Front. Struct. Civ. Eng. 2025, 19(6): 919–932
https://doi.org/10.1007/s11709-025-1191-6

RESEARCH ARTICLE

Boosting-based ensemble machine learning models to predict the bond behavior between concrete and fiber-reinforced polymer bars

Irwan AFRIADI, Chanachai THONGCHOM*, Divesh Ranjan KUMAR, Warit WIPULANUSAT, Suraparb KEAWSAWASVONG

Department of Civil Engineering, Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani 12120, Thailand

Corresponding author. E-mail: tchanach@engr.tu.ac.th

ABSTRACT This study addresses the application of advanced boosting-based ensemble machine learning techniques such as extreme gradient boosting (XGBoost), random forest (RF), category-aware gradient boosting (CATBoost), and adaptive boosting (ADABoost) algorithms to study the bond behavior of fiber-reinforced polymer (FRP) bars in reinforced concrete (RC) beams. To forecast the peak load (P_{max}) of the bond behavior between the FRP bars and concrete, five total input variables, namely, the elastic modulus of the bar (E_f), the tensile strength of the bar (F_f), the compressive strength of the concrete (f'_c), the diameter of the bar (d_b), and the bar embedment length (l_b), were selected for machine learning model construction. The accuracy of the constructed predictive machine learning models was compared using several metric performances. However, rank analysis has also been used to ascertain which models perform the best. According to the findings of rank analysis using several metric performances, XGBoost outperformed RF, ADABoost, and CATBoost. Utilizing the developed advanced machine learning methods to examine the bond behavior of FRP bars in RC beams yields tangible advantages for the construction sector. This approach refines the design precision, minimizes expenses, and elevates the overall effectiveness and longevity of structures reinforced with FRP.

KEYWORDS reinforced concrete, FRP, XGBoost, CATBoost, ADABoost, Random Forest

1 Introduction

The major cause of the degradation of reinforced concrete (RC) structures is the rusting of steel reinforcing bars or rebars. This has a large effect on the cost of maintaining and repairing the reinforced structures. fiber-reinforced polymer (FRP) bars are a more attractive substitute for traditional steel bars for addressing the corrosion problem in RC constructions [1]. FRP (Glass fiber-reinforced polymer (GFRP), Carbon fiber-reinforced polymer (CFRP), Aramid fiber-reinforced polymer (AFRP), Basalt fiber-reinforced polymer (BFRP)) made of glass, carbon, aramid, or basalt provide a large majority of products on the market. Vinyl ester or epoxy systems are examples of thermosetting resins with polymer matrices that frequently used as an alternative material to developed the FRP [2]. They can be easily shaped into sheets, bars, or other shapes as per the user convenient [3]. FRPs are advantageous materials for the repair of concrete structures because they are lightweight, possess high strength, and exhibit significant resistance to corrosion. Applying FRPs externally to concrete constructions efficiently enhances their flexural, shear, axial, and seismic capacities [4].

El Refai et al. [5] conducted several experimental laboratory tests. Direct pullout testing was executed on 36 cylinders RC with FRP bars and 12 cylinders RC with GFRP bars. An experimental examination called a pull-out examination is employed to calculate how well the bond strength of concrete adheres to reinforcing bars [6].

Article history: Received Feb 23, 2025; Accepted Mar 25, 2025

The pull-out examination was operated in accordance with established protocols, such as those described in previous work [7].

According to Rolland et al. [2], for pull-out specimens with the same diameter bar used in the experimental test, it is very promising that GFRP rebars revealed greater bond strength outcomes than specimens that employed reference deformed steel rebars. Godat et al. [8] conducted another laboratory experiments on various types of FRP bars to study the bonding behavior of FRP bars with concrete. Tests were conducted on the capacity of CFRP, BFRP, and GFRP bars to separate from recycled aggregate concrete. The use of BFRP bars can enhance the bonding behavior after the peak strength is reached. These findings imply that the bonding behavior of FRP may depend on the type of bars and several factors. Nepomuceno et al. [9] review that the numerical formula for measuring the bond strength, which incorporates the contribution of different FRP bar surface types, offers significantly improved accuracy in predicting the bond strength of concrete-FRP bars compared to the existing equations reviewed in the literature. Machine learning (ML) models may be trained on an experimental data set to provide tools that are more affordable, quicker, and more dependable than traditional solutions. Moreover, ML models can achieve greater generalizability when large amount of data set are used during training of machine learning models [10].

ML, a subfield of artificial intelligence, makes it possible for researchers and engineers to gain new insights and address the complex nonlinear interactions between the parameters. Several MLs are employed by several researchers to do prediction [11–26] owing to its capability to analyze large data, hidden pattern and provide high accuracy prediction. Recent researches have explored various ML-based approaches to predict bonding performance of concrete-FRP bars. Tang et al. [27] utilized several ML including Random Forest (RF) and extreme gradient boosting (XGBoost) to forecast the rupture pattern and bond strength of FRP-coral aggregate concrete The results found XGBoost identified as the optimal algorithm. Research have been conducted on the bonding capacity and development duration of fiber-reinforced polymer bars inserted into concrete, as noted by Basaran et al. [28]. Additionally, bond strength was estimated using coding equations and machine learning methods and Gaussian Progress Regressor model was found to be the highest prediction accuracy. Zhang et al. [29] forecasted interfacial bond capacity of FRP-concrete and the XGBoost model revealed the highest accuracy. The flexural capacity of FRP-strengthened RC beams was predicted by Zhang et al. [3] using machine learning, and in their findings, it can be observed that the machine learning models have demonstrated greater prediction accuracies than empirical methods. More precise predictions are possible due to ML technique's ability to recognize trends and relationships in experimental data. ML models also avoid overfitting problems and are capable of good generalization. Several ML algorithms were used by Barkhordari and Jawdhari [30] to forecast the length of plastic hinge for structural RC walls of reinforced concrete. This is because the application of ML in seismic engineering design and evaluation is growing to precisely anticipate the behavior of structures. Zhou et al. [31] examined how well current models predict the strength of the link of FRP and concrete. To assess the prediction accuracy of different models, the authors gathered a sizable quantity of experimental data, namely, 969 test results from single-lap shear experiments on FRP-concrete interfacial connections. A study by Kim et al. [32] used an ensemble machine learning technique to measure the capability of the connection of FRP-concrete. To precisely determine the interfacial cohesive characteristics of concrete-FRPs, Su et al. [33] presented an Artificial Neural Network machine learning-based technique. Abuodeh et al. [34] estimated the shear capacity and performance of beams reinforced with FRP that is externally connected sheets using machine learning techniques, more precisely, a robust backpropagating neural network (RBPNN). Using machine learning methods, Su et al. [35] investigated the prediction of the interfacial bond strength of FRPs-concrete. Wang et al. [10] introduced a framework for estimating the bond capacity of the FRP-to-concrete interface called the metaheuristic neuron-vector machine (MNVIM). Yuan et al. [36] provided a data-driven method to identify the bond-slip pattern of the FRP-concrete interface applying ML and Bayesian optimization. To increase the prediction accuracy, the authors examine various machine learning models and tweak their hyperparameters.

Specifically, this study uses advanced regression ML to explore the complicated behavior of FRP bonds in forecasting the peak load (P_{max}) based on multiple experimental tests that have been conducted in the past. Compared to data-driven empirical models, machine learning models have demonstrated greater prediction accuracy [3]. The field of structural and civil engineering can benefit from ML's exploration of the variables influencing bond behavior and its importance. Researchers and engineers may improve structural designs, make well-informed decisions, and guarantee the safety and long-term sustainability of FRP-reinforced structures by comprehending the bonding performance of FRPs-concrete. Applying advanced boosting-based ensemble models for bond strength prediction, offering a significant improvement over conventional ML technique. Additionally, there is a lack of comprehensive data sets covering a wide range of FRP bar properties and geometric configurations. This study aims to address

these gaps by applying state-of-the-art ensemble learning models, namely XGBoost, adaptive boosting (ADABoost), category-aware gradient boosting (CATBoost), and RF, to predict peak bond load (P_{max}) in FRP-reinforced concrete structures.

2 Research significance

Corrosion-related issues are now being resolved by substituting FRP bars for steel bars. Nevertheless, a number of disadvantages of FRP rebars, including their nonhomogeneous properties, low rigidity and linear elastic action, lead to entirely distinct mechanisms for stress transmission between the rebars and the nearby concrete. This is why understanding the bonding behavior of two materials, FRP reinforcing bars and concrete, is crucial before adding them to concrete structures.

In structural engineering and construction, analyzing the bond performance of FRP materials is essential, especially when trying to improve the lifetime and performance of civil infrastructure. Importantly, the bonding properties of FRP composites with substrate materials determine how well they support concrete buildings. To maximize these bonding connections and obtain deeper insights, it is becoming increasingly important to apply cutting-edge technologies such as ML. In this research study, several ensemble machine learning models were proposed for forecasting the ultimate load of pull-out tests between FRPs and concrete. Table 1 shows the available model details from the literature.

3 The collected database

Numerous researchers have carried out extensive tests using various configurations since it is crucial to forecast the ultimate load (P_{max}) of FRP-to-concrete composites. The pull-out examination, as presented in Fig. 1, is the main focus of this research. For every test, the ultimate load (P_{max}) test values, geometric parameters, and material attributes are usually provided. According to earlier research, the elastic modulus of the bar (E_f) [38], tensile strength of the bar (F_f) [39], compressive strength of the concrete (f'_c) [40], diameter of the bar (d_b) [9], surface bar treatment [41], type of FRP bar [41], and bar

Table 1 Model details are available in the literature

No.	Ref.	Input parameters	Datasets used	AI Approach	R^2
1	Zhou et al. [31]	P_{max}, f'_c, K_f, b_f, κ_w, κ_L	969	ANN	0.930
2	Zhang and Xue [37]	L, E_fA_f, D_g/W_g, f_e, f'_c, P_{max}	145	GEP	0.852
				RF	0.912
3	Basaran et al. [28]	Type, surface, d_b, $\sqrt{f'_c}$, C/d_b, L/d_b, A_{tr}/snd_b, position, P_{max}	432	GPR	0.96
				ANN	0.95
				MLR	0.78
				SVMR	0.83
4	Su et al. [35]	f'_c, F_f, E_f, t_f, L_f, b_f, b_c, P_{max}	136	SVM	0.79
				ANN	0.77
				MLR	0.74
5	Kim et al. [32]	f'_c, E_f, t_f, L_f, b_f, b_c, P_{max}	855	CatBoost	0.96
				XGBoost	0.95
				HGBoost	0.95
				RF	0.95
				ANN	0.93
6	Wang et al. [10]	b_c, f'_c, E_f, t_f, L_f, b_f, P_{max}	855	SVR	0.827
				RBFNN	0.803
				LSVR	0.620
				BPNN	0.817
				LSVR+RBFNN	0.808
				EO-LSVR	0.799
				EO-RBFNN	0.830
				MNVIM	0.839

Note: t_f is the thickness of FRP sheet, P_{max} is ultimate load of test, and for the other symbols please refer to the citation.

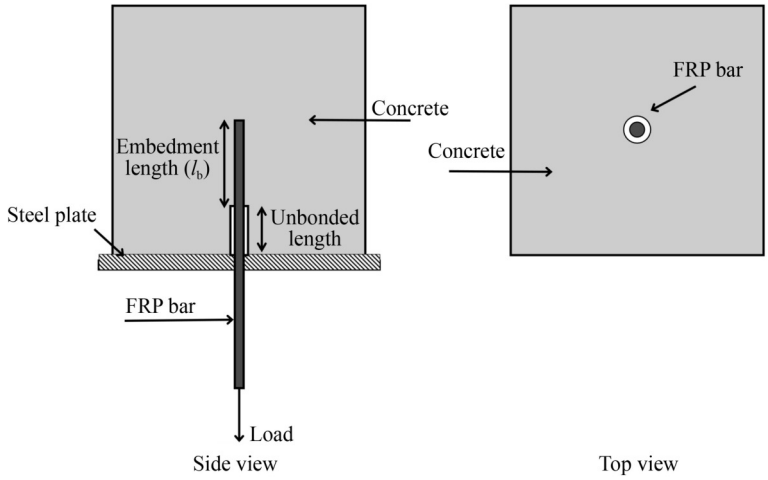

Fig. 1 Pull-out test setup.

embedment length (l_b) [38] all affect the FRP-concrete bond findings. Since the type of FRP bar and surface bar treatment are not considered numerical data, these parameters are excluded from the analysis.

This study compiled several experimental test databases of information from 1, 101 specimens that were subjected to pull-out tests from 19 published literatures [6–8, 39–54]. The database is separated into two stages, testing and training data, using a 70/30 ratio, which is a common choice for small databases [55]. For certain tests, the P_{max} results are not displayed. To overcome this, the author manually calculated the findings using the previous research formula, which is displayed in Eq. (1), where τ_{max} is the maximum bond strength. The concrete compressive strength was determined by evaluating various control specimens to the compressive strength of a cylinder (f_c') to compare the findings from various investigations [56]. Equation (2) was used to convert the compressive strength values (f_{cu}) obtained from the concrete cubes to f_c' [57].

$$\tau_{max} = \frac{P_{max}}{\pi d_b l_b}, \tag{1}$$

$$f_c' = 0.78 f_{cu}. \tag{2}$$

4 Methods

To forecast the final load of the pull-out test, this study used the random forest, XGBoost, ADABoost, and CATBoost models are trained in this study based on experiment data set collected from various literature. Finding the optimal performance among several proposed regressor algorithms is the goal of this study. The XGBoost, CATBoost, ADABoost, and random forest methodologies are introduced in this section.

4.1 Random forest regression

Breiman [58] originally presented the RF technique. It is a popular method that creates a series of decision trees for regression and classification. Bootstrap sampling allows RF to create a diverse set of forest trees. Each tree in the forest is trained on a unique subset drawn from the original data set's training data. As each tree grows, RF gives it an extra boost in selection [32].

4.2 XGBoost regression

Extreme-gradient boosting, or XGBoost, is a well-liked and potent ML method that uses a gradient boosting technique [32]. Kim et al. [32] claim that XGBoost technology is meant to be an accurate and mountable technique for tree boosting. XGBoost is characterized by the following features: parallel tree learning with a cache-conscious column block, sparsity awareness of the split function, estimated crack results based on a one-sided quantile draw, and reformulating the objective problem and adding regularization expression. By increasing processing and memory capacity, XGBoost quickly accelerates learning to its greatest potential. While XGBoost incorporates adjustments to mitigate overfitting and address various extended challenges, its primary safeguard against overfitting lies in a regularized model framework.

4.3 CATBoost regression

According to Huang et al. [59], CATBoost is a gradient boosting method designed specifically to manage features with categories in an effective manner. By processing category features faster during tree splitting, CATBoost improves accuracy and performance. Furthermore, CATBoost offers a technique called minimal variance sampling that helps regularize boosting models. In addition, CATBoost uses symmetric trees to generate

trees, which provides faster results than other ensemble approaches. The approach offers a variety of hyperparameters for customization and custom callback functions, allowing for flexibility and adaptability for a wide range of modeling scenarios [32].

4.4 ADABoost regression

ADABoost, or adaptive boosting, is an iterative boosting method aimed at improving the categorization of minority classes. First, a variant weight is assigned to each observation via the ADABoost method. After a few cycles, the weights of the incorrectly identified observations increase, while the properly recognized observations have smaller weights. The weights allocated to the observations serve as indicators of the class to which they belong, thereby minimizing classification errors and greatly improving classifier performance [60]. The ADABoost algorithm is an ensemble learning method that improves weak learner accuracy by modifying the sample weight distribution.

4.5 Metric performances

The precision quality of the models used was assessed using several metric performances, including the coefficient of determination (R^2), nash–sutcliffe efficiency (*NS*), performance index (*PI*), root mean square error (*RMSE*), variance accounting factor (*VAF*), and mean absolute percentage error (*MAPE*) [61]. Table 2 displays the mathematical formulas with the ideal values.

5 The data set statistical analysis and hyperparameter

5.1 Data analysis and model processing

The data set for this study included the following five input variables: bar embedment length (l_b), diameter bar (d_b), tensile bar (F_f), modulus elastic (E_f), concrete compressive strength (f'_c), and one output parameter, which is the ultimate load (P_{max}). The surface bar treatment and type of FRP bar are excluded; hence, they are not all numerical categories. Table 3 displays the data set's statistical descriptive statistics. The bar embedment (l_b) ranges from 3.3 to 21.25 mm, the diameter bar (d_b) ranges from 9.53 to 285 mm, the elastic modulus (E_f) ranges from 35.74 to 200 GPa, the tensile bar (F_f) ranges from 540 to 2800 MPa, the compressive strength of concrete (f'_c) ranges from 20.67 to 114.34 MPa, and the ultimate load (P_{max}) ranges from 3.11 to 121.79 kN.

Furthermore, Fig. 2 presents the relationship correlation heatmap matrix for each of the input and output variables. The correlation coefficient spans from −1 to 1. A coefficient of a perfectly correlated positive linear relationship is represented by a value of 1, and a perfect negative linear correlation by a value of −1, and a value of 0 suggests no linear correlation at all. According to the results, there is a greater statistical correlation between d_b and l_b with the output parameter, i.e., the ultimate load (P_{max}), while the others have a lower statistical correlation with the ultimate load (P_{max}). The mathematical formula used for the Pearson correlation is shown in Eq. (3).

$$\sigma = \frac{\sum(X_i - \underline{X})(Y_i - \underline{Y})}{\sqrt{\sum(X_i - \underline{X})^2 \cdot \sum(Y_i - \underline{Y})^2}}, \tag{3}$$

where σ is the Pearson correlation, X_i and Y_i are the individual variable values, and $\underline{X}$ and $\underline{Y}$ are the averages of the individual variable values. Apart from the Pearson correlation coefficient, which reflects the correlation among data sets, normalization of all data sets was also used to generalize the scale of all data sets. All the input and output variables are normalized between 0 and 1 using the min–max technique, as shown in Eq. (4).

Table 2 Mathematical equation of metric performances

Performance metrics	Mathematical expression	Ideal value
Root Mean Square Error (*RMSE*)	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$	0
Mean Absolute Percentage Error (*MAPE*)	$MAPE = \sum_{i=1}^{n}\frac{1}{n}\left\lvert\frac{y_i - \hat{y}_i}{y_i}\right\rvert \times 100$	0
Coefficient of determination (R^2)	$R^2 = \frac{\sum_{i=1}^{n}\left(y_i - \underline{y}_i\right)^2 - \sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}\left(y_i - \underline{y}_i\right)^2}$	1
Nash-Sutcliffe Efficiency (*NS*)	$NS = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}\left(y_i - \underline{y}_i\right)^2}$	1
Variance Account Factor (*VAF*)	$VAF = \left(1 - \frac{var(y_i - \hat{y}_i)}{var(y_i)}\right) \times 100$	100
Performance Index (*PI*)	$PI = adj.R^2 + (0.01 \times VAF) - RMSE$	2

Note: The *i*th measured ultimate load factor is denoted by y_i, the *i*th model projected ultimate load factor is represented by $\hat{y}_i$, and the average value of the measured ultimate load factor is represented by $\underline{y}_i$. The total number of data sets under examination is denoted by n.

Table 3 Statistical data analysis results

Statistics	d_b	l_b	E_f	F_f	f'_c	P_{max}
Minimum	9.53	3.3	35.74	540	20.67	3.11
Maximum	285	21.25	200	2800	114.34	121.79
Mean	62.65	11.06	68.21	1290.92	46.90	37.26
Standard Error	1.01	0.09	1.10	15.03	0.78	0.62
Median	50	10.1	50.67	1100	36.114	33.437
Mode	50	12	44	1100	27.768	26.03
Standard Deviation	33.49	3.02	36.46	498.84	25.75	20.73
Sample Variance	1121.78	9.09	1329.06	248841.22	662.95	429.70
Kurtosis	15.16	1.01	0.43	0.30	0.36	0.50
Skewness	3.29	0.51	1.42	1.13	1.24	0.88

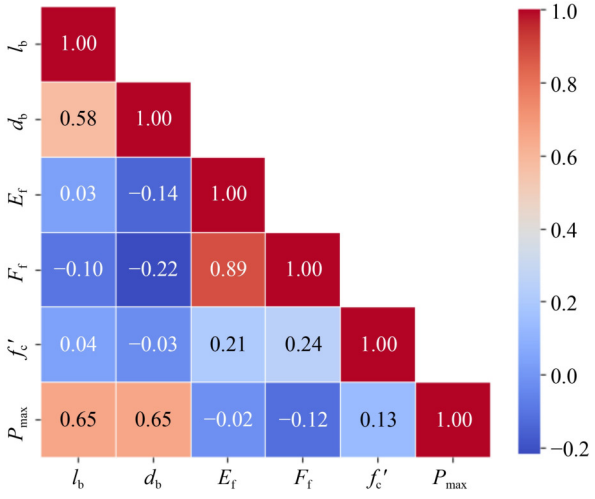

Fig. 2 Pearson correlation heatmap.

$$\gamma_n = \frac{\gamma_{act} - \gamma_{min}}{\gamma_{max} - \gamma_{min}}, \quad (4)$$

where γ_{act} represents the parameter's actual value, γ_{max} and γ_{min} represent the highest and lowest values of the parameters, respectively, and γ_n represents the normalized value of the parameters.

5.2 Hyperparameter tuning

The model's performance is directly impacted by the hyperparameters of each algorithm, which control the speed and types of patterns the model detects and analyzes. To enhance the hyperparameters in this study, a random search (RS) technique was applied. In the RS method, random combinations are selected to generate a grid of hyperparameter values from previously provided hyperparameter values for the purpose of training and grading the model. First, RS was used to find appropriate values for the hyperparameters in this research study. After that, the hyperparameter configurations are manually checked and modified to obtain a better performance value for each model algorithm. The RF, XGBoost, ADABoost, and CATBoost hyperparameters are generates by tuning strategy used in this study. Table 4 lists the optimal hyperparameters for the RF, XGBoost, ADABoost, and CATBoost models. Figure 3 shows the approach flowchart, which includes the data preparation, model selection, analysis of the acquired results and conclusion.

Table 4 Hyperparameter tuning

Method	Hyperparameter	Score
Random Forest	random_state	1
XGBoost	random_state	1
	learning_rate	0.1
ADABoost	learning_rate	0.0001
	*n*_estimators	100
	random_state	1
CATBoost	learning_rate	0.5
	12_leaf_reg	8
	iterations	200
	depth	8

6 Results and discussion

6.1 Comparison of predictive performance

In this section, comparative analysis has been performed to assess the efficiency of proposed four different model namely RF, XGBoost, ADABoost, and CATBoost. The pullout test data are taken into consideration from earlier published literature to create the maximum load (P_{max}) prediction model. Seventy percent of the entire data set (770 data points) available for model development, which is the training phase, is used in this suggested work. The performance of the proposed model was later confirmed using the remaining 30% of the data (331 data points),

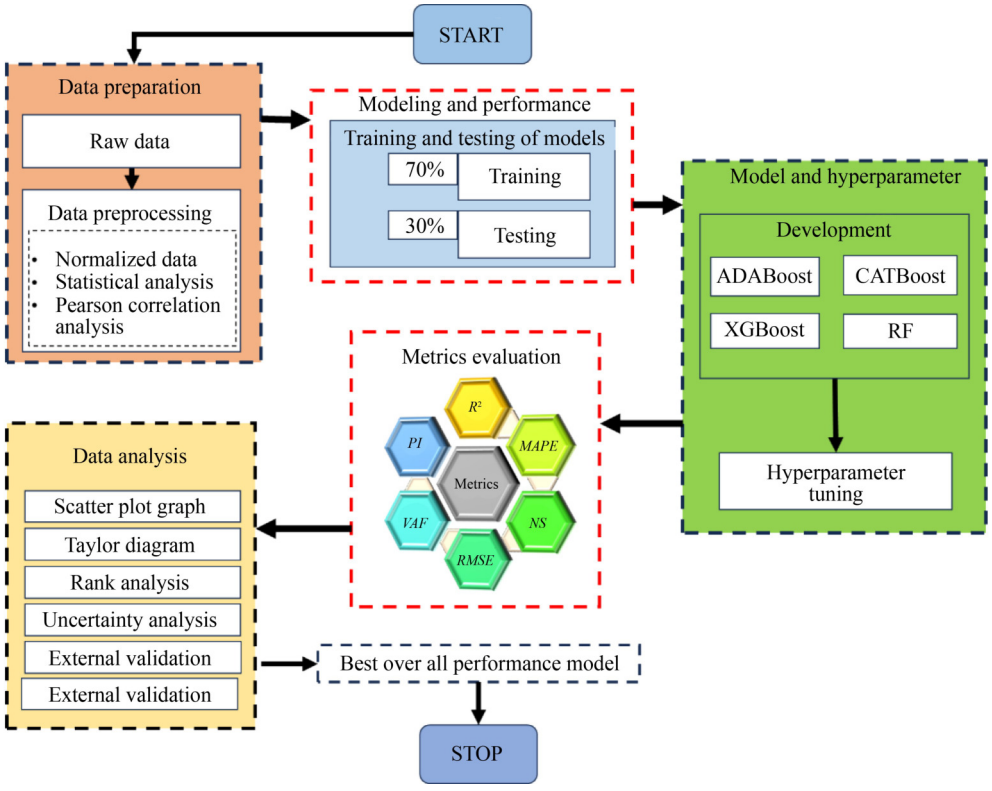

Fig. 3 Methodology flowchart.

which was the testing phase. Every ensemble regression model that was suggested was built using the Google Colaboratory and relied on the Python machine learning packages. This variant was built with an Intel Core i7 processor at 4.0 GHz with 8 GB of RAM and an 8550U CPU. Ensemble approaches simply require a minimum amount of processing capacity, in contrast to deep learning models. They can execute computations without relying on GPU servers [32].

Table 5 shows the metric performance of the ML models. A comparison of the performances of several metrics obtained from different machine learning methods revealed that XGBoost was the best model for prediction in the training phase, with an R^2 value of 0.99997; second-rank models were obtained from the RF model, with an R^2 value of 0.99996; and CATBoost and ADABoost were the third and fourth-rank models, with R^2 values of 0.99986 and 0.99952, respectively. However, in the testing phase, ADABoost achieved the best prediction performance, with R^2 value of 0.99942; second models were obtained from the XGBoost model, with an R^2 value of 0.99931; and RF and CATBoost in third and fourth place, with R^2 values of 0.99907 and 0.99570, respectively.

6.2 Measured scatter plot of the model-predicted data

A Taylor diagram was used to evaluate and examine each proposed model performance on both the testing and training data sets. The performance of machine learning models is compared using a graphical representation known to be a Taylor chart or plot. Taylor [62] developed

Table 5 Metric performance of ML models

Model	Set	R^2	*MAPE*	*NS*	*RMSE*	*VAF*	*PI*
ADABoost	Train	0.99952	0.63871	0.999516	0.00380	99.95166	1.99523
	Test	0.99942	1.10412	0.999391	0.00440	99.93913	1.99441
CATBoost	Train	0.99986	0.790313	0.999857	0.002065	99.98574	1.99765
	Test	0.99570	4.27653	0.995522	0.011927	99.55702	1.97933
XGBoost	Train	0.99997	0.329979	0.999971	0.000927	99.99713	1.99902
	Test	0.99931	1.370191	0.999236	0.004927	99.92535	1.99364
RF	Train	0.99996	0.166245	0.999954	0.001173	99.99540	1.99874
	Test	0.99907	0.976125	0.999029	0.005555	99.90301	1.99256

this technique to demonstrate and contrast the capabilities of various machine-learning models. Figures 4 and 5 display the Taylor chart results, which represent the whole performance of the model over the testing and also the training phases.

In this section, a scatter plot is created with the line $y = x$ connecting the measured and predicted ultimate load (P_{max}) factors. A point on the line ($x = y$) represents a perfect forecast of the technique's ability, while a forecast is nearer to the line ($x = y$) denotes a more accurate model. For the training and testing data displayed on Fig. 6, the estimator results of the ensemble methods of XGBoost produce cleaner findings and are less scattered from the line $x = y$. Model quality was increased, and overfitting was decreased because of XGBoost's boosting schemes. The ability of XGBoost to assign regularizations, allow it to focus on improving performance in instances that are difficult for other models to access. By using the decision stump model, XGBoost can provide good generalizability without the risk of overfitting.

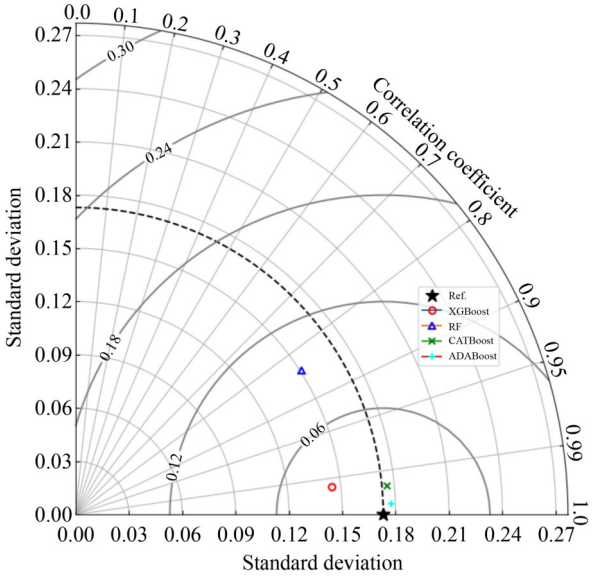

Fig. 4 Taylor chart testing phase.

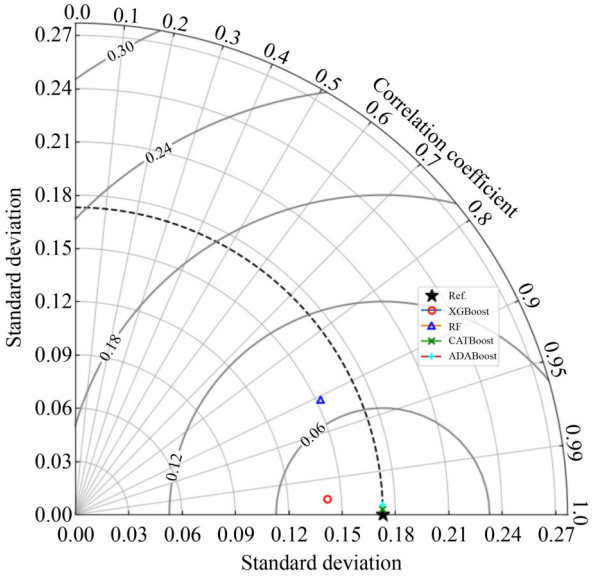

Fig. 5 Taylor chart training phase.

6.3 Rank analysis

Rank analysis is a method used to assess a model's performance. The model generates scores based on the performance metrics that are gathered throughout the training and testing stages. A higher score indicates a stronger relationship within a certain category, whereas lower results may indicate uncertainty. The technique that has the greatest overall score is ranked first, while the technique with the lowest overall score is placed in the last rank. This technique helps in assessing the capacity of suggested models for prediction. Table 6 presents the rank analysis findings for all the models provided in this study. With a total possible score of 40, the XGBoost technique outperformed the other techniques, with total scores of 33, 30, and 17 for the RF, ADABoost, and CATBoost models, respectively.

6.4 Uncertainty analysis

Uncertainty analysis was performed during both training and testing phases to evaluate the reliability of the predictive models. This process quantifies uncertainties arising from factors such as intrinsic randomness, input variation, and hyperparameter adjustment, providing insight into the robustness of the applied machine learning approaches. The mathematical Eqs. (5) and (6), mean prediction error (e) and standard deviation (σ_e), were calculated to determine the uncertainty bandwidth of the proposed models.

$$e = \frac{1}{n}\sum_{i=1}^{n} e_i, \tag{5}$$

$$\sigma_e = \sqrt{\frac{\sum_{i=1}^{n}(e_i - e)}{n-1}}, \tag{6}$$

where e_i represents the individual predicted P_{max} and n stands for the summation of databases. A negative error signifies underestimation, while a positive error reflects an overestimation of P_{max}. The 95% confidence band width of Wilson score technique was applied to measure the uncertainty bandwidth. Figures 7 and 8 depict the uncertainty bandwidths of the proposed models during

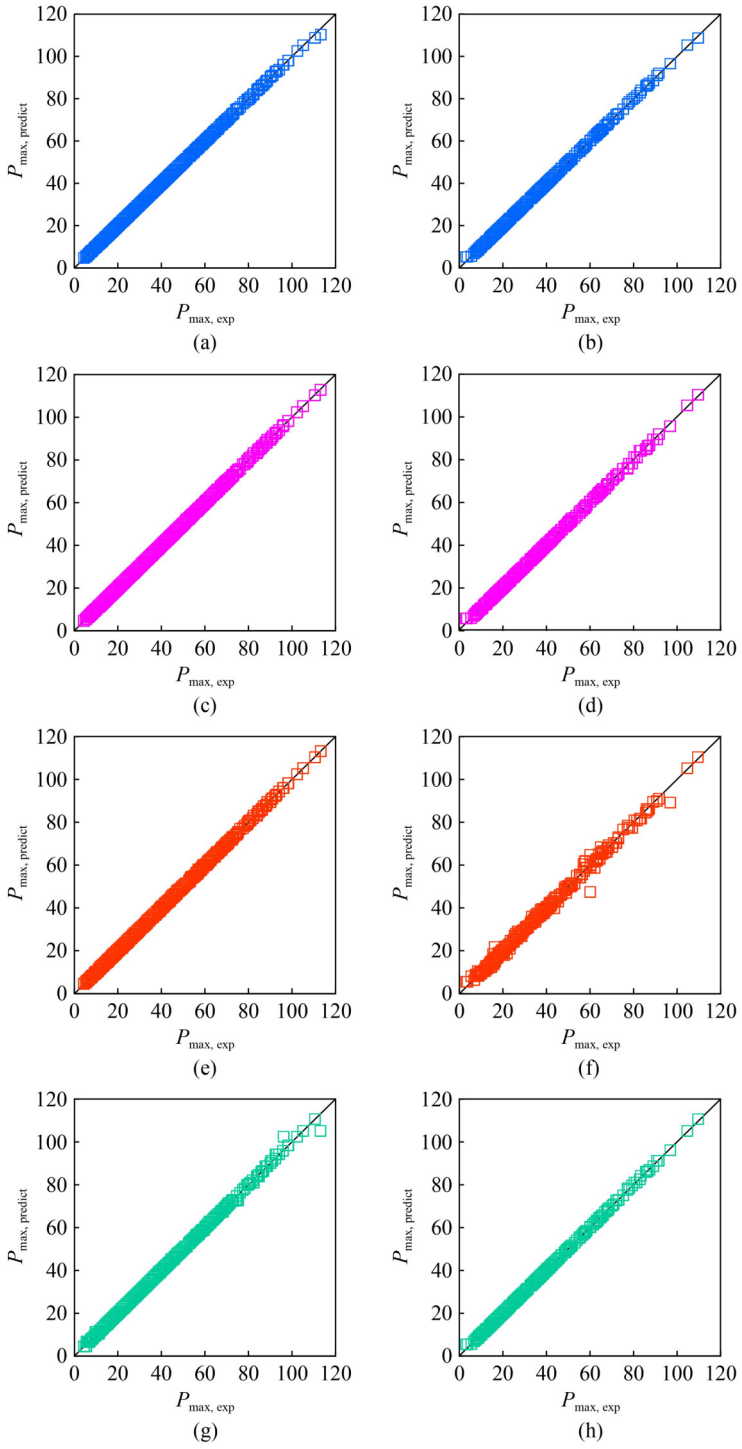

Fig. 6 Machine learning scatter plots: (a) RF training phase; (b) RF testing phase; (c) XGBoost training phase; (d) XGBoost testing phase; (e) CATBoost training phase; (f) CATBoost testing phase; (g) ADABoost training phase; (h) ADABoost testing phase.

training and testing stages. A narrower uncertainty bandwidth indicates improved predictive capability of the ML model. XGBoost model shows the narrower bandwidth indicate the best predictive capability.

The findings in Table 7 show that the XGBoost model has a smaller bandwidth (The bounds were 0.215 (upper) and −0.216 (lower) during training, and 1.044 (upper) and −1.222 (lower) during testing) compared to the other models. This indicates that XGBoost provides more consistent and accurate predictions than RF, ADABoost, and CATBoost, making it the most effective model for this analysis.

6.5 External validation

External validation is crucial for ensuring the reliability

Table 6 Rank analysis result

Metric Performance	ADABoost		CATBoost		XGBoost		RF	
	Train	Test	Train	Test	Train	Test	Train	Test
R^2								
Value	0.99952	0.99942	0.99986	0.99570	0.99997	0.99931	0.99996	0.99907
Score	1	4	2	1	4	3	3	2
MAPE								
Value	0.63871	1.10412	0.79031	4.27653	0.32998	1.37019	0.16625	0.97613
Score	2	3	1	1	3	2	4	4
NS								
Value	0.99952	0.99939	0.99986	0.99552	0.99997	0.99924	0.99995	0.99903
Score	1	4	2	1	4	3	3	2
RMSE								
Value	0.00380	0.00440	0.00207	0.01193	0.00093	0.00493	0.00117	0.00555
Score	1	4	2	1	4	3	3	2
VAF								
Value	99.95166	99.93913	99.98574	99.55702	99.99713	99.92535	99.99540	99.90301
Score	1	4	2	1	4	3	3	2
PI								
Value	1.99523	1.99441	1.99765	1.97933	1.99902	1.99364	1.99874	1.99255
Score	1	4	2	1	4	3	3	2
Subtotal	7	23	11	6	23	17	19	12
Total	30		17		40		33	
Rank	Third rank		Fourth rank		First rank		Second rank	

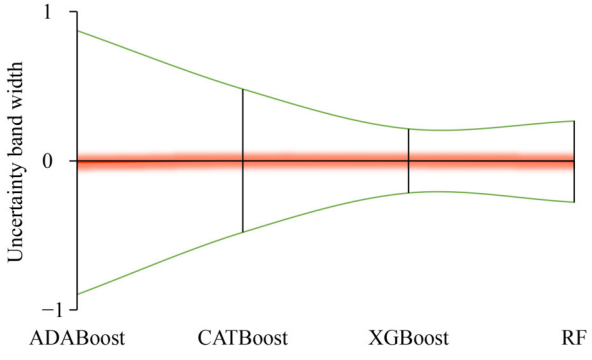

Fig. 7 Uncertainty analysis training stage.

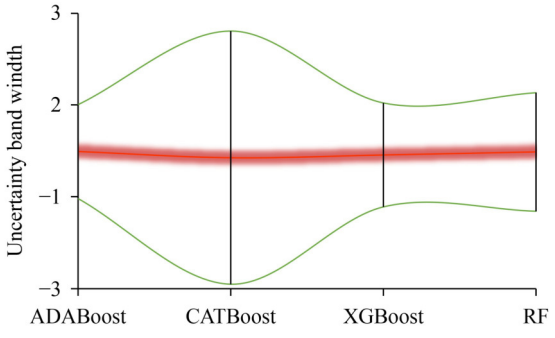

Fig. 8 Uncertainty analysis testing stage.

of prediction methods in estimating P_{max}. Proper methodological considerations must be taken into account when designing or evaluating an external validation study. In this research, ML models were tested through external validation to identify the most accurate and dependable model for predicting P_{max}. The symbols of k and k' represent the regression line slopes for the experimental and predicted values of the P_{max}. If either of the regression line slopes (k or k') is close to one at the origin, its slope should approximate 1. The coefficient of determination for the experimental value compared to ML projected value and vice versa, are represented as R'^2 and R''^2, respectively. Additionally, the relative error of ML experimental (m) and relative error of Ml prediction (n) performance indices must remain below the 0.1 threshold to meet validation criteria. The confirmed indicator (R_m) should exceed 0.5, following the recommendations of Jitchaijaroen et al. [63]. The external validation equations and results under optimal conditions are summarized in Table 8. The findings confirm that the proposed techniques are effective for predicting P_{max}. In particular, the XGBoost, RF, and ADABoost models have been successfully validated for P_{max} estimation. While

Table 7 Results of uncertainty analysis

Model	Stage	e	σ_e	Uncertainty bandwidth	Upper bound	Lower bound
ADABoost	Train	–0.012	0.451	0.884	0.872	–0.897
	Test	–0.016	0.522	1.023	1.006	–1.039
CATBoost	Train	0.001	0.245	0.480	0.481	–0.480
	Test	–0.147	1.408	2.759	2.612	–2.907
XGBoost	Train	–0.001	0.110	0.216	0.215	–0.216
	Test	–0.089	0.578	1.133	1.044	–1.222
RF	Train	–0.006	0.139	0.273	0.267	–0.278
	Test	–0.025	0.659	1.291	1.266	–1.316

Table 8 Result of external validation

Validation parameters equation	Ideal condition	ADABoost		CATBoost		XGBoost		RF	
		Train	Test	Train	Test	Train	Test	Train	Test
$k = \frac{\sum_{i=0}^{n}(a_i \times y_i)}{\sum_{i=0}^{n} y_i^2}$	$0.85 < k < 1.15$	1.001	1.002	1.000	1.006	1.000	1.004	1.000	1.002
$k' = \frac{\sum_{i=0}^{n}(a_i \times y_i)}{\sum_{i=0}^{n} a_i^2}$	$0.85 < k < 1.15$	0.999	0.998	1.000	0.993	1.000	0.996	1.000	0.998
$R'^2 = 1 - \frac{\sum_{i=1}^{n} y_i^2 (1-k)^2}{\sum_{i=1}^{n}(y_i - \bar{y})}$	Closer to 1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$R''^2 = 1 - \frac{\sum_{i=1}^{n} a_i^2 (1-k')^2}{\sum_{i=1}^{n}(a_i - \bar{a})}$	Closer to 1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$R_m = R^2 \times \left(1 - \sqrt{\lvert R^2 - R'^2 \rvert}\right)$	$R_m > 0.5$	0.978	0.975	0.988	0.930	0.995	0.973	0.993	0.969
$m = \frac{R^2 - R'^2}{R^2}$	$\lvert m \rvert < 0.1$	0.000	–0.001	0.000	–0.004	0.000	–0.001	0.000	–0.001
$n = \frac{R^2 - R''^2}{R^2}$	$\lvert n \rvert < 0.1$	0.000	–0.001	0.000	–0.004	0.000	–0.001	0.000	0.001

Note: a and actual y denotate the experimental and predicted value of P_{max}, respectively. The average experimental value and average predicted value are symbolized by and $\bar{a}$ and $\bar{y}$, respectively.

CATBoost model ranked lowest, it still met all the necessary criteria and was deemed acceptable.

6.6 Feature Importance (FI)

A key method for interpreting the outcomes of ML techniques is post-training evaluation, with FI analysis serving as a crucial approach in this process [64]. FI analysis quantifies the impact of each input feature (independent parameter) to the predicted output, helping to identify the most influential variables in the technique [65]. In this study, feature importance analysis using Gini importance (derived from tree-based models) was employed. The FI analysis results for each test are illustrated in Fig. 9. The most influential parameter in XGBoost, ADABoost, and RF is l_b, with FI percentages of 50.81%, 42.55%, and 50.80%, respectively. In contrast, for CATBoost, the highest-ranking parameter is F_f, with a contribution of 26.78%. Using FI analysis in model inference facilitates a clearer understanding of cause-and-effect relationships between key data properties and outcomes, providing insights into the decision-making process of the technique

7 Conclusions

In this research study, advanced regression ensemble machine learning models, namely, ADABoost, XGBoost, CATBoost, and random forest, were successfully applied to calculate the ultimate load (P_{max}) based on FRP-concrete pull-out test data. The 1101 case histories of FRP-concrete pull-out test data gathered from several literature sources were used to train and test in the suggested machine learning models, as previously mentioned. Five input variables, namely, the bar embedment length (l_b), diameter bar (d_b), tensile bar (F_f), modulus elastic strength (E_f), concrete compressive strength (f_c'), and one output parameter, the ultimate load (P_{max}), were adopted. Performance metrics such as the R^2, *RMSE*, *NS*, *PI*, *VAF*, and *MAPE* were computed to compare the accuracy of the suggested machine learning models. Afterward, rank analysis was used to evaluate the

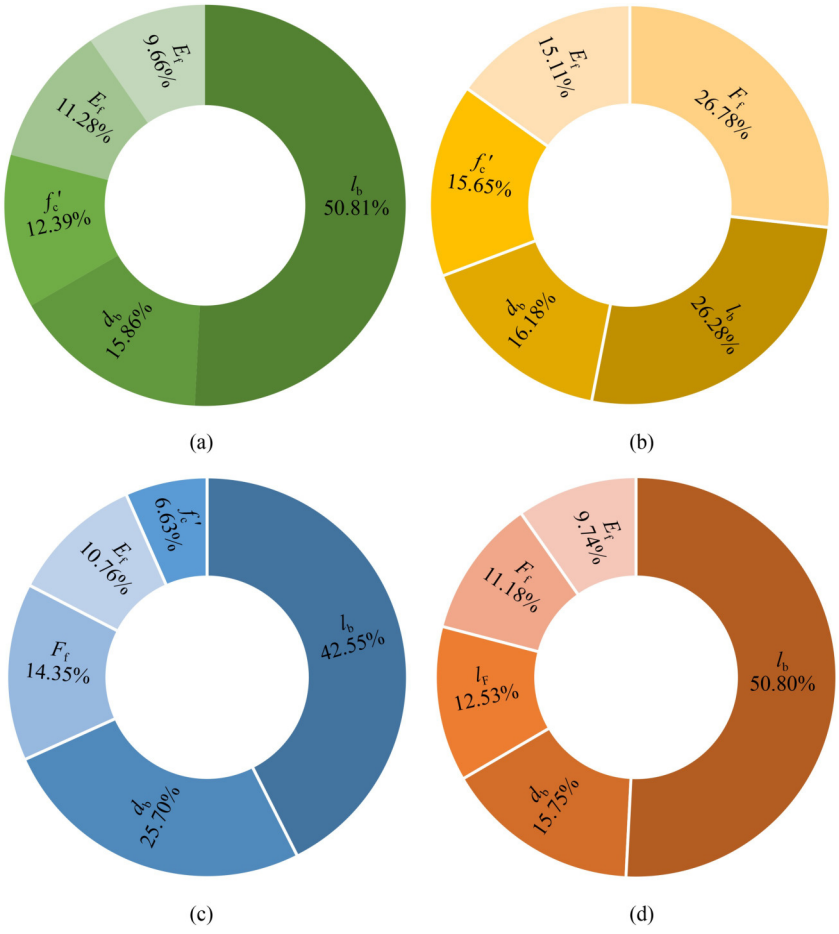

Fig. 9 Feature importance analysis: (a) ADABoost; (b) CATBoost; (c) XGBoost; (d) RF.

metric performances of each model. Among the four machine learning models, XGBoost model was identified as the best performing model based on rank analysis performance, followed by RF, ADABoost, and CATBoost, respectively. Therefore, the FI also applied to determine the influential parameter. The most influential parameter in XGBoost, RF, and ADABoost is l_b, with FI percentages of 42.55%, 50.80%, and 50.81%, respectively. In contrast, for CATBoost, the highest-ranking parameter is F_f, with a contribution of 26.78%.

Future research must address some of this study's weaknesses, however. Future studies could use the model on larger data sets when predicting the ultimate load (P_{max}) to improve the accuracy of the models that have been suggested. Therefore, using other machine learning (integration of physics-informed neural networks (PINNs) and other deep machine learning), comparing ML models with proposed equation, adding more input parameters (surface treatment, type of bar, etc.), and applying sensitivity analysis in future work could offer a clearer understanding of the importance of various input parameters the next stages of this research may be able to predict the ultimate load (P_{max}).

Acknowledgements First author (Irwan Afriadi) would like to thank to Thammasat International Student Recuitment (TISR). Also, this work was supported by Thammasat University Research Unit in Structural and Foundation Engineering.

References

1. Irshidat M R. Improved bond behavior between FRP reinforcing bars and concrete with carbon nanotubes. Construction & Building Materials, 2020, 257: 119562
2. Rolland A, Quiertant M, Khadour A, Chataigner S, Benzarti K, Argoul P. Experimental investigations on the bond behavior between concrete and FRP reinforcing bars. Construction & Building Materials, 2018, 173: 136–148
3. Zhang S Y, Chen S Z, Jiang X, Han W S. Data-driven prediction of FRP strengthened reinforced concrete beam capacity based on interpretable ensemble learning algorithms. Structures, 2022, 43: 860–877
4. Wu Y, Zhou Z, Yang Q, Chen W. On shear bond strength of FRP-concrete structures. Engineering Structures, 2010, 32(3): 897–905
5. El Refai A, Ammar M A, Masmoudi R. Bond performance of basalt fiber-reinforced polymer bars to concrete. Journal of Composites for Construction, 2015, 19(3): 04014050

6. Baena M, Torres L, Turon A, Barris C. Experimental study of bond behaviour between concrete and FRP bars using a pull-out test. Composites. Part B, Engineering, 2009, 40(8): 784–797
7. Chen L, Liang K, Shan Z. Experimental and theoretical studies on bond behavior between concrete and FRP bars with different surface conditions. Composite Structures, 2023, 309: 116721
8. Godat A, Aldaweela S, Aljaberi H, Al Tamimi N, Alghafri E. Bond strength of FRP bars in recycled-aggregate concrete. Construction & Building Materials, 2021, 267: 120919
9. Nepomuceno E, Sena-Cruz J, Correia L, D'Antino T. Review on the bond behavior and durability of FRP bars to concrete. Construction & Building Materials, 2021, 287: 123042
10. Wang W C, Nguyen N M, Cao M T. Smart ensemble machine learner with hyperparameter-free for predicting bond capacity of FRP-to-concrete interface: Multi-national data. Construction & Building Materials, 2022, 345: 128158
11. Liu B, Vu-Bac N, Rabczuk T. A stochastic multiscale method for the prediction of the thermal conductivity of Polymer nanocomposites through hybrid machine learning algorithms. Composite Structures, 2021, 273: 114269
12. Liu B, Lu W. Surrogate models in machine learning for computational stochastic multi-scale modelling in composite materials design. International Journal of Hydromechatronics., 2022, 5(4): 336–365
13. Liu B, Vu-Bac N, Zhuang X, Lu W, Fu X, Rabczuk T. Al-DeMat: A web-based expert system platform for computationally expensive models in materials design. Advances in Engineering Software, 2023, 176: 103398
14. Liu B, Vu-Bac N, Zhuang X, Fu X, Rabczuk T. Stochastic full-range multiscale modeling of thermal conductivity of Polymeric carbon nanotubes composites: A machine learning approach. Composite Structures, 2022, 289: 115393
15. Liu B, Vu-Bac N, Zhuang X, Fu X, Rabczuk T. Stochastic integrated machine learning based multiscale approach for the prediction of the thermal conductivity in carbon nanotube reinforced polymeric composites. Composites Science and Technology, 2022, 224: 109425
16. Liu B, Lu W, Olofsson T, Zhuang X, Rabczuk T. Stochastic interpretable machine learning based multiscale modeling in thermal conductivity of Polymeric graphene-enhanced composites. Composite Structures, 2024, 327: 117601
17. Shishegaran A, Saeedi M, Mirvalad S, Korayem A H. Computational predictions for estimating the performance of flexural and compressive strength of epoxy resin-based artificial stones. Engineering with Computers, 2023, 39(1): 347–372
18. Shishegaran A, Khalili M R, Karami B, Rabczuk T, Shishegaran A. Computational predictions for estimating the maximum deflection of reinforced concrete panels subjected to the blast load. International Journal of Impact Engineering, 2020, 139: 103527
19. Shishegaran A, Saeedi M, Kumar A, Ghiasinejad H. Prediction of air quality in Tehran by developing the nonlinear ensemble model. Journal of Cleaner Production, 2020, 259: 120825
20. Shishegaran A, Varaee H, Rabczuk T, Shishegaran G. High correlated variables creator machine: Prediction of the compressive strength of concrete. Computers & Structures, 2021, 247: 106479
21. Shishegaran A, Boushehri A N, Ismail A F. Gene expression programming for process parameter optimization during ultrafiltration of surfactant wastewater using hydrophilic polyethersulfone membrane. Journal of Environmental Management, 2020, 264: 110444
22. Naghsh M A, Shishegaran A, Karami B, Rabczuk T, Shishegaran A, Taghavizadeh H, Moradi M. An innovative model for predicting the displacement and rotation of column-tree moment connection under fire. Frontiers of Structural and Civil Engineering, 2021, 15(1): 194–212
23. Karami B, Shishegaran A, Taghavizade H, Rabczuk T. Presenting innovative ensemble model for prediction of the load carrying capacity of composite castellated steel beam under fire. Structures, 2021, 33: 4031–4052
24. Shishegaran A, Karami B, Safari Danalou E, Varaee H, Rabczuk T. Computational predictions for predicting the performance of steel 1 panel shear wall under explosive loads. Engineering Computations, 2021, 38(9): 3564–3589
25. Shishegaran A, Moradi M, Naghsh M A, Karami B, Shishegaran A. Prediction of the load-carrying capacity of reinforced concrete connections under post-earthquake fire. Journal of Zhejiang University. Science A, 2021, 22(6): 441–466
26. Bigdeli A, Shishegaran A, Naghsh M A, Karami B, Shishegaran A, Alizadeh G. Surrogate models for the prediction of damage in reinforced concrete tunnels under internal water pressure. Journal of Zhejiang University. Science A, 2021, 22(8): 632–656
27. Tang Y, Zhou C, Wang W, Zhang W, Cheng Y. Machine learning driven bond performance prediction between FRP bars and coral aggregate concrete. Construction & Building Materials, 2024, 442: 137684
28. Basaran B, Kalkan I, Bergil E, Erdal E. Estimation of the FRP-concrete bond strength with code formulations and machine learning algorithms. Composite Structures, 2021, 268: 113972
29. Zhang F, Wang C, Liu J, Zou X, Sneed L H, Bao Y, Wang L. Prediction of FRP-concrete interfacial bond strength based on machine learning. Engineering Structures, 2023, 274: 1–15
30. Barkhordari M S, Jawdhari A. Machine learning based prediction model for plastic hinge length calculation of reinforced concrete structural walls. Advances in Structural Engineering, 2023, 26(9): 1714–1734
31. Zhou Y, Zheng S, Huang Z, Sui L, Chen Y. Explicit neural network model for predicting FRP-concrete interfacial bond strength based on a large database. Composite Structures, 2020, 240: 111998
32. Kim B, Lee D E, Hu G, Natarajan Y, Preethaa S, Rathinakumar A P. Ensemble machine learning-based approach for predicting of FRP–concrete interfacial bonding. Mathematics, 2022, 10(2): 231
33. Su M, Peng H, Yuan M, Li S. Identification of the interfacial cohesive law parameters of FRP strips externally bonded to concrete using machine learning techniques. Engineering Fracture Mechanics, 2021, 247: 107643
34. Abuodeh O R, Abdalla J A, Hawileh R A. Prediction of shear strength and behavior of RC beams strengthened with externally bonded FRP sheets using machine learning techniques. Composite Structures, 2020, 234: 111698
35. Su M, Zhong Q, Peng H, Li S. Selected machine learning approaches for predicting the interfacial bond strength between

FRPs and concrete. Construction & Building Materials, 2021, 270: 121456
36. Yuan C, He C, Xu J, Liao L J, Kong Q Z. Bayesian optimization for selecting efficient machine learning regressors to determine bond-slip model of FRP-to-concrete interface. Structures, 2022, 39: 351–364
37. Zhang R, Xue X. A predictive model for the bond strength of near-surface-mounted FRP bonded to concrete. Composite Structures, 2021, 262: 113618
38. Başaran B, Kalkan İ, Beycioğlu A, Kasprzyk I. A review on the physical parameters affecting the bond behavior of FRP bars embedded in concrete. Polymers, 2022, 14(9): 1796
39. Yoo S J, Hong S H, Yoon Y. Bonding behavior and prediction of helically ribbed CFRP bar embedded in ultra high-performance concrete (UHPC). Case Studies in Construction Materials, 2023, 19: e02253
40. Shan Z, Liang K, Chen L. Bond behavior of helically wound FRP bars with different surface characteristics in fiber-reinforced concrete. Journal of Building Engineering, 2023, 65: 105504
41. Solyom S, Di Benedetti M, Balázs G L. Bond of FRP bars in air-entrained concrete: Experimental and statistical study. Construction & Building Materials, 2021, 300: 124193
42. Taha A, Alnahhal W, Alnuaimi N. Bond durability of basalt FRP bars to fiber reinforced concrete in a saline environment. Composite Structures, 2020, 243: 112277
43. F. M. Fahmy M, A. S. Ahmed S E, Wu Z S. Bar surface treatment effect on the bond-slip behavior and mechanism of basalt FRP bars embedded in concrete. Construction & Building Materials, 2021, 289: 122844
44. Zhao D, Zhou Y, Xing F, Sui L, Ye Z, Fu H. Bond behavior and failure mechanism of fiber-reinforced polymer bar–engineered cementitious composite interface. Engineering Structures, 2021, 243: 112520
45. Hussain S, Khan M Z N, Khan H A. Bond performance of basalt FRP bar against aggressive environment in high-strength concrete with varying bar diameter and bond length. Construction & Building Materials, 2022, 349: 128779
46. Liu S, Bai C, Zhang J, Zhao J, Hu Q. Experimental and theoretical study on bonding performance of FRP bars-Recycled aggregate concrete. Construction & Building Materials, 2022, 361: 129614
47. Peng K D, Zeng J J, Huang B T, Huang J Q, Zhuge Y, Dai J G. Bond performance of FRP bars in plain and fiber-reinforced geopolymer under pull-out loading. Journal of Building Engineering, 2022, 57: 104893
48. Trabacchin G, Sebastian W, Zhang M. Experimental and analytical study of bond between basalt FRP bars and geopolymer concrete. Construction & Building Materials, 2022, 315: 125461
49. Zeng J J, Liao J J, Zhuge Y, Guo Y C, Zhou J K, Huang Z H, Zhang L. Bond behavior between GFRP bars and seawater sea-sand fiber-reinforced ultra-high strength concrete. Engineering Structures, 2022, 254: 113787
50. Zhou Y, Wu G, Li L, Guan Z, Guo M, Yang L, Li Z. Experimental investigations on bond behavior between FRP bars and advanced sustainable concrete. Polymers, 2022, 14(6): 1132
51. Al-Hamrani A, Alnahhal W. Bond durability performance of sand-coated BFRP bars in high-strength fiber reinforced concrete. Composite Structures, 2023, 321: 117306
52. Liang K, Chen L, Shan Z, Su R K L. Experimental and theoretical study on bond behavior of helically wound FRP bars with different rib geometry embedded in ultra-high-performance concrete. Engineering Structures, 2023, 281: 115769
53. Liu S, Bai C, Zhang J, Zhao K, Li Q, Jin G. Experimental study on bonding performance of GFRP bars-recycled aggregate concrete under sulfate attack environment. Construction & Building Materials, 2023, 379: 131231
54. Zhou W, Feng P, Lin H, Zhou P. Bond behavior between GFRP bars and coral aggregate concrete. Composite Structures, 2023, 306: 116567
55. Mohri M, Rostamizadeh A, Talwalkar A. Foundations of Machine Learning. Cambridge, MA: MIT press, 2018
56. ASTM. Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens, ASTM C39/C39M-17b. West Conshohocken, PA: ASTM, 2017
57. Guo Z. Principles of Reinforced Concrete. Waltham, MA: Butterworth-Heinemann, 2014
58. Breiman L. Random forests. Machine Learning, 2001, 45(1): 5–32
59. Huang G, Wu L, Ma X, Zhang W, Fan J, Yu X, Zeng W, Zhou H. Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. Journal of Hydrology, 2019, 574: 1029–1041
60. Abd Rahman H, Wah Y B, Bulgiba A. Comparisons of ADABOOST, KNN, SVM and logistic regression in classification of imbalanced dataset. Soft Computing in Data Science, 2015, 545: 54–64
61. Kumar D R, Samui P, Wipulanusat W, Keawsawasvong S, Sangjinda K, Jitchaijaroen W. Bearing capacity of eccentrically loaded footings on rock masses using soft computing techniques. Engineering and Science, 2023, 24: 929
62. Taylor K E. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research, 2001, 106(D7): 7183–7192
63. Jitchaijaroen W, Keawsawasvong S, Wipulanusat W, Kumar D R, Jamsawang P, Sunkpho J. 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
64. Rengasamy D, Mase J M, Kumar A, Rothwell B, Torres M T, Alexander M R, Winkler D A, Figueredo G P. Feature importance in machine learning models: A fuzzy information fusion approach. Neurocomputing, 2022, 511: 163–174
65. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 2020, 58: 82–115