correlation but significantly deviates in terms of standard deviation or has a high RMSE, it reveals specific strengths and weaknesses in its performance. Ideally, the area representing flawless models would perfectly align with the reference area. This signifies that the dataset or model is a superb fit for the reference data in relation to all chosen metrics. In an ideal scenario, the area representing perfect models would coincide with the reference area, indicating that the dataset or model is an excellent match to the reference data concerning all selected metrics.

The triangle diagram is a valuable tool in research papers for visually comparing and evaluating datasets or models based on multiple metrics. Researchers can use these data to assess the performance of these datasets or models in relation to a reference dataset, helping to identify strengths, weaknesses, and areas for improvement. The diagram's visual nature makes it accessible to a broad audience and aids in clearly communicating research findings.

This research paper presents a visual comparison of the performance of various deep learning models for embedded depth prediction. The study evaluated the performances of the RNN, LSTM, Bi-LSTM, and DNN models. The primary focus of this research is to assess and compare the accuracy of these models in terms of their error rates and proximity to a reference area. The results are presented through a visual representation in Figs. 8 and 9, which depict the error areas and proximity of the models to a reference area. Fig. 8 illustrates that the RNN model exhibits a comparatively larger error area, indicating higher prediction errors. The RNN model's performance is far from that of the reference area, suggesting suboptimal predictive accuracy. This observation aligns with the known limitations of standard RNNs in capturing long-term dependencies. LSTM, as shown in Fig. 8, performs better than RNN but still displays a noticeable error area. Although this approach represents an improvement over the basic RNN, it fails to attain optimal accuracy. Bi-LSTM, as depicted in Fig. 9, exhibits further improvement in reducing the error area compared to that of LSTM. The error area appears smaller, suggesting that the bidirectional nature of the model helps capture dependencies in both forward and backward directions, leading to improved predictions. The most promising results were obtained with the DNN model, as illustrated in Fig. 9. It achieves the smallest error area and is closest to the reference area, indicating superior predictive performance. This finding implies that the DNN architecture is well suited for embedded depth prediction tasks, outperforming the RNN, LSTM, and Bi-LSTM models in terms of accuracy.

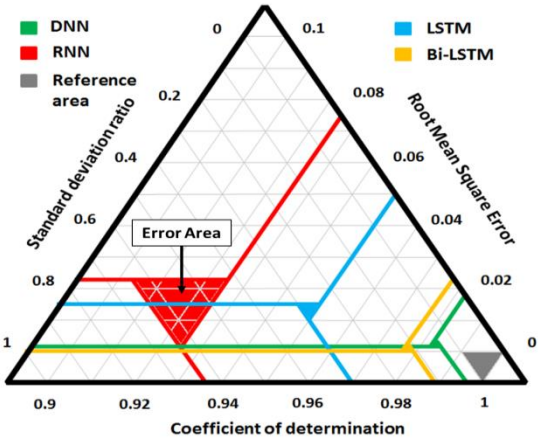


Fig. 8 Triangle diagram for training data.

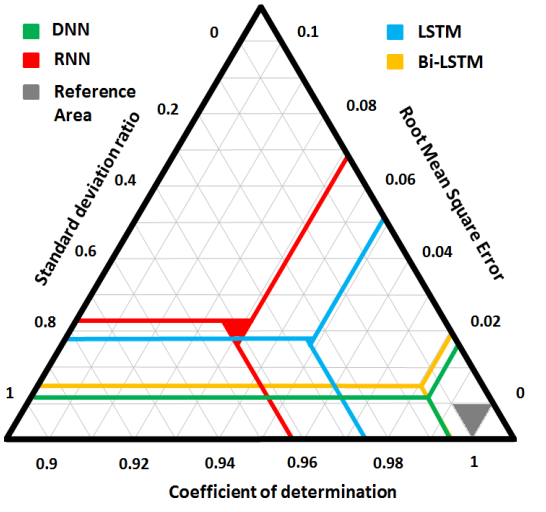


Fig. 9 Triangle diagram for testing data.

4.6 External validation

External validation involves the comparison of actual and predicted values of embedment depth using certain statistical criteria.[45] This process is used to assess the model's generalizability, which refers to its capacity to perform effectively based on its performance on the testing dataset. External validation is important for several reasons. First, it helps to ensure that the model is not simply overfitting the training data. Overfitting occurs when a model learns the training data too well, which hampers its ability to generalize effectively to new data. Second, external validation can help to identify any biases in the model. To assess generalizability, several rigorous statistical penalties were evaluated using the following mathematical expressions presented in Eqs. (15)

and (16):

$$k = \frac{\sum_{i=0}^{n}(P_i \times d_i)}{\sum_{i=0}^{n} P_i^2} \quad (15)$$

$$k' = \frac{\sum_{i=0}^{n}(P_i \times d_i)}{\sum_{i=0}^{n} d_i^2} \quad (16)$$

Where the variables k and k' represent the slope of the regression line passing through the origin of the actual and predicted embedment depths and the predicted and actual embedment depths, respectively. P_{mean} and d_{mean} represent the predicted mean and actual mean of the embedment depth, respectively. The coefficient of determination of the predicted versus the observed embedment depth (*i.e.*, R_o^2) and the observed versus the predicted embedment depth (*i.e.*, R'^2_o) is calculated using Eqs. (17) and (18). On the other hand, stabilization criteria were defined using the variable R_s, which can be calculated using Eq. (19).

$$R_o^2 = 1 - \frac{\sum_{i=1}^{n} P_i^2 (1-k)^2}{\sum_{i=1}^{n} (P_i - P_{mean})^2} \quad (17)$$

$$R'^2_o = 1 - \frac{\sum_{i=1}^{n} d_{o,i}^2 (1-k\prime)^2}{\sum_{i=1}^{n} (d_i - d_{mean})^2} \quad (18)$$

$$R_s = R^2 \times \left(1 - \sqrt{|R^2 - R_o^2|}\right) \quad (19)$$

According to the external validation criteria, any model can be accepted as generalizable and robust if it meets at least one of the two following criteria, as mentioned below. A machine learning model is considered fully acceptable and generalizable if it meets all four criteria:[46] (i) $0.85 \leq k \text{ or } k' \leq 1.15$ (ii) $R^2 \geq 0.6$; (iii) R_o^2 or R'^2_o should be close to R^2, (*i.e.*, the ratio of the absolute difference of R^2 from R_o^2 or R'^2_o to R^2 must be less than 0.1, and (iv) $R_s \geq 0.5$. The results of the external validation of all the proposed models for the testing phase are summarized in Table 5. From the presented results, it is evident that all the proposed models meet all four external validation criteria. However, the DNN model obtained the best performance among all four proposed models.

Table 5. Results of external validation.

	Criteria						Condition			
	R^2	K	K'	R'^2	R''^2	R_m	1^{st}	2^{nd}	3^{rd}	4^{th}
DNN	0.996	1.01	0.99	1.00	1.00	0.993	✓	✓	✓	✓
RNN	0.958	1.09	0.90	0.96	0.97	0.900	✓	✓	✓	✓
LSTM	0.975	1.06	0.93	0.99	0.99	0.888	✓	✓	✓	✓
BI-LSTM	0.995	1.00	1.00	1.00	1.00	0.993	✓	✓	✓	✓

Comparative analysis

An artificial neural network (ANN) is utilized as a prediction model to determine the embedment depth of cantilever sheet pile walls.[47] Soil characteristics such as cohesiveness (c), angle of internal friction (φ), and unit weight (γ) are the input factors for this prediction. It uses an ANN hybrid model with various optimization techniques, such as teaching-learning-based optimization, artificial bee colony optimization, ant colony optimization, ant lion optimization, imperialist optimization, the imperialist competitive algorithm, and shuffled complex evolution. In this study, the first-order reliability method (FORM) is employed to assess the reliability of cantilever sheet pile walls. Various metrics, including the reliability index, rank analysis, Taylor diagram, accuracy matrix, DDR criteria, AIC, and OBJ criterion, are utilized to assess the performance of the prediction models. The study concluded that hybrid models combining ANNs and teaching-learning-based optimization (ANN-TLO) are more effective at predicting the dependability of cantilever sheet pile walls based on experimental data. In this section, a comparative study was conducted using the same data input with respect to the current models displayed in Table 6. In this comparison, the DNN from the current study demonstrated superior performance to the other models, as evidenced by higher R^2 and lower RMSE values.

Table 6. Comparisons of the previous study and the present study.

Model	R^2		RMSE		Reference
	Training	Testing	Training	Testing	
ANN-ABC	0.777	0.686	0.114	0.144	[47]
ANN-ACO	0.945	0.921	0.053	0.073	
ANN-ALO	0.958	0.938	0.046	0.064	
ANN-ICA	0.990	0.984	0.022	0.035	
ANN-SCE	0.991	0.987	0.021	0.030	
ANN-TLO	0.991	0.987	0.021	0.029	
DNN	0.996	0.996	0.017	0.016	Present study
RNN	0.936	0.958	0.075	0.069	
LSTM	0.972	0.975	0.050	0.052	
Bi-LSTM	0.990	0.995	0.024	0.019	

5. Summary and conclusion

This section details the development and hyperparameter optimization of the proposed machine learning models for estimating the embedment depth of cantilever sheet pile walls. The models were developed using Python, with essential data preprocessing performed with the numpy and pandas libraries. For model execution and output computation, libraries such as Keras and TensorFlow were employed. All the models utilized the rectified linear activation function (ReLU), and their performances were assessed using the mean absolute error (MAE), with the "Adam" optimizer employed for objective function optimization.

The performance parameters were summarized for each model in both the training and testing phases. The DNN model exhibited the highest R^2 value of 0.996 and outperformed the

other models in various performance metrics, achieving first place in both stages. The RNN and LSTM models demonstrated strong performance during the training phase, with R^2 values of 0.936 and 0.972, respectively. However, actual advancements occurred when the DNN and Bi-LSTM models were implemented, which achieved exceptional R^2 values of 0.996 and 0.990 in the training phase and 0.996 and 0.995 in the testing phase, respectively. This study visually represents the predicted versus measured values for the training and testing datasets, highlighting the accuracy of the predictions, particularly for the DNN and Bi-LSTM models.

A triangle diagram provides a comprehensive comparison of the models, highlighting the advantages and disadvantages of each model. Evidently, the DNN model outperformed the others, achieving a smaller error area and closer proximity to the reference area. External validation, an essential step in assessing the model's generalizability, confirmed that all proposed models met rigorous statistical criteria. However, the DNN model demonstrated superior performance in this evaluation, further emphasizing its robustness. Finally, compared to the findings of a prior study by Pradeep *et al.* (2022), the DNN model in the present study outperformed other models, including hybrid models incorporating ANNs based on R^2 and RMSE.

In conclusion, the proposed DNN model demonstrated outstanding performance in estimating the embedment depth of cantilever sheet pile walls, surpassing other models and exhibiting high accuracy, reliability, and generalizability. This study's comprehensive approach, including model development, hyperparameter tuning, performance assessment, and external validation, contributes significantly to the field of geotechnical engineering and predictive modeling. However, there are several limitations, including its focus solely on sandy soil overlaying clay soil, inadequate data quality, and limited sample size. Future research can address these limitations by employing advanced data-cleaning techniques, implementing more rigorous feature selection methods, and integrating practical data sources. These measures can effectively overcome the existing constraints and enhance the model's overall performance.

Acknowledgment

This work was supported by Thammasat University Research Unit in Data Science and Digital Transformation.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

References

[1] S. Jiang, C. Du, L. Sun, Numerical analysis of sheet pile wall structure considering soil-structure interaction, *Geomechanics and Engineering*, 2018, **16**, 309-320, doi: 10.12989/GAE.2018.16.3.309.

[2] S. E. N. Ingenier, Título: Comparison of numerical and classical analytical method for sheet pile wall analyses, MÁSTER EN INGENIERÍA GEOLÓGICA, 2014.

[3] Q. Yan, X. Yan, Numerical simulation research and use of the steel sheet pile supporting structure in vertical excavation, Numerical Analysis-Theory and Application, IntechOpen, 201, doi: 10.5772/intechopen.84032.

[4] L. Eskandari, B. Kalantari, Basic types of sheet pile walls and their application in the construction industry-a review, *Electronic Journal of Geotechnical Engineering*, 2011, **16**, 1533-1541

[5] S. Roy, S. Kumar Bhalla, Role of Geotechnical Properties of Soil on Civil Engineering Structures, *Resources and Environment*, 2017, **7**, 103-109, doi: 10.5923/j.re.20170704.03.

[6] F. Deckner, K. Viking, S. Hintze, Wave patterns in the ground: case studies related to vibratory sheet pile driving, *Geotechnical and Geological Engineering*, 2017, **35**, 2863-2878, doi: 10.1007/s10706-017-0285-x.

[7] G. J. W. King, Analysis of cantilever sheet-pile walls in cohesionless soil, *Journal of Geotechnical Engineering*, 1995, **121**, 629-635, doi: 10.1061/(asce)0733-9410(1995)121: 9(629).

[8] S. Gajan, Normalized relationships for depth of embedment of sheet pile walls and soldier pile walls in cohesionless soils, *Soils and Foundations*, 2011, **51**, 559-564, doi: 10.3208/sandf.51.559.

[9] A. G. Ray, D. K. Baidya, Reliability-based analysis of cantilever sheet pile walls backfilled with different soil types using the finite-element approach, *International Journal of Geomechanics*, 2015, **15**, 6015001, doi: 10.1061/(asce)gm.1943-5622.0000475.

[10] K. Terzaghi, Stress distribution in dry and in saturated sand above a yielding trap-door, Proceedings of the International Conference of Soil Mechanics, Harvard University, Cambridge, USA, 1936, 307-311.

[11] W. J. M. Rankine, II On the Stability of Loose Earth, *Philosophical Transaction of the Royal Society of London*, 1857, **147**, 9-27, doi: 10.1098/rstl.1857.0003.

[12] H. Patil, R. Deshmukh, P. J. Salunke, Analysis of sheet pile in dense and loose soil using finite element method, *Materials Today: Proceedings*, 2023, **77**, 654-661, doi: 10.1016/j.matpr.2022.11.285.

[13] A. Debnath, S. K. Pal, A numerical analysis on anchored sheet pile wall subjected to surcharge strip loading, *Journal of Engineering Research*, 2023, **11**, 62-74, doi: 10.1016/j.jer.2023.100088.

[14] S. Chen, Y. Guan, J. Dai, Behaviour of anchored sheet pile quay stabilized with deep cement mixing columns in soft soil: Centrifuge and numerical modelling, *Computers and Geotechnics*, 2023, **160**, 105504, doi: 10.1016/j.compgeo.2023.105504.

[15] N. Kumar, Prabhansu, Artificial intelligence and machine learning applications in energy storage system: technology overview and perspectives, Emerging Trends in Energy Storage Systems and Industrial Applications. Amsterdam: Elsevier, 2023,

1-26, doi: 10.1016/b978-0-323-90521-3.00014-4.
[16] S. O. Abioye, L. O. Oyedele, L. Akanbi, A. Ajayi, J. M. Davila Delgado, M. Bilal, O. O. Akinade, A. Ahmed, Artificial intelligence in the construction industry: a review of present status, opportunities and future challenges, *Journal of Building Engineering*, 2021, **44**, 103299, doi: 10.1016/j.jobe.2021.103299.
[17] I. H. Sarker, Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions, *SN Computer Science*, 2021, **2**, 420, doi: 10.1007/s42979-021-00815-1.
[18] M. Soori, B. Arezoo, R. Dastres, Artificial intelligence, machine learning and deep learning in advanced robotics, a review, *Cognitive Robotics*, 2023, **3**, 54-70, doi: 10.1016/j.cogr.2023.04.001.
[19] I. H. Sarker, AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems, *SN Computer Science*, 2022, **3**, 158, doi: 10.1007/s42979-022-01043-x.
[20] K. Bayoudh, R. Knani, F. Hamdaoui, A. Mtibaa, A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets, *The Visual Computer*, 2022, **38**, 2939-2970, doi: 10.1007/s00371-021-02166-7.
[21] M. E. Paoletti, J. M. Haut, J. Plaza, A. Plaza, Deep learning classifiers for hyperspectral imaging: a review, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, **158**, 279-317, doi: 10.1016/j.isprsjprs.2019.09.006.
[22] A. Sherstinsky, Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network, *Physica D: Nonlinear Phenomena*, 2020, **404**, 132306, doi: 10.1016/j.physd.2019.132306.
[23] F. Harrou, A. Zeroual, M. M. Hittawe, Y. Sun, Recurrent and convolutional neural networks for traffic management. Road Traffic Modeling and Management. Amsterdam: Elsevier, 2022, 197-246, doi: 10.1016/b978-0-12-823432-7.00011-2.
[24] P. B. Weerakody, K. W. Wong, G. Wang, W. Ela, A review of irregular time series data handling with gated recurrent neural networks, *Neurocomputing*, 2021, **441**, 161-178, doi: 10.1016/j.neucom.2021.02.046.
[25] R. Bu, Z. Ren, H. Ge, J. Chen, Online NBTI-induced partially depleted (PD) SOI degradation and recovery prediction utilizing long short-term memory (LSTM), *Microelectronics Reliability*, 2023, **142**, 114932, doi: 10.1016/j.microrel.2023.114932.
[26] B. Xu, C. K. Pooi, K. M. Tan, S. Huang, X. Shi, H. Y. Ng, A novel long short-term memory artificial neural network (LSTM)-based soft-sensor to monitor and forecast wastewater treatment performance, *Journal of Water Process Engineering*, 2023, **54**, 104041, doi: 10.1016/j.jwpe.2023.104041.
[27] Y. Yu, Y. Yao, Z. Liu, Z. An, B. Chen, L. Chen, R. Chen, A Bi-LSTM approach for modelling movement uncertainty of crowdsourced human trajectories under complex urban environments, *International Journal of Applied Earth Observation and Geoinformation*, 2023, **122**, 103412, doi: 10.1016/j.jag.2023.103412.
[28] W. F. Alfwzan, M. M. Selim, S. Althobaiti, A. M. Hussin, Application of Bi-LSTM method for groundwater quality assessment through water quality indices, *Journal of Water Process Engineering*, 2023, **53**, 103889, doi: 10.1016/j.jwpe.2023.103889.
[29] J. Yin, Z. Deng, A. V.-M. Ines, J. Wu, E. Rasu, Forecast of short-term daily reference evapotranspiration under limited meteorological variables using a hybrid bi-directional long short-term memory model (Bi-LSTM), *Agricultural Water Management*, 2020, **242**, 106386, doi: 10.1016/j.agwat.2020.106386.
[30] K.-K. Phoon, F. H. Kulhawy, Characterization of geotechnical variability, *Canadian Geotechnical Journal*, 1999, **36**, 612-624, doi: 10.1139/cgj-36-4-612.
[31] N. Alamanis, Uncertainties and optimization in geotechnical engineering, *American Scientific Research Journal for Engineering, Technology, and Sciences*, 2017, **38**, 92-111.
[32] Y. Wu, X. Zhou, Y. Gao, L. Zhang, J. Yang, Effect of soil variability on bearing capacity accounting for non-stationary characteristics of undrained shear strength, *Computers and Geotechnics*, 2019, **110**, 199-210, doi: 10.1016/j.compgeo.2019.02.003.
[33] M. Kawa, I. Bagińska, M. Wyjadłowski, The reliability analysis of sheet pile wall located in soil with random properties based on cptu results, *Engineering Transactions*, 2017, **65**, 193-200, doi: 10.24423/engtrans.745.2017.
[34] G. Acharya, T. A. Cochrane, T. Davies, E. Bowman, The influence of shallow landslides on sediment supply: a flume-based investigation using sandy soil, *Engineering Geology*, 2009, **109**, 161-169, doi: 10.1016/j.enggeo.2009.06.008.
[35] C. H. Benson, H. Zhai, X. Wang, Estimating hydraulic conductivity of compacted clay liners, *Journal of Geotechnical Engineering*, 1994, **120**, 366-387, doi: 10.1061/(asce)0733-9410(1994)120: 2(366).
[36] A. A. Basma, Reliability-based design of sheet pile structures, *Reliability Engineering & System Safety*, 1991, **33**, 215-230, doi: 10.1016/0951-8320(91)90060-k.
[37] K. R. Arora, Soil Mechanics and Foundation Engineering (Geotechnical Engineering): In SI Units; *Standard publishers*, 2008, ISBN 8180141128.
[38] H. Akoglu, User's guide to correlation coefficients, *Turkish Journal of Emergency Medicine*, 2018, **18**, 91-93, doi: 10.1016/j.tjem.2018.08.001.
[39] D. Singh, B. Singh, Investigating the impact of data normalization on classification performance, *Applied Soft Computing*, 2020, **97**, 105524, doi: 10.1016/j.asoc.2019.105524.
[40] D. R. Kumar, P. Samui, W. Wipulanusat, S. Keawsawasvong, K. Sangjinda, W. Jitchaijaroen, Soft computing techniques for predicting penetration and uplift resistances of dual pipelines in cohesive soils, *Engineered Science*, 2023, **24**, 897, doi: 10.30919/es897.
[41] D. R. Kumar, P. Samui, W. Wipulanusat, S. Keawsawasvong, K. Sangjinda, W. Jitchaijaroen, Machine learning approaches for prediction of the bearing capacity of ring foundations on rock

masses, *Earth Science Informatics*, 2023, **16**, 4153-4168, doi: 10.1007/s12145-023-01152-y.
[42] D. R. Kumar, P. Samui, W. Wipulanusat, S. Keawsawasvong, K. Sangjinda, W. Jitchaijaroen, Bearing capacity of eccentrically loaded footings on rock masses using soft computing techniques, *Engineered Science*, 2023, **24**, 929, doi: 10.30919/es929.
[43] M. Kumar, R. Biswas, D. R. Kumar, P. Samui, M. R. Kaloop, M. Eldessouki, Soft computing-based prediction models for compressive strength of concrete, *Case Studies in Construction Materials*, 2023, **19**, e02321, doi: 10.1016/j.cscm.2023.e02321.
[44] R. Kumar, A. Kumar, D. Ranjan Kumar, Buckling response of CNT based hybrid FG plates using finite element method and machine learning method, *Composite Structures*, 2023, **319**, 117204, doi: 10.1016/j.compstruct.2023.117204.
[45] A. Golbraikh, A. Tropsha, Beware of Q2!, *Journal of Molecular Graphics and Modelling*, 2002, **20**, 269-276, doi: 10.1016/s1093-3263(01)00123-1.
[46] W. Jitchaijaroen, S. Keawsawasvong, W. Wipulanusat, D. R. Kumar, P. Jamsawang, J. Sunkpho, Machine learning approaches for stability prediction of rectangular tunnels in natural clays based on MLP and RBF neural networks, *Intelligent Systems with Applications*, 2024, **21**, 200329, doi: 10.1016/j.iswa.2024.200329.