

Article

Soft-Computing Techniques for Predicting Seismic Bearing Capacity of Strip Footings in Slopes

Divesh Ranjan Kumar ¹, Pijush Samui ¹, Warit Wipulanusat ^{2,*}, Suraparb Keawsawasvong ²,
Kongtawan Sangjinda ² and Wittaya Jitchaijaroen ²

¹ Department of Civil Engineering, National Institute of Technology Patna, Patna 800005, India; diveshk.ph21.ce@nitp.ac.in (D.R.K.); pijush@nitp.ac.in (P.S.)

² Department of Civil Engineering, Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani 12120, Thailand; ksurapar@engr.tu.ac.th (S.K.); kongtawan.sang@dome.tu.ac.th (K.S.); wittaya.jitc@dome.tu.ac.th (W.J.)

* Correspondence: wwarit@engr.tu.ac.th

Abstract: In this study, various machine learning algorithms, including the minimax probability machine regression (MPMR), functional network (FN), convolutional neural network (CNN), recurrent neural network (RNN), and group method of data handling (GMDH) models, are proposed for the estimation of the seismic bearing capacity factor (N_c) of strip footings on sloping ground under seismic events. To train and test the proposed machine learning model, a total of 1296 samples were numerically obtained by performing a lower-bound (LB) and upper-bound (UB) finite element limit analysis (FELA) to evaluate the seismic bearing capacity factor (N_c) of strip footings. Sensitivity analysis was performed on all dimensionless input parameters (i.e., slope inclination (β); normalized depth (D/B); normalized distance (L/B); normalized slope height (H/B); the strength ratio ($c_u/\gamma B$); and the horizontal seismic acceleration (k_h)) to determine the influence on the dimensionless output parameters (i.e., the seismic bearing capacity factor (N_c)). To assess the performance of the proposed models, various performance parameters—namely the coefficient of determination (R^2), variance account factor (VAF), performance index (PI), Willmott's index of agreement (WI), the mean absolute error (MAE), the weighted mean absolute percentage error (WMAPE), the mean bias error (MBE), and the root-mean-square error (RMSE)—were calculated. The predictive performance of all proposed models for a bearing capacity factor (N_c) prediction was compared by using the testing dataset, and it was found that the MPMR model achieved the highest R^2 values of 1.000 and 0.957 and the lowest RMSE values of 0.000 and 0.038 in both the training and testing phases, respectively. The parametric analyses, rank analyses, REC curves, and the AIC showed that the proposed models were quite effective and reliable for the estimation of the bearing capacity factor (N_c).

Keywords: bearing capacity; slopes; strip footing; soft computing; MPMR; GMDH



Citation: Kumar, D.R.; Samui, P.; Wipulanusat, W.; Keawsawasvong, S.; Sangjinda, K.; Jitchaijaroen, W. Soft-Computing Techniques for Predicting Seismic Bearing Capacity of Strip Footings in Slopes. *Buildings* **2023**, *13*, 1371. <https://doi.org/10.3390/buildings13061371>

Academic Editor: Harry Far

Received: 25 March 2023

Revised: 18 May 2023

Accepted: 22 May 2023

Published: 24 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One of the most common geotechnical issues with strip footings on slopes is their stability. Many scholars have investigated the stability of this problem using a variety of techniques, such as semi-empirical methods (e.g., Hansen [1]; Satvati et al. [2]; Khalvati et al. [3]; finite element methods (e.g., Georgiadis [4,5]); limit equilibrium techniques (e.g., Meyerhof [6]); and limit analysis (e.g., Davis and Booker [7]; Kusakabe et al. [8]; Shiau et al. [9]; Georgiadis [10]). However, these works did not consider the seismic body force on the overall stability of footings on slopes. Indeed, the seismic bearing capacity problem is essential for understanding the nature of earthquake zones. To capture this seismic effect, the pseudo-static technique is a convenient technique. The horizontal and vertical seismic coefficients (k_h and k_v) are defined as a function of gravity acceleration. These coefficients have been considered a common and extensively applied method for assessing the stability of footings in earthquake zones.

Previous studies have determined the seismic bearing capacity of footings on slopes regarding pseudo-static seismic forces. Many methods have been used to compute seismic bearing capacity solutions, such as the limit equilibrium method (e.g., Budhu and Al-Karni [11]; Kumar and Mohan [12]) and the limit analysis method (e.g., Farzaneh et al. [13]; Kumar and Ghosh [14]; Yamamoto [15]; Georgiadis and Chrysouli [16]). One of the numerical techniques that can provide rigorous plastic solutions to stability problems is the finite element limit analysis (FELA) method, which was successfully employed by Kumar and Chakraborty [17] to provide the seismic bearing capacity factor for strip footings on cohesionless slopes. In addition, Chakraborty and Kumar [18] and Chakraborty and Mahesh [19] also used the FELA technique to determine the strip footings' seismic bearing capacities on sloping ground under seismic events. Recently, Luo et al. [20] and Lai et al. [21] also carried out the FELA method to investigate the seismic bearing capacity of strip footings on cohesive soils.

Nevertheless, the MPMR, FN, CNN, RNN, and GMDH models have not yet been applied, based on the findings of the authors' literature review, to estimate the bearing capacity factor (N_c) of strip footings on sloping ground under seismic events; meanwhile, these methods have been widely used to predict the nonlinear behavior of engineering problems. For instance, the MPMR model was used to estimate the rock strain, the axial capacity of bored piles, the uplift capacity of a suction caisson, etc., and the results showed that the predictive performance of the MPMR model was high [22–24]. The FN and GMDH methods were applied to estimate the axial capacity of bored piles and the settlement of pile groups in clay [23,25]. Many of the applications of the CNN and RNN methods, which were applied to predict the accurate and desired target values, have been identified in the literature [26–29].

Taking these considerations into account, this study implements five advanced machine learning models to evaluate the seismic bearing capacity of strip footings on sloping ground under seismic events. The strip footings in clay with inclination (β) and height (H) are taken as a problem statement, as shown in Figure 1. Machine learning models—namely the minimax probability machine regression (MPMR), functional network (FN), convolutional neural network (CNN), recurrent neural network (RNN), and the group method of data handling (GMDH) model—were constructed, analyzed, and discussed when using artificial datasets that were generated from the FELA method. The prediction performance of the proposed models, i.e., MPMR, FN, CNN, RNN, and GMDH, were thoroughly examined in terms of eight statistical parameters, score analyses, sensitivity analyses, and regression error characteristic (REC) curves to identify the best-performing models. In this study, the authors propose the geotechnical practitioners' models that are the quickest and easiest to use. These models are based on software, and require only a basic understanding of computer programming to determine the strip footings' seismic bearing capacity on sloping ground under seismic events. Addressing the geotechnical engineers' challenges for finding the seismic bearing capacity of strip footings on sloping ground under seismic events is a challenging task. The proposed advanced ML models have many practical implications in the seismic design of soil structures in geotechnical engineering.

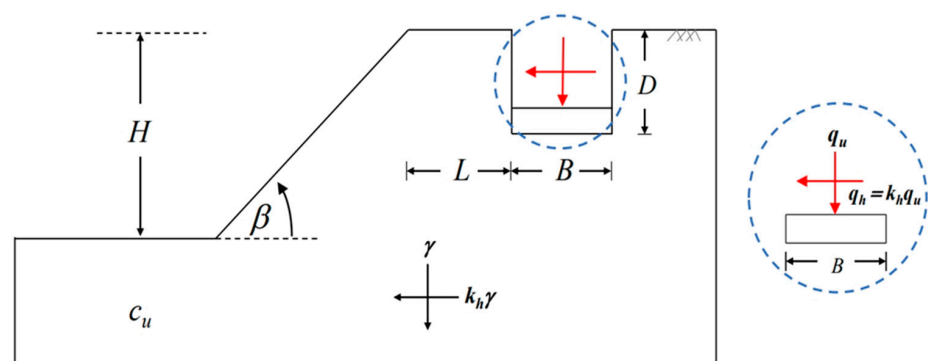


Figure 1. The problem definition of a strip footing on a slope.

2. Data Collection

Figure 1 depicts the problem definition for a strip footing on a slope, where β denotes the slope inclination; H denotes the slope height; B denotes the width of the footing; D denotes the depth of the footing; and L denotes the distance from the top of the slope to the edge of the footing. The soil is cohesive with a unit weight (γ) and undrained shear strength (c_u). This study employs the pseudo-static approach, which is a simplified method used in earthquake engineering to estimate the seismic forces acting on structures or underground works. In this approach, the earthquake-induced ground motion is simplified to a static force that acts on the structure. The static force is calculated by multiplying the seismic coefficient (a factor that depends on the seismic hazard and the characteristics of the soil) by the weight of the structure. The resulting force is then applied to the structure as a static load. By applying this approach, we assumed that both the footing and slope are subjected to horizontal seismic acceleration. The horizontal body force is set to $k_h\gamma$, where k_h is the horizontal seismic coefficient. The footing is set to be subjected to q_u in the vertical direction, and $k_h q_u$ in the horizontal direction, where q_u is the ultimate bearing capacity. Note that, in this study, the vertical seismic coefficient (k_v) is neglected since it has little effect on the stability of structures or underground works. In most earthquakes, the vertical ground motion is usually smaller than the horizontal motion, and its frequency content is different from that of the horizontal motion. Therefore, it is commonly assumed that the vertical seismic coefficient can be neglected without a significant loss of accuracy in estimating the seismic forces acting on the structure or underground work. More details on this problem can be found in Lai et al. [21].

Based on Lai et al. [21], the seismic bearing capacity (N_c) can be expressed as a function of six dimensionless parameters as follows:

$$N_c = \frac{q_u}{c_u} = f\left(\beta, \frac{H}{B}, \frac{L}{B}, \frac{D}{B}, \frac{c_u}{\gamma B}, k_h\right) \quad (1)$$

Note that the seismic bearing capacity factor (N_c) in this study is the ultimate vertical bearing capacity of the footing divided by the undrained shear strength of soil, which is similar to the classic Terzaghi's bearing capacity factors. The other dimensionless parameters are D/B , the normalized depth; L/B , the normalized distance; H/B , the normalized slope height; and $c_u/\gamma B$, the strength ratio. The selected ranges of these six dimensionless inputs are shown in Table 1 and are according to Lai et al. [21].

Table 1. List of parametric values used in this study.

Input Parameters	Selected Values
β	15°, 30°, 45°, 60°
H/B	1, 2, 4
L/B	0, 1, 2, 4
D/B	0, 1, 2
$c_u/\gamma B$	1.5, 2.5, 5
k_h	0.1, 0.2, 0.3

According to Lai et al. [21], FELA techniques were employed to determine the lower bound (LB) and upper bound (UB) solutions of N_c , where OptumG2 [30]—which is a type of FELA software—was carried out to obtain all numerical results. Note that an automatically adaptable mesh refinement technique [31], which was used to enhance the precision of the upper and lower solutions, was also utilized in the study by Lai et al. [21]. The setting of the mesh refinement was set to automatically develop from 5000 to 10,000 elements throughout the 5 adaptive meshing iterations by following certain previous studies (e.g., Keawsawasvong and Ukritchon [32,33]; Shiau et al. [34]; Keawsawasvong et al. [35]). Note that the difference between the UB and LB solutions was within 3% for all numerical results. An example of the OptumG2 model of the footing on a slope is shown in Figure 2, where

the potential slip surface can be clearly determined by using the FELA method with the automatically adaptable mesh refinement technique.

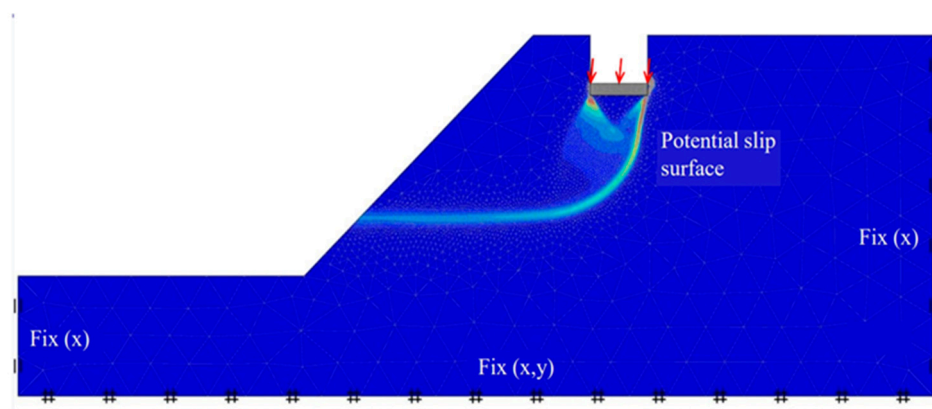


Figure 2. Numerical model, boundary condition, and failure mechanism.

2.1. Statistical Analysis of the Dataset

Based on the above dataset, various statistical descriptions, such as the range mean standard deviation (STDEV), skewness, and kurtosis value of the input and output parameters are presented in Table 2. As per the statistical description presented in Table 2, the slope inclination (β) varies from 15° to 60° ; the normalized slope height (H/B) ranges from 1 to 4; the normalized depth (D/B) ranges from 0 to 2; the horizontal seismic acceleration (k_h) ranges from 0.1 to 0.3; the strength ratio ($c_u/\gamma B$) ranges from 1.5 to 5; the normalized distance (L/B) ranges from 0 to 4; and the seismic bearing capacity (N_c) ranges from 0 to 8.48. The value of skewness for the normalized slope height (H/B), the strength ratio ($c_u/\gamma B$), and the normalized distance (L/B) variables obtained a higher value, so these variables deviate more from their mean value than the other variables. The kurtosis value for all variables was negative, meaning that the taken dataset had a lower peak in a symmetric distribution.

Table 2. Statistical description of the input and output data.

Statistics	b	H/B	D/B	k_h	c_u/gB	L/B	N_c
Max.	60	4	2	0.3	5	4	8.48
Min.	15	1	0	0.1	1.5	0	0
St. dev.	16.78	1.25	0.82	0.08	1.47	1.48	1.58
Mean	37.5	2.3	1	0.2	3	1.75	5.2
Skewness	0.000	0.382	0.000	0.000	0.471	0.435	−0.256
Kurtosis	−1.361	−1.501	−1.501	−1.501	−1.501	−1.154	−0.769

Notably, certain seismic bearing capacity factors can depend on each other. Hence, the correlation heatmap is derived for each input and output variable and is shown in Figure 3. When the input variables' correlation coefficient has a high positive or negative value, it can be difficult to ascertain the impact of these factors on the output. It can be concluded that the normalized depth (D/B), normalized distance (L/B), and strength ratio ($c_u/\gamma B$) have a positive correlation coefficient with the seismic bearing capacity (N_c); furthermore, the normalized slope height (H/B), slope inclination (β), and horizontal seismic acceleration (k_h) have a significant negative correlation with the seismic bearing capacity (N_c), which means that each input variable has quantified association strength with the output variable.

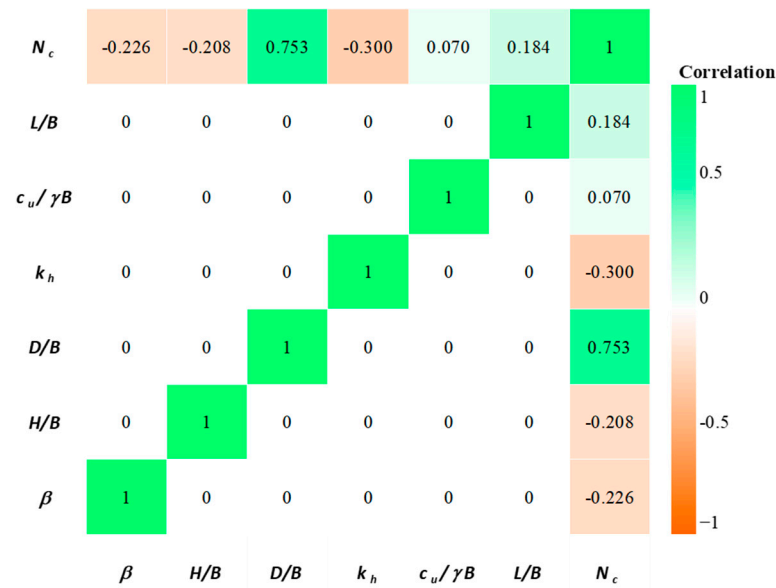


Figure 3. Correlation heatmap matrix.

In this study, five advanced computational models were used to evaluate the seismic bearing capacity (N_c) based on influential variables, such as slope inclination (β); normalized depth (D/B); normalized distance (L/B); normalized slope height (H/B); strength ratio ($c_u/\gamma B$); and the horizontal seismic acceleration (k_h). To reduce the dimensional effect and to improve the accuracy of the proposed models, we first normalized the variables between 0 and 1 using the min-max approach (with Equation (2)) because the scales of the variables utilized in the model's construction are not the same [36,37].

$$D_N = \frac{D_{act} - D_{min}}{D_{max} - D_{min}} \quad (2)$$

where D_N and D_{act} denote the normalized and actual value variables, respectively, and D_{min} and D_{max} denote the minimum and maximum values of the variables, respectively. In the study, a total of 1296 samples were randomly divided into two parts: the training datasets and the testing datasets. The training phase, which includes 70% of the whole dataset (i.e., 907 datasets) was utilized for training the model. The testing dataset, which contains 30% of the whole dataset (i.e., 389 dataset), was used to test the model. The training data can overfit the model, which may lead to 'memorization' instead of 'generalization'. Among the various methods available, the dropout method was used to solve this problem and to avoid overfitting.

2.2. Performance Evaluation Indicators

For evaluation of the model's performance, the following evaluation indicators are widely used: (1) the coefficient of determination (R^2); (2) the variance account factor (VAF); (3) the performance index (PI); (4) Willmott's index of agreement (WI); (5) the mean absolute error (MAE); (6) the weighted mean absolute percentage error ($WMAPE$); (7) the mean bias error (MBE); and (8) the root-mean-square error ($RMSE$). The following Equations (3) to (10) indicate the definition of the abovementioned indices [38–40].

$$R^2 = \frac{\sum_{i=1}^n (d_i - d_{avg})^2 - \sum_{i=1}^n (d_i - y_i)^2}{\sum_{i=1}^n (d_i - d_{avg})^2} \quad (3)$$

$$VAF = \left(1 - \frac{var(d_i - y_i)}{var(d_i)} \right) \times 100 \quad (4)$$

$$PI = adj.R^2 + (0.01 \times VAF) - RMSE \quad (5)$$

$$WI = 1 - \left[\frac{\sum_{i=1}^n (d_i - y_i)^2}{\sum_{i=1}^n \{|y_i - d_{avg}| + |d_i - d_{avg}|\}^2} \right] \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - d_i| \quad (7)$$

$$WMAPE = \frac{\sum_{i=1}^n \left| \frac{d_i - y_i}{d_i} \right| \times d_i}{\sum_{i=1}^n d_i} \quad (8)$$

$$MBE = \frac{1}{n} \sum_{i=1}^n (y_i - d_i) \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i - y_i)^2} \quad (10)$$

where d_i and y_i denote the actual and predicted i^{th} values of the seismic bearing capacity, respectively; “ n ” denotes the total number of datasets used in the training and testing phases; and d_{avg} denotes the average value of the actual seismic bearing capacity. Some of these statistical indices, including R^2 , VAF , PI , and WI , were recognized as accuracy parameters, whereas others, such $RMSE$, $WMAPE$, MAE , and MBE , were classified as error parameters. A perfect model would have a predicted a value that was identical to or very close to the ideal value (see Table 3).

Table 3. Ideal values of the statistical parameters.

Statistical Parameters	R^2	$WMAPE$	$RMSE$	VAF	PI	WI	MAE	MBE
Ideal Values	1	0	0	100	2	1	0	0

3. Methodology of Soft-Computing Techniques

3.1. Minimax Probability Machine Regression (MPMR)

The minimax probability machine regression (MPMR) technique was suggested by Lanckreit et al. [41] as a powerful probabilistic machine learning approach. For the linear classification problem, the MPMC algorithm was first used in which the minimum probability of the correctly classified future data was maximized and furthered by the help of the Mercer kernels function nonlinear version of this theorem. The MPMR algorithm’s goal, which is based on a nonlinear regression framework, is to maximize the minimal probability for the correct regression of the actual data with minimum probability, which falls between the upper and lower bounds of the actual regression. Following the minmax probability machine classification (MPMC) technique, the following expression is used for regression:

$$Sup_{E[z]} = \sum_{x \sim (\mu, \Sigma_z)} Pr \{ a^T x \geq b \} \quad (11)$$

where $x \in (\mu, \Sigma_z)$ is a random vector representing the class of statistical information, $Sup_{E[z]}$ represents the supremum over the distribution having the mean $\mu \in R^n$ and covariance matrix $\Sigma_z \in R^{n \times n}$, and a and b represent the constant value.

MPMC is a classification method created by Strohmam and Grudic [42]. It is used as a binary classifier for separating data into two sets of points to implement the MPMR algorithm, in which one set of classes is produced by shifting all the regression data along the $+\varepsilon$ side of the output variable axis, and the other is obtained by shifting all the regression data along the $-\varepsilon$ side of the output variable axis. The MPMR model is based on a regression

surface, and the classification boundary between these two separating surfaces is $\pm\epsilon$. This paper uses the MPMR algorithm to estimate the undrained seismic bearing capacity factor. Assume a set of training data was generated using an unknown regression function $f(x)$, where $f: R^d \rightarrow R$ with the following structure. Now, the minimum probability may be estimated directly with the help of the following regression equation, which is based on kernel formulation.

$$y = \sum_{i=1}^n \alpha_i K(x_i, x) + b \quad (12)$$

where $K(x_i, x)$ represents the kernel function, y represents the output of the MPMR algorithm, and n represents the total number of datasets of independent variables. α_i and b represent the output of the MPMR model. As a kernel function, the radial basis function (RBF) is employed, i.e., $K(x_i, x) = \exp\left[-(x_i, x)(x_i, x)^T / (2\sigma^2)\right]$, in which σ is defined as the width of the RBF. In this study, the slope inclination (β), normalized depth (D/B), normalized distance (L/B), normalized slope height (H/B), the strength ratio ($c_u/\gamma B$), and the horizontal seismic acceleration (k_h) are used as inputs of the MPMR model, and the seismic bearing capacity (N_c) is used as the output of the MPMR model. Thus, $y = N_c$ and $K(x_i, x) = f\left(\beta, \frac{H}{B}, \frac{L}{B}, \frac{D}{B}, \frac{c_u}{\gamma B}, k_h\right)$. However, the kernel function can be investigated for its potential in developing a regression model. (Please refer to Strohmann and Grudic's [42] work for comprehensive methodology details.) The MPMR model was constructed using MATLAB software.

3.2. Functional Network (FN)

Castillo et al. [43] introduced a new technique: the updated neural network. To estimate the unknown neuron functions, an FN algorithm was used on both data and domain knowledge. This is considered to be the advantage of FN over the ANN algorithm. Initially, the FN method had a complex topology, but it can now be reduced to simple terminology. Thus, FNs solve the "black box" issue with neural networks by combining domain knowledge with data knowledge to infer the problem's topology. FNs utilize data to estimate the unknown neuron functions and domain knowledge to ascertain the network's topology. In an FN, it is expected that the neural functions have several arguments and are vector-valued functions, whereas the ANN algorithm uses the sigmoidal function. Parametric and structural learning are used to learn and estimate the functions. In contrast, artificial neural networks have predefined neural functions. Compared to ANNs, the FN's intermediary layers allow many neuron outputs to be coupled to the same unit. Functional networks can be classified as either structural learning or parametric learning, depending on the approach taken to train them. The network's initial topology is constructed using the designer's actual resources in structural learning. Using functional equations, we can further reduce the complexity of the problem. On the other hand, parametric learning relies on a combination of functional families to estimate neuron function. The parameters in question are calculated using the data at hand. Three distinct kinds of components make up a functional network. They include data stores (input, output, and processing layers), processors, and directed link sets. The following mathematical Equation (13) is proposed to approximate the neural function.

$$f_i(x) = \sum_{j=1}^n a_{ij} \phi_{ij}(X) \quad (13)$$

where (X) represents the input vector and ϕ_{ij} represents the shape function; these can be polynomial functions, such as $(1, x, x^2, x^3, \dots, x^n)$; a trigonometric function, such as $(\sin(x), \cos(x), \tan(x), \dots, \sin(2x))$; exponential functions, such as $(e^x, e^{2x}, \dots, e^{nx})$; or any other acceptable function. Associative optimization functions are used to obtain a system of linear or nonlinear algebraic equations. Working with a functional network necessitates prior knowledge of the functional equation. Cauchy's functional equation appears more frequently than any other type in the class of functional equations. The degree/order of the function and the type of basic function (exponential, polynomial, sine,

cosine, or tangent) are utilized to determine the FN's effectiveness. To construct the FN model, the tan basic function (BF) was adopted. The basic architecture of the FN, which is presented in Figure 4, was used to predict the bearing capacity (N_c) of the strip footing.

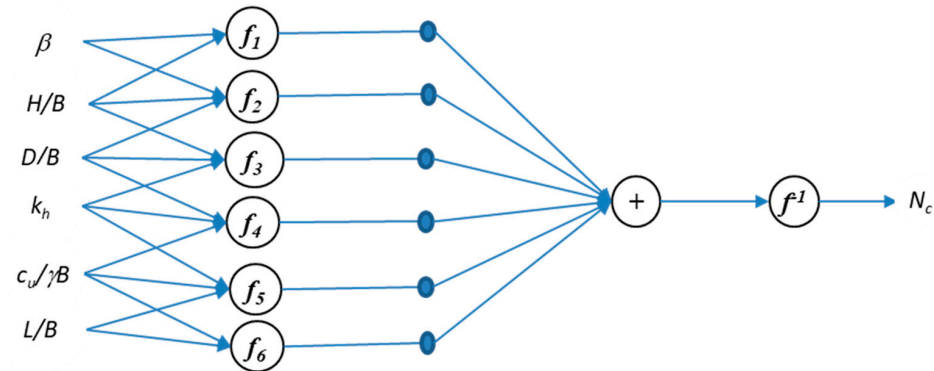


Figure 4. Basic Architecture of Functional Network.

3.3. Convolutional Neural Network (CNN)

A convolutional neural network (CNN) is a deep, feed-forward NN with global sliding, local connections, and weight sharing; as such, it can solve the problem of excessive parameters and prolonged training time as the hidden layer increases, thus making the network extraordinarily applicable and generalizable [44]. It consists primarily of a one-dimensional CNN (1D-CNN) and a high-dimensional CNN. Furthermore, the 1D-CNN is frequently employed for time series and natural language processing; meanwhile, for image processing and video processing, 2D-CNN and 3D-CNN have been used (as in Wang et al. [45]). For this prediction problem, it should be noted that a 2D-CNN has been employed. It comprises one input, two convolutional layers, two pooling layers, one fully connected layer, and an output layer.

Convolutional layers are typically the product of the input matrix multiplied by the filters. Filters, sometimes known as kernels, are used to identify and categorize features within the incoming data. Using pooling layers can reduce the spatial dimensions of the incoming data. In contrast, a fully connected layer adheres to the structure of a CNN and is made up of a number of hidden layers, which further combine the features that are extracted. The whole model structure was built, and the main work focuses on the training of the model. The objective of the training was to determine the optimal values for all parameters, including the weights and biases, to minimize the loss function. This function is used to quantify the degree to which the measured value deviates from the anticipated value. The basic structure of the CNN model is presented in Figure 5.

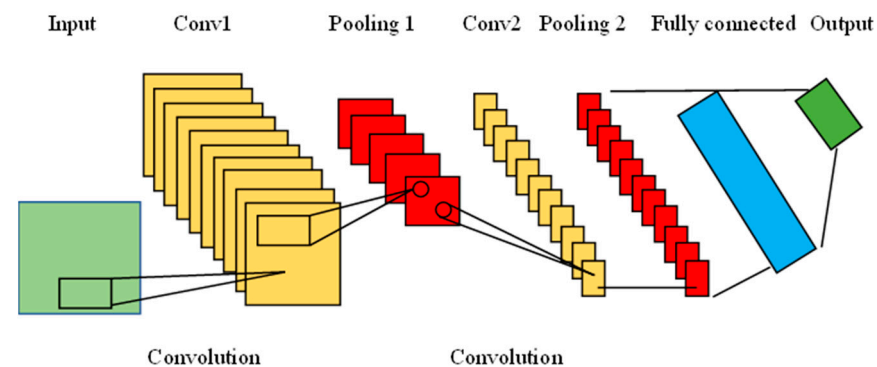


Figure 5. Basic structure of a CNN model.

3.4. Recurrent Neural Networks (RNN)

Recurrent neural networks (RNNs) are dynamic neural networks that are used to solve time series problems [46]. RNNs are supervised machine learning models that take sequence data as their input. They are distinguished from other machine-learning (ML) model architectures by their use of recurrent connections, which means that the output of one cell is related to the output of the previous cell. More specifically, the network memorizes the information from the cell's output that came from the previous output. In contrast to a standard neural network, it possesses a recursive loop, as is demonstrated in Figure 6.

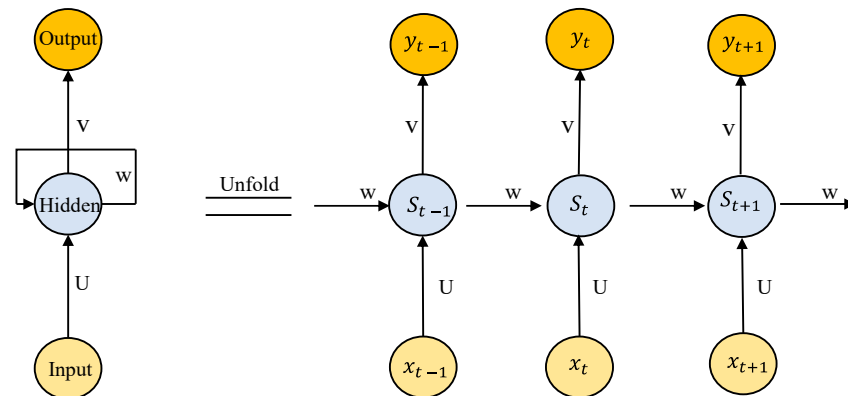


Figure 6. Basic Architecture of a Simple RNN model.

The parameters of RNNs are trained by employing backpropagation through time (BPTT) values. The BPTT values propagate the difference between the ground truth and the output at time t , and backward to time $t - 1$. Likewise, an error at time $t - 1$ is propagated at time $t - 2$, and the training is then conducted retroactively. Here, the fundamental equations of a simple RNN are depicted in Equations (14) and (15).

$$S_t = f(Ux_t + WS_{t-1} + b_h) \quad (14)$$

$$y_t = f(VS_t + b_o) \quad (15)$$

where the variables x_t , S_t , y_t represent the input, the hidden, and the output layers at time t , respectively. Additionally, W , U , and V are the shared parameters with W , indicating the weight of the input layers. U stands for the weight of the current state and V denotes the weight of outputs. Here, $f(\cdot)$ is an activation function, and b_h and b_o are the biases of the hidden and output layers, respectively.

The RNN that is typically used is both straightforward and effective. In actuality, however, it can be challenging to train a model for problems that involve a significant amount of time between the target and the antecedent-related events [47]. In context, the RNN cannot keep good memory if the time interval is large and suffers from a vanishing gradient problem (BPTT).

3.5. Group Method of Data Handling (GMDH)

Ivakhnenko developed GMDH in 1971 [48] as an inductive learning algorithm, and it has been extensively applied to the field of civil engineering to analyze complex and nonlinear problems. The GMDH network is often referred to as a polynomial neural network due to the feed-forward of the neural network's structure [49]. The GMDH network, unlike other networks, continuously changes throughout the training process. The identification problem is essentially defined as finding a function \hat{f} that may be approximately used in place of the actual one f to predict output \hat{y} for a given input vector $X = (x_1, x_2, x_3, \dots, x_n)$, which is as close as feasible to its actual output y . Consequently,

when considering m observations of a multiple input, the single-output data pairs y_i can be written as

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \quad (i = 1, 2, \dots, m) \quad (16)$$

GMDH-type neural networks can now be trained to predict output values \hat{y}_i for every given input vector $X = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})$, i.e.,

$$\hat{y}_i = \hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \quad (i = 1, 2, \dots, m) \quad (17)$$

To minimize the square of the difference between the actual and predicted output, that is, to find a GMDH-type neural network, we used

$$\sum_{i=1}^m [\hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) - y_i]^2 \rightarrow \min \quad (18)$$

A complex discrete form of the Volterra functional series was used in the form of

$$y = a_0 + \sum_1^n a_i x_i + \sum_1^n \sum_1^n a_{ij} x_i x_j + \sum_1^n \sum_1^n \sum_1^n a_{ijk} x_i x_j x_k + \dots \quad (19)$$

which are known as the Kolmogorov–Gabor polynomial [50] and can be used to express general relationships between input and output variables.

This comprehensive mathematical description can be expressed as a system of partial quadratic polynomials with just two variables (neurons), taking the form of

$$\hat{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \quad (20)$$

Regression techniques were used to determine the coefficients a_i in Equation (20) to minimize the difference between the calculated output \hat{y} and the actual output y for each pair of input variables x_i, x_j . This was conducted to determine the coefficients of each quadratic function G_i and to ensure that the output in the entire set of input–output data pairs fit as closely as possible.

$$E = \frac{\sum_{i=1}^m (y_i - G_i)^2}{m} \rightarrow \min \quad (21)$$

The GMDH algorithm's basic form involves selecting all possible combinations of two independent variables from a total of n input variables to build the regression polynomial in the form of Equation (20), whereby, in the sense of the least-squares method, it can best fit the dependent data $(y_i, i = 1, 2, \dots, m)$. As a result, $\binom{n}{2} = \frac{n(n-1)}{2}$ neurons will be developed from the observations $\{(y_i, x_{ip}, x_{iq}) \mid (i = 1, 2, \dots, m)\}$ for various $p, q \in \{1, 2, \dots, n\}$ in the feed-forward network's first hidden layer. Therefore, it is now possible to create m data triples $\{(y_i, x_{ip}, x_{iq}) \mid (i = 1, 2, \dots, m)\}$ from random selection by utilizing such $p, q \in \{1, 2, \dots, n\}$ in matrix form.

$$\begin{bmatrix} x_{1p} & x_{1q} & \vdots & y_1 \\ x_{2p} & x_{2q} & \vdots & y_2 \\ \dots & \dots & \dots & \dots \\ x_{Mp} & x_{Mq} & \vdots & y_M \end{bmatrix}$$

For each row of m data triples, the quadratic subexpression in the form of Equation (20) can be used to easily generate the following matrix equation:

$$Aa = Y \quad (22)$$

where a represents the quadratic polynomial's unknown coefficient vector in Equation (20),

$$a = \{a_0, a_1, a_2, a_3, a_4, a_5\} \text{ and } Y = \{y_1, y_2, y_3, \dots, y_m\}^T \quad (23)$$

and Y represents the vector of the observation's output value. It is clear that matrix A is as follows:

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & x_{1p}^2 & x_{1q}^2 \\ 1 & x_{2p} & x_{2q} & x_{2p}x_{2q} & x_{2p}^2 & x_{2q}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{Mp} & x_{Mq} & x_{Mp}x_{Mq} & x_{Mp}^2 & x_{Mq}^2 \end{bmatrix}$$

Using the least-squares method with multiple regression analysis, the normal equations are solved as follows:

$$a = (A^T A)^{-1} A^T Y \quad (24)$$

which, for the entire set of m data triples, determines the vector containing the optimal coefficients of the quadratic Equation (20). Depending on the network's connectivity structure, this process is repeated for each neuron in the subsequent hidden layer. Such a solution, however, is rather prone to round-off errors and, more critically, to the singularity of these equations when it is obtained directly from solving normal equations.

4. Results and Discussion

4.1. Tuning Hyperparameters of the Proposed Models

For the MPMR technique, the design values of the error insensitive zone (e) and the width (s) of the radial basis function are 0.005 and 0.3, respectively. For the FNN, the cos function was used with a four-degree polynomial. The GMDH model was developed based on eight layers with three neurons and $a = 0.6$. In constructing the CNN model, the following hyperparameters were used to obtain the best result. Table 4 displays the optimal values for the deterministic parameters of the CNN and RNN models. The method of predicting the seismic bearing capacity factor (N_c) of the strip footing is presented by using a flow chart, as shown in Figure 7.

Table 4. Details of the hyperparametric configurations for the CNN and RNN models.

Hyperparameters	CNN	RNN
Number of hidden layers	3	3
Batch size	150	150
Activation function	ReLU	ReLU
Dense layer	64	64
Number of epochs	500	500
Loss function	mean_squared_error	mean_squared_error
Optimizer	adam	adam

4.2. Performance Evaluations of the Proposed Models

Figures 8–12 show the scatter plots of the actual seismic bearing capacity (N_c) that was obtained through the FELA solution as well as the model-predicted seismic bearing capacity (N_c) for both the training and testing phases for the MPMR, FN, CNN, RNN, and GMDH models, respectively; these were constructed to provide a more in-depth look at the performance. The line $y = x$ shows the ideal model with the actual output value = the model-predicted output value. Figure 2 demonstrates this clearly; all of the data points cluster closely around line $y = x$, suggesting that the MPMR model is the best fit. Most data points lie between the dotted line, which indicates a $\pm 20\%$ deviation of the predicted output from the actual regression line $y = x$. Both the training and testing phases were observed to have significantly less variation in the MPMR model, followed by the CNN, RNN, FN, and GMDH models.

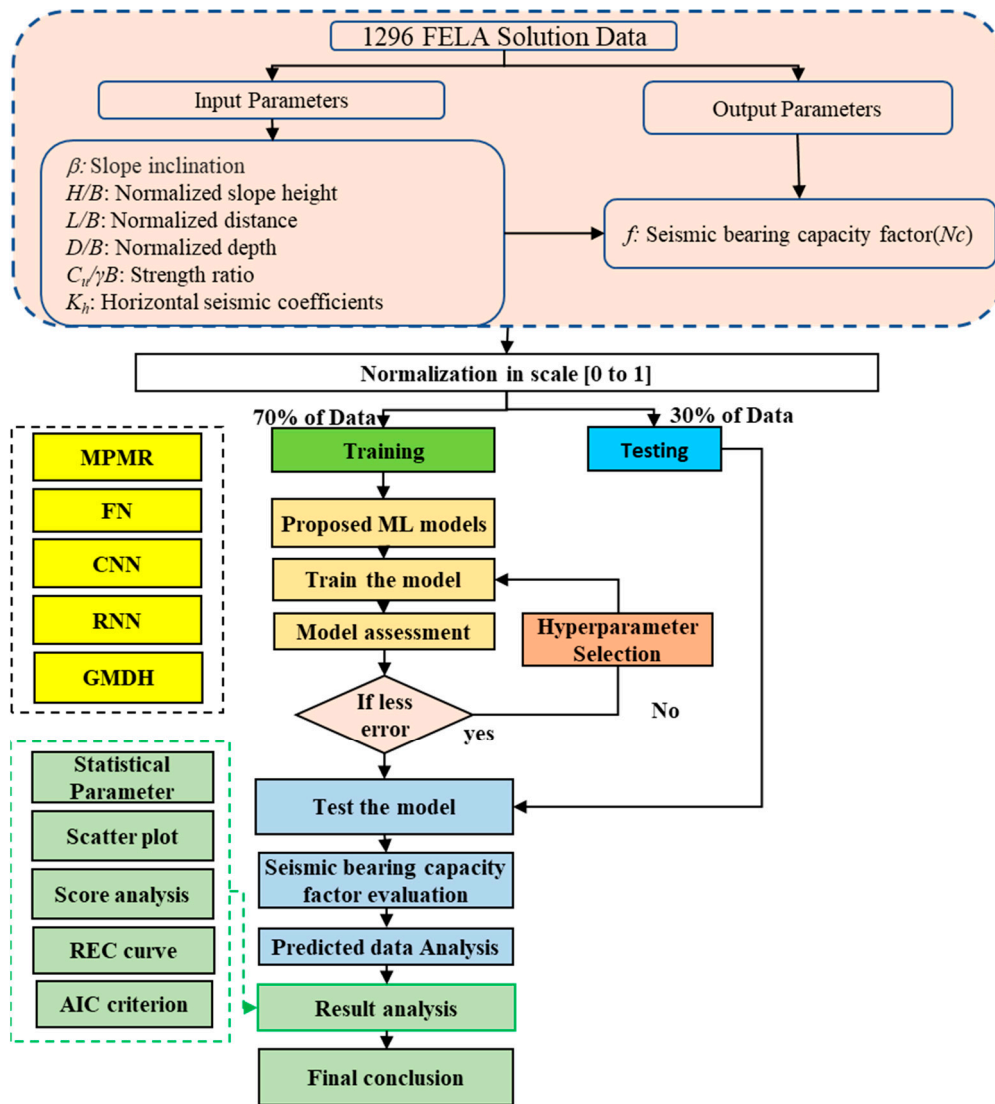


Figure 7. The flowchart of the methodology in predicting the seismic bearing capacity factor (N_c) of the strip footings.

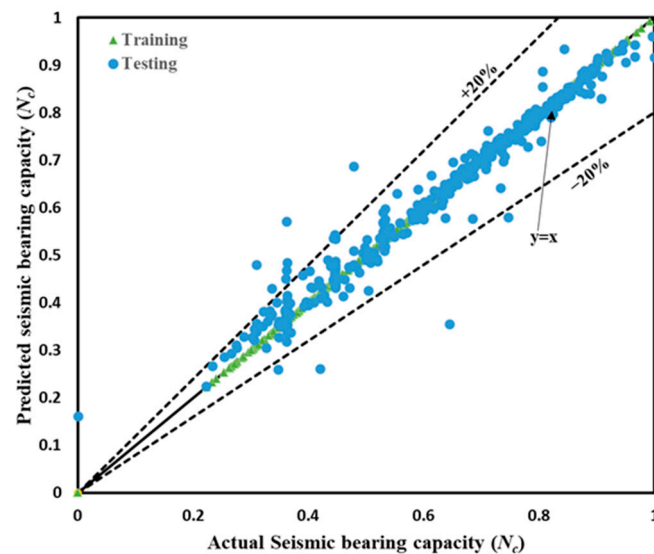


Figure 8. Scatter plot of the actual and predicted seismic bearing capacities (N_c) for the MPMR model.

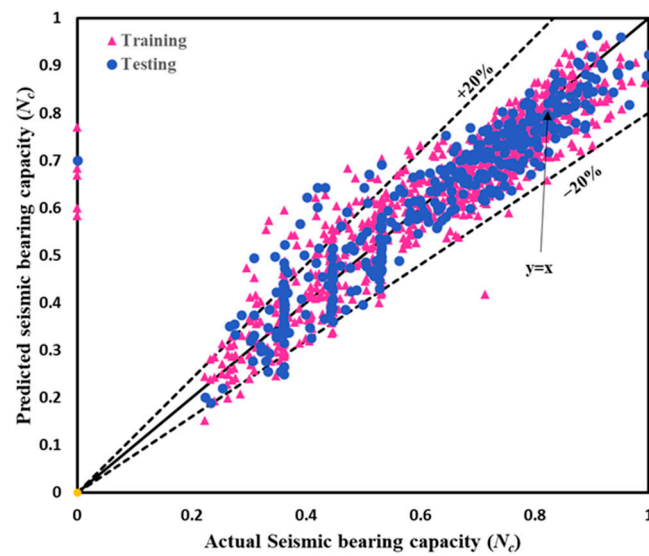


Figure 9. Scatter plot of the actual and predicted seismic bearing capacities (N_c) for the FN model.

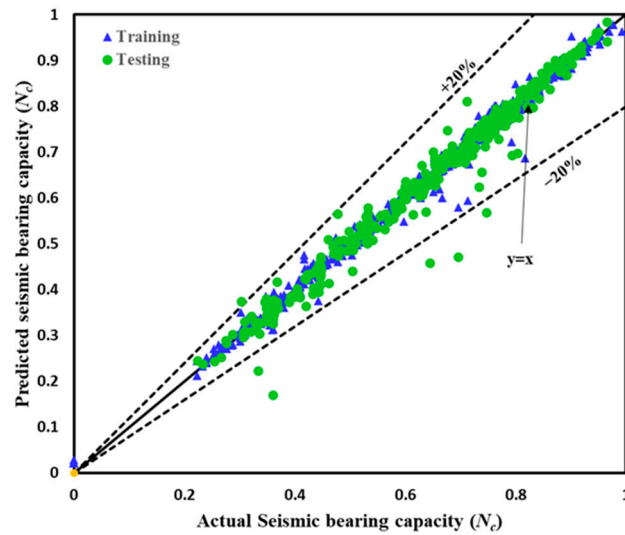


Figure 10. Scatter plot of the actual and predicted seismic bearing capacities (N_c) for the CNN model.

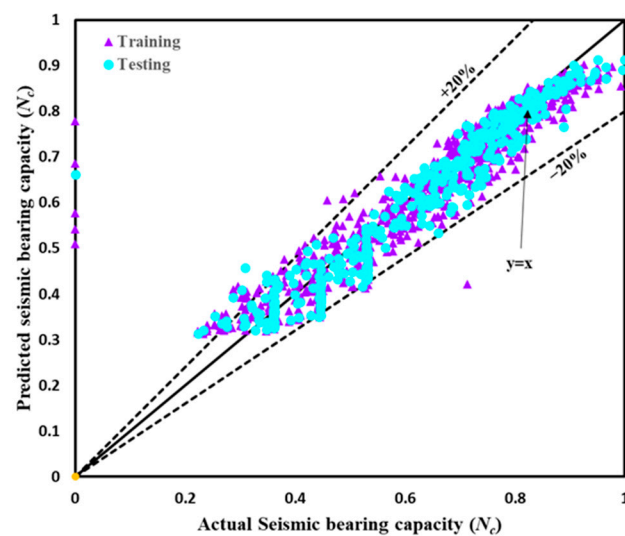


Figure 11. Scatter plot of the actual and predicted seismic bearing capacities (N_c) for the RNN model.

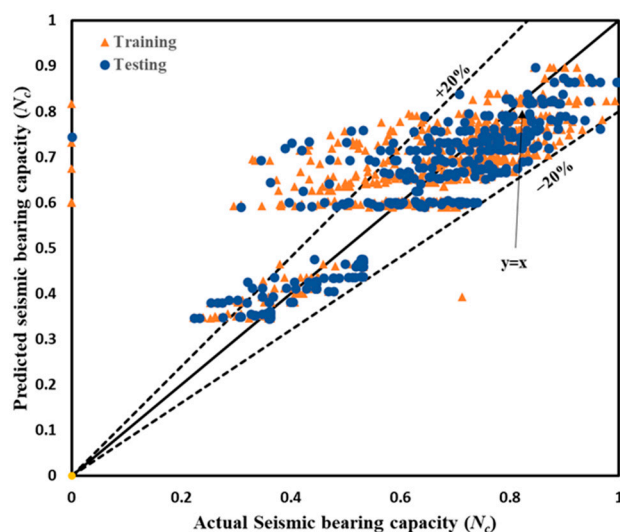


Figure 12. Scatter plot of the actual and predicted seismic bearing capacities (N_c) for the GMDH model.

4.3. Performance Parameters

To assess the performance of the proposed models, certain statistical parameters, such as R^2 , VAF, PI, WI, MAE, WMAPE, MBE, and RMSE, were evaluated, the results of which are presented in Table 5. The results in Table 5 give some quantitative information on the performance of each algorithm for both the training and testing phases. It also shows the ranks for the models with better performance. The models that achieve statistical parameter values close to their respective ideal values (which are presented in Table 3) are considered the most efficient. Generally, the models that attained the higher values for the accuracy parameter and the lower values for the error parameter are considered the best. The proposed MPMR models attained the maximum accuracy ($R^2 = 1$) and least degree of error ($RMSE = 0.00$), followed by the CNN ($R^2 = 0.9945$, $RMSE = 0.0140$), RNN ($R^2 = 0.8791$, $RMSE = 0.0655$), FN ($R^2 = 0.8231$, $RMSE = 0.0785$), and GMDH ($R^2 = 0.722$, $RMSE = 0.0985$) models during the training phase. Furthermore, the CNN model attained the maximum accuracy ($R^2 = 0.9754$) and least degree of error ($RMSE = 0.0297$), followed by the MPMR, RNN, FN, and GMDH models during the testing phase. Overall, the MPMR model outperformed the CNN, RNN, FN, and GMDH models based on the other index results (which are presented in Table 5).

Table 5. The statistical parameter values.

Model	Phase	R^2	VAF	PI	WI	MAE	WMAPE	MBE	RMSE
MPMR	Train	1	100	2	1	0	0	0	0
	Test	0.9577	95.6775	1.8751	0.9884	0.0214	0.0347	0.0036	0.0387
FN	Train	0.8231	82.3142	1.5666	0.9496	0.0508	0.0841	0.0000	0.0785
	Test	0.8605	86.0316	1.6493	0.9606	0.0480	0.0776	0.0044	0.0694
CNN	Train	0.9945	99.4461	1.9749	0.9986	0.0085	0.0141	0.0018	0.0140
	Test	0.9754	97.4407	1.9197	0.9937	0.0167	0.0270	0.0022	0.0297
RNN	Train	0.8791	87.8739	1.6916	0.9658	0.0385	0.0638	0.0076	0.0655
	Test	0.9143	91.3719	1.7705	0.9749	0.0371	0.0600	0.0142	0.0562
GMDH	Train	0.7220	72.1942	1.3436	0.9153	0.0663	0.1098	0.0009	0.0985
	Test	0.7444	74.4356	1.3909	0.9228	0.0654	0.1057	0.0054	0.0938

4.4. Rank Analysis

A rank analysis is the most straightforward and extensively used technique for evaluating model performance and comparing robustness. The maximum score depends on the number of models considered in the analysis (i.e., five). In this analysis, we assigned the score based on the statistical parameter values for both the training and testing phases sepa-

rately. Which of the proposed models produces the best outcomes is based on the statistical parameters that are assigned the highest possible score (i.e., five) and vice versa. If two models produce the same statistical result, their ranking ratings could be identical. When scoring a model, the training and test scores are added together by using Equation (25) to obtain a total score. The model that attained the highest score is ranked as one, and the model that earned the lowest score is ranked as five. From this analysis, it can be concluded that the MPMR model attained the highest total score (72), followed by the CNN (70), RNN (44), FN (35), and GMDH (19) models (as presented in Table 6). Thus, the MPMR model gives the most accurate result, followed by the CNN, RNN, FN, and GMDH models when calculating the seismic bearing capacity (N_c).

$$\text{Total score} = \sum_{i=1}^m S_i + \sum_{j=1}^n S_j \quad (25)$$

where S_i and S_j represent the score of the individual statistical parameters for the training and testing phases, respectively. Additionally, m and n represent the number of statistical parameters used for the rank analysis.

Table 6. Rank analysis of all proposed models based on statistical parameters.

Parameters		MPMR		FN		CNN		RNN		GMDH	
		TR	TS	TR	TS	TR	TS	TR	TS	TR	TS
R^2	Score	5	4	2	2	4	5	3	3	1	1
RMSE	Score	5	4	2	2	4	5	3	3	1	1
PI	Score	5	4	2	2	4	5	3	3	1	1
WI	Score	5	4	2	2	4	5	3	3	1	1
MAE	Score	5	4	2	2	4	5	3	3	1	1
WMAPE	Score	5	4	2	2	4	5	3	3	1	1
MBE	Score	5	4	4	3	2	5	1	1	3	2
VAE	Score	5	4	2	2	4	5	3	3	1	1
Sub Total		40	32	18	17	30	40	22	22	10	9
Total Score		72		35		70		44		19	
Rank		1		4		2		3		5	

4.5. Sensitivity Analysis

Sensitivity analysis aims to ascertain the impact on the model's target variables (i.e., N_c) regarding the changes to the model's input variables, such as β , D/B , L/B , H/B , $c_u/\gamma B$, and k_h . It is a method that is used for determining the results of a choice when only some of the possible outcomes are known—in other words, an analyst can learn how a shift in a single variable affects a result if they construct a model with that collection of variables. In this analysis, the impact of the input variables on the output variable is determined using the cosine amplitude technique [51]. The data prepared to perform the study were stored in a data array form (V) as follows.

$$V = \{v_1, v_2, v_3, \dots, v_n\} \quad (26)$$

where V represents the input vector of length " n ", and v_i represents the length vector of dimension " m ", which is presented as follows in Equation (27):

$$v_i = \{v_{i1}, v_{i2}, v_{i3}, \dots, v_{im}\} \quad (27)$$

The correlation between the strength of the relation C_{ij} and the dataset of v_i and v_j was calculated by using Equation (28).

$$C_{ij} = \frac{\sum_{k=1}^m v_{ik}v_{jk}}{\sqrt{\sum_{k=1}^m v_{ik}^2 \sum_{k=1}^m v_{jk}^2}} \quad (28)$$

The relative importance between the input parameters and the seismic bearing capacity (N_c) of the footing are presented with a pie chart in Figure 13.

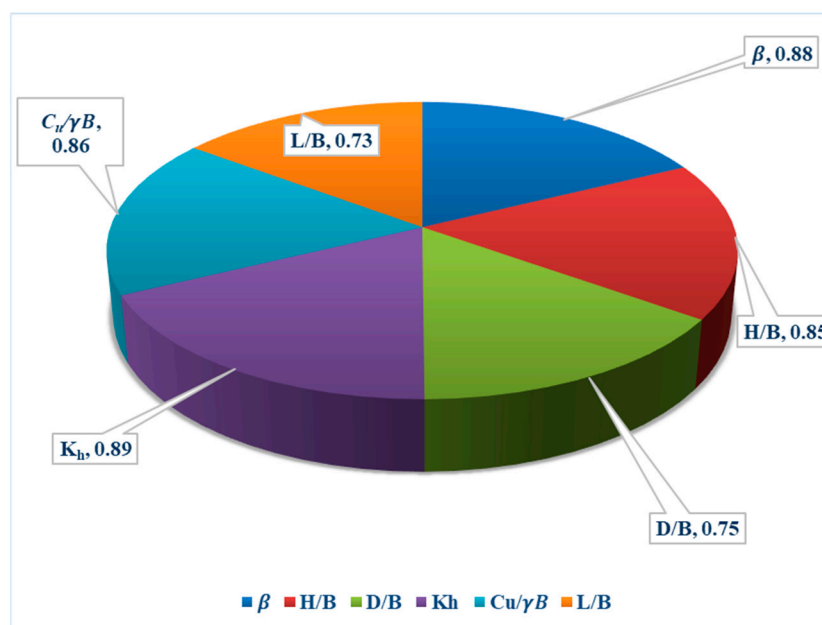


Figure 13. Illustration of the sensitivity analysis result.

From the obtained sensitivity analysis results, it can be concluded that the horizontal seismic acceleration (k_h) has the greatest influence on seismic bearing capacity (N_c) with a value of 0.89, followed by slope inclination (β) with a value of 0.88 and strength ratio ($c_u/\gamma B$) with a value of 0.86. The other parameters, normalized slope height (H/B), normalized depth (D/B), and normalized distance (L/B), have a minimal effect on the seismic bearing capacity (N_c), with values of 0.85, 0.75, and 0.73, respectively. Finally, it can be concluded that all five parameters strongly influenced the seismic bearing capacity (N_c). Hence, the effect of all input parameters was considered when predicting the output. Additionally, sensitivity analysis can serve as a guide for prioritizing which input parameters to use when developing a model.

4.6. Regression Error Characteristic (REC) Curve

A receiver operating characteristic (ROC) curve is a graphical representation of a classifier's performance in a binary classification problem. Although ROC curves only apply to classification problems, regression error characteristic (REC) curves can be used to visualize the performance of regressor models. By plotting the percentage of points with accurate predictions within the tolerance interval against the absolute error tolerance, the x - and y -axes of a regression function represent the margin of error and the precision, respectively. The curve thus obtained is a rough approximation of the error's cumulative distribution function. The predicted error is estimated by the area over the REC curve (AOC). Models perform better when their AOC value is lower. Therefore, ROC curves offer a visual representation of model performance that is both fast and reliable.

Figures 14 and 15 represent the REC curves of all proposed models for both the training and testing phases, respectively.

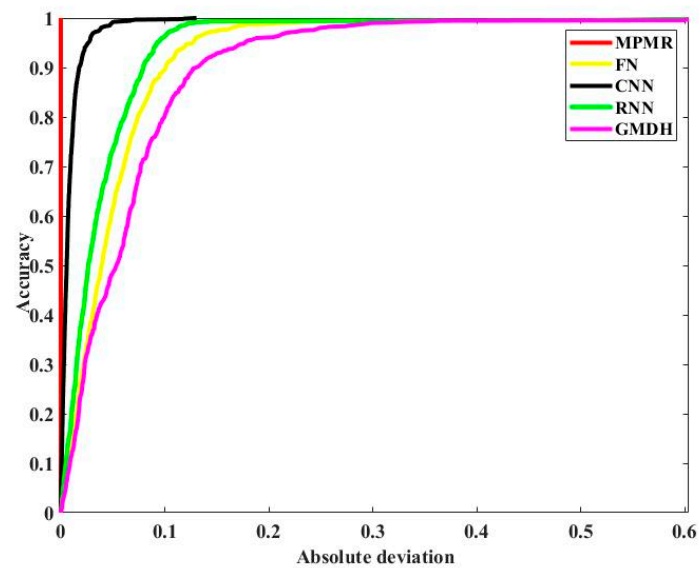


Figure 14. Illustration of the REC curve for the training phase.

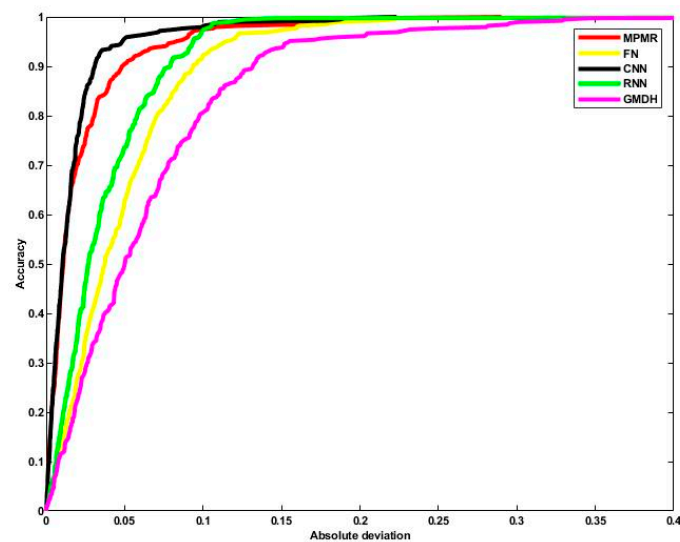


Figure 15. Illustration of the REC curve for the testing phase.

Table 7 represents the AOC value for all the proposed models for both the training and testing phases. From the results presented in Figures 14 and 15, it can be concluded that MPMR is the most accurate model and GMDH is the least accurate model in terms of prediction accuracy. The obtained value of AOC is shown in Table 7. MPMR has the most negligible AOC value (0.000025), followed by the CNN (0.0084), RNN (0.0381), FN (0.0503), and GMDH (0.0658) models in the training phase, and CNN has the lowest AOC value (0.0164), followed by the MPMR (0.0211), RNN (0.0362), FN (0.0471), and GMDH (0.0644) models in the testing phase. Finally, it can be concluded that both the MPMR and CNN models give the most accurate results in the training and testing phases when compared to the RNN, FN, and GMDH models.

Table 7. The AOC value for all proposed methods.

Phase	MPMR	FN	CNN	RNN	GMDH	Ideal Value
Training	2.51×10^{-5}	0.0503	0.0084	0.0381	0.0658	0
Testing	0.0211	0.0471	0.0164	0.0362	0.0644	0

4.7. Akaike Information Criterion (AIC)

Akaike [52] established the Akaike information criterion (AIC) to determine whether trained models are generalizable. The Akaike information criterion (AIC) is used to evaluate the relative quality of statistical models for a given dataset. The AIC estimates the generalization potential of each model concerning other models when given a set of models for a particular dataset. As a result, the AIC offers a model selection method. The AIC value for each model is calculated using Equation (29).

$$AIC = n \times \ln(RMSE^2) + 2k \quad (29)$$

where n represents the number of datasets used to train the model and k represents the total number of input parameters used to train the model. For the best performing model, the AIC value should be the lowest [53,54]. In this study, as presented in Table 8, the MPMR model attained the lowest AIC value (−18,953.65 for training and −2518.42 for testing) compared to the other models. Thus, it can be concluded that the MPMR model has a greater generalization potential, followed by the CNN, RNN, FN, and GMDH models. The comparison of the AIC value for all the models is presented in Figure 16.

Table 8. The AIC value for all models.

Model	MPMR	FN	CNN	RNN	GMDH	Ideal Value
Training	−18,953.65	−4603.04	−7728.65	−4933.07	−4192.56	Lowest value
Testing	−2518.42	−2063.89	−2723.42	−2227.19	−1829.08	Lowest value

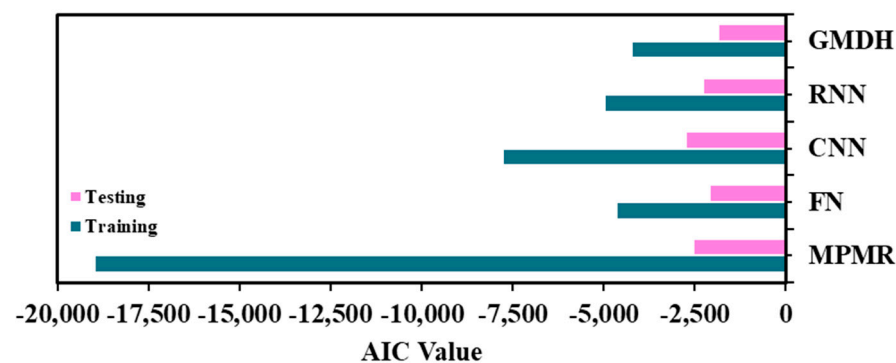


Figure 16. A comparison of the AIC value for all models.

5. Conclusions

This paper presents five advanced machine learning algorithms, i.e., the MPMR, FN, CNN, RNN, and GMDH models, for the estimation of the bearing capacity (N_c) performance of strip footings that are resting on undrained cohesive slopes. The suggested model was constructed by first training and then testing all models on a set of 1296 FELA solutions. A dataset of 1296 FELA solutions with six individual dimensionless variables, i.e., $(\beta, \frac{H}{B}, \frac{L}{B}, \frac{D}{B}, \frac{c_u}{\gamma B}, k_h)$, was taken as the input, and the seismic bearing capacity factor (N_c) was considered as the output during the construction of the model. A sensitivity analysis was performed to determine the influence of the dimensionless input parameters on the seismic bearing capacity factor. The proposed model's efficiency and performance were then analyzed using statistical indicators such as R^2 , VAF , PI , WI , MAE , $WMAPE$, MBE , and $RMSE$. Subsequently, rank analyses, REC curves, and the AIC were used to compare the complete performance of the various proposed models. Based on the obtained results, it is evident that the proposed MPMR model attained the highest prediction accuracy in predicting the N_c of strip footings. Through this study, the following conclusions can be drawn: (1) the proposed advanced machine learning model is a valuable tool for estimating—with less computational effort and greater precision—the seismic bearing

capacity factor of strip footings; (2) the MPMR model has a particularly high potential for predicting the desired N_c value of strip footings, as was evident from the values obtained from the performance parameters, rank analyses, REC curves, and the AIC; (3) the proposed models can be easily implemented for practical application, as well as for numerous applications in seismic research; and (4) the model has a particularly low computational cost of approximately 30 s. Overall, the MPMR model was found to be the best-performing model, followed by the CNN, FN, RNN, and GMDH models. The comparative results indicated that all the proposed models in this study have a better accuracy and ability to estimate the seismic bearing capacity factor of strip footings. However, they have certain limitations that need to be explored in the future, which are identified as follows: (i) the proposed models should be trained for large datasets to predict the desired target value of the seismic bearing capacity factor (N_c) more accurately; (ii) the models created for estimating the seismic bearing capacity factor (N_c) of strip footings are valid for the defined range of dimensionless input parameters; (iii) a comparison should be conducted on the performance of the proposed models via several standard machine learning models and with metaheuristic optimization algorithms on the testing dataset; and (iv) the proposed models do not apply to soils with multiple layers. Layered impacts can be studied in greater depth with additional research.

Author Contributions: Conceptualization, P.S. and W.W. and S.K.; Methodology, S.K.; Software, D.R.K. and K.S.; Formal analysis, D.R.K. and S.K.; Resources, W.W.; Data curation, K.S. and W.J.; Writing—original draft, D.R.K. and W.J.; Writing—review & editing, P.S., W.W. and S.K.; Supervision, P.S. and S.K.; Funding acquisition, W.W. All authors have read and agreed to the published version of the manuscript.

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

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding authors on reasonable request.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Hansen, B.J. *A General Formulae for Bearing Capacity*; Bulletin No. 11; Danish Geotechnical Institute: Lyngby, Denmark, 1961.
2. Satvati, S.; Alimohammadi, H.; Rowshanzamir, M.; Hejazi, S.M. Bearing Capacity of Shallow Footings Reinforced with Braid and Geogrid Adjacent to Soil Slope. *Int. J. Geosynth. Ground Eng.* **2020**, *6*, 41. [[CrossRef](#)]
3. Khalvati Fahliani, H.; Arvin, M.R.; Hataf, N.; Khademhosseini, A. Experimental Model Studies on Strip Footings Resting on Geocell-Reinforced Sand Slopes. *Int. J. Geosynth. Ground Eng.* **2021**, *7*, 24. [[CrossRef](#)]
4. Georgiadis, K. The Influence of Load Inclination on the Undrained Bearing Capacity of Strip Footings on Slopes. *Comput. Geotech.* **2010**, *37*, 311–322. [[CrossRef](#)]
5. Georgiadis, K. Undrained Bearing Capacity of Strip Footings on Slopes. *J. Geotech. Geoenviron. Eng.* **2010**, *136*, 677–685. [[CrossRef](#)]
6. Meyerhof, G.G. The Ultimate Bearing Capacity of Foundations on Slopes. In Proceedings of the 4th International Conference on Soil Mechanics and Foundation Engineering, London, UK, 12–24 August 1957; Volume 1, pp. 384–386.
7. Davis, E.H.; Brooker, J.R. Some Adaptations of Classical Plasticity Theory for Soil Stability Problems. In Proceedings of the Symposium on the Role of Plasticity in Soil Mechanics, Cambridge, UK, 13–15 September 1973; p. 24.
8. Kusakabe, O.; Kimura, T.; Yamaguchi, H. Bearing Capacity of Slopes under Strip Loads on the Top Surfaces. *Soils Found.* **1981**, *21*, 29–40. [[CrossRef](#)] [[PubMed](#)]
9. Shiau, J.S.; Merifield, R.S.; Lyamin, A.V.; Sloan, S.W. Undrained Stability of Footings on Slopes. *Int. J. Geomech.* **2011**, *11*, 381–390. [[CrossRef](#)]
10. Georgiadis, K. An Upper-Bound Solution for the Undrained Bearing Capacity of Strip Footings at the Top of a Slope. *Geotechnique* **2010**, *60*, 801–806. [[CrossRef](#)]
11. Budhu, M.; Al-Karni, A. Seismic Bearing Capacity of Soils. *Geotechnique* **1994**, *44*, 185–187. [[CrossRef](#)]
12. Kumar, J.; Mohan Rao, V.B.K. Seismic Bearing Capacity of Foundations on Slopes. *Geotechnique* **2003**, *53*, 347–361. [[CrossRef](#)]
13. Farzaneh, O.; Mofidi, J.; Askari, F. Seismic Bearing Capacity of Strip Footings near Cohesive Slopes Using Lower Bound Limit Analysis. In Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering: Challenges and Innovations in Geotechnics, ICSMGE 2013, Paris, France, 2–6 September 2013; Volume 2, pp. 1467–1470.

14. Kumar, J.; Ghosh, P. Seismic Bearing Capacity for Embedded Footings on Sloping Ground. *Geotechnique* **2006**, *56*, 133–140. [[CrossRef](#)]
15. Yamamoto, K. Seismic Bearing Capacity of Shallow Foundations near Slopes Using the Upper-Bound Method. *Int. J. Geotech. Eng.* **2010**, *4*, 255–267. [[CrossRef](#)]
16. Georgiadis, K.; Chryssouli, E. Seismic Bearing Capacity of Strip Footings on Clay Slopes. In *Proceedings of the 15th European Conference on Soil Mechanics and Geotechnical Engineering*; IOS Press: Amsterdam, The Netherlands, 2011; pp. 723–728.
17. Kumar, J.; Chakraborty, D. Seismic Bearing Capacity of Foundations on Cohesionless Slopes. *J. Geotech. Geoenviron. Eng.* **2013**, *139*, 1986–1993. [[CrossRef](#)]
18. Chakraborty, D.; Kumar, J. Seismic Bearing Capacity of Shallow Embedded Foundations on a Sloping Ground Surface. *Int. J. Geomech.* **2015**, *15*, 4014035. [[CrossRef](#)]
19. Chakraborty, D.; Mahesh, Y. Seismic Bearing Capacity Factors for Strip Footings on an Embankment by Using Lower-Bound Limit Analysis. *Int. J. Geomech.* **2016**, *16*, 6015008. [[CrossRef](#)]
20. Luo, W.; Zhao, M.; Xiao, Y.; Zhang, R.; Peng, W. Seismic Bearing Capacity of Strip Footings on Cohesive Soil Slopes by Using Adaptive Finite Element Limit Analysis. *Adv. Civ. Eng.* **2019**, *2019*, 4548202. [[CrossRef](#)]
21. Lai, V.Q.; Lai, F.; Yang, D.; Shiau, J.; Yodsomjai, W.; Keawsawasvong, S. Determining Seismic Bearing Capacity of Footings Embedded in Cohesive Soil Slopes Using Multivariate Adaptive Regression Splines. *Int. J. Geosynth. Ground Eng.* **2022**, *8*, 46. [[CrossRef](#)]
22. Thangavel, P.; Samui, P. Determination of the Size of Rock Fragments Using RVM, GPR, and MPMR. *Soils Rocks* **2022**, *45*, e2022008122. [[CrossRef](#)]
23. Mohanty, R.; Suman, S.; Das, S.K. *Modeling the Axial Capacity of Bored Piles Using Multi-Objective Feature Selection, Functional Network and Multivariate Adaptive Regression Spline*, 1st ed.; Elsevier Inc.: Amsterdam, The Netherlands, 2017; ISBN 9780128113196.
24. Das, S.K.; Suman, S. Prediction of Lateral Load Capacity of Pile in Clay Using Multivariate Adaptive Regression Spline and Functional Network. *Arab. J. Sci. Eng.* **2015**, *40*, 1565–1578. [[CrossRef](#)]
25. Kumar, M.; Samui, P. Reliability Analysis of Pile Foundation Using GMDH, GP and MARS. In *Proceedings of the CIGOS 2021, Emerging Technologies and Applications for Green Infrastructure. Proceedings of the 6th International Conference on Geotechnics, Civil Engineering and Structures, Ha Long, Vietnam, 28–29 October 2021*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1151–1159.
26. Dey, P.; Chaulya, S.K.; Kumar, S. Hybrid CNN-LSTM and IoT-Based Coal Mine Hazards Monitoring and Prediction System. *Process Saf. Environ. Prot.* **2021**, *152*, 249–263. [[CrossRef](#)]
27. Tiwari, S.K.; Kumaraswamidhas, L.A.; Gautam, C.; Garg, N. An Auto-Encoder Based LSTM Model for Prediction of Ambient Noise Levels. *Appl. Acoust.* **2022**, *195*, 108849. [[CrossRef](#)]
28. Tiwari, S.K.; Kumaraswamidhas, L.A.; Prince, Kamal, M.; Ur Rehman, M. A Hybrid Deep Learning Model for Prediction and Parametric Sensitivity Analysis of Noise Annoyance. *Environ. Sci. Pollut. Res.* **2023**, *30*, 49666–49684. [[CrossRef](#)]
29. Chen, L.; Chen, W.; Wang, L.; Zhai, C.; Hu, X.; Sun, L.; Tian, Y.; Huang, X.; Jiang, L. Convolutional Neural Networks (CNNs)-Based Multi-Category Damage Detection and Recognition of High-Speed Rail (HSR) Reinforced Concrete (RC) Bridges Using Test Images. *Eng. Struct.* **2023**, *276*, 115306. [[CrossRef](#)]
30. Optum Computational Engineering: Copenhagen, Denmark. Available online: <https://optumce.com/> (accessed on 21 May 2023).
31. Ciria, H.; Peraire, J.; Bonet, J. Mesh Adaptive Computation of Upper and Lower Bounds in Limit Analysis. *Int. J. Numer. Methods Eng.* **2008**, *75*, 899–944. [[CrossRef](#)]
32. Keawsawasvong, S.; Ukritchon, B. Undrained Stability of a Spherical Cavity in Cohesive Soils Using Finite Element Limit Analysis. *J. Rock Mech. Geotech. Eng.* **2019**, *11*, 1274–1285. [[CrossRef](#)]
33. Keawsawasvong, S.; Ukritchon, B. Undrained lateral capacity of I-shaped concrete piles. *Songklanakarin J. Sci. Technol.* **2017**, *39*, 751–758.
34. Shiau, J.; Chudal, B.; Mahalingasivam, K.; Keawsawasvong, S. Pipeline Burst-Related Ground Stability in Blowout Condition. *Transp. Geotech.* **2021**, *29*, 100587. [[CrossRef](#)]
35. Keawsawasvong, S.; Thongchom, C.; Likitlersuang, S. Bearing Capacity of Strip Footing on Hoek-Brown Rock Mass Subjected to Eccentric and Inclined Loading. *Transp. Infrastruct. Geotechnol.* **2021**, *8*, 189–202. [[CrossRef](#)]
36. Kumar, M.; Biswas, R.; Kumar, D.R.; Pradeep, T.; Samui, P. Metaheuristic Models for the Prediction of Bearing Capacity of Pile Foundation. *Geomech. Eng.* **2022**, *31*, 129–147. [[CrossRef](#)]
37. Kumar, D.R.; Samui, P.; Burman, A. Prediction of Probability of Liquefaction Using Soft Computing Techniques. *J. Inst. Eng. Ser. A* **2022**, *103*, 1195–1208. [[CrossRef](#)]
38. Naser, M.Z.; Alavi, A.H. Error Metrics and Performance Fitness Indicators for Artificial Intelligence and Machine Learning in Engineering and Sciences. *Archit. Struct. Constr.* **2021**, 1–19. [[CrossRef](#)]
39. Chai, T.; Draxler, R.R. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?—Arguments against Avoiding RMSE in the Literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [[CrossRef](#)]
40. Kumar, D.R.; Samui, P.; Burman, A. Prediction of Probability of Liquefaction Using Hybrid ANN with Optimization Techniques. *Arab. J. Geosci.* **2022**, *15*, 1587. [[CrossRef](#)]
41. Lanckriet, G.R.G.; El Ghaoui, L.; Bhattacharyya, C.; Jordan, M.I. A Robust Minimax Approach to Classification. *J. Mach. Learn. Res.* **2002**, *3*, 555–582.

42. Strohmann, T.; Grudic, G. A Formulation for Minimax Probability Machine Regression. *Adv. Neural Inf. Process. Syst.* **2002**, *15*, 785–792.
43. Castillo, E.; Cobo, A.; Gómez-Nesterkin, R.; Hadi, A.S. A General Framework for Functional Networks. *Netw. Int. J.* **2000**, *35*, 70–82. [[CrossRef](#)]
44. Hinton, G.E.; Osindero, S.; Teh, Y.-W. A Fast Learning Algorithm for Deep Belief Nets. *Neural Comput.* **2006**, *18*, 1527–1554. [[CrossRef](#)] [[PubMed](#)]
45. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic Graph Cnn for Learning on Point Clouds. *ACM Trans. Graph.* **2019**, *38*, 1–12. [[CrossRef](#)]
46. Elman, J.L. Finding Structure in Time. *Cogn. Sci.* **1990**, *14*, 179–211. [[CrossRef](#)]
47. Gers, F.A.; Schmidhuber, J.; Cummins, F. Learning to Forget: Continual Prediction with LSTM. *Neural Comput.* **2000**, *12*, 2451–2471. [[CrossRef](#)]
48. Ivakhnenko, A.G.; Ivakhnenko, G.A. The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH). *Pattern Recognit. Image Anal.* **1995**, *5*, 527–535.
49. Farlow, S.J. *Self-Organizing Method in Modeling: GMDH*; Type Algorithm; CRC Press: Boca Raton, FL, USA, 1984.
50. Mueller, J.-A.; Lemke, F. *Self-Organising Data Mining*; Extracting Knowledge from Data; Libri GmbH: Hamburg, Germany, 2000.
51. Biswas, R.; Bardhan, A.; Samui, P.; Rai, B.; Nayak, S.; Armaghani, D.J. Efficient Soft Computing Techniques for the Prediction of Compressive Strength of Geopolymer Concrete. *Comput. Concr.* **2021**, *28*, 221–232. [[CrossRef](#)]
52. Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Contr.* **1974**, *19*, 716–723. [[CrossRef](#)]
53. Pradeep, T.; Samui, P. Prediction of Rock Strain Using Hybrid Approach of Ann and Optimization Algorithms. *Geotech. Geol. Eng.* **2022**, *40*, 4617–4643. [[CrossRef](#)]
54. Guven, A.; Kişi, Ö. Estimation of Suspended Sediment Yield in Natural Rivers Using Machine-Coded Linear Genetic Programming. *Water Resour. Manag.* **2011**, *25*, 691–704. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.