



Contents lists available at **SCCE**

Journal of Soft Computing in Civil Engineering

Journal homepage: [www.jssoftcivil.com](http://www.jssoftcivil.com)



## Developing Soft-Computing Models for Simulating the Maximum Moment of Circular Reinforced Concrete Columns

Buthainah Al-kharabsheh<sup>1\*</sup>, Hasan Alqawasmeh<sup>2</sup>

1. Ph.D., Civil Engineering Department, Faculty of Engineering, Al Albayt University, 25113 Mafraq, Jordan

2. Ph.D., Civil Engineering Department, Faculty of Engineering Technology Al-Balqa Applied University, 11134 Amman, Jordan

Corresponding author: [alkharabsheh@aabu.edu.jo](mailto:alkharabsheh@aabu.edu.jo)

 <https://doi.org/10.22115/SCCE.2023.369096.1556>

### ARTICLE INFO

Article history:

Received: 08 November 2022

Revised: 24 June 2023

Accepted: 17 July 2023

Keywords:

Circular reinforced concrete columns;

Moment capacity;

Machine learning models;

Soft-computing techniques.

### ABSTRACT

There has been a significant rise in research using soft-computing techniques to predict critical structural engineering parameters. A variety of models have been designed and implemented to predict crucial elements such as the load-bearing capacity and the mode of failure in reinforced concrete columns. These advancements have made significant contributions to the field of structural engineering, aiding in more accurate and reliable design processes. Despite this progress, a noticeable gap remains in literature. There's a notable lack of comprehensive studies that evaluate and compare the capabilities of various machine learning models in predicting the maximum moment capacity of circular reinforced concrete columns. The present study addresses a gap in the literature by examining and comparing the capabilities of various machine learning models in predicting the ultimate moment capacity of spiral reinforced concrete columns. The main models explored include AdaBoost, Gradient Boosting, and Extremely Gradient Boosting. The  $R^2$  value for Histogram-Based Gradient Boosting, Random Forest, and Extremely Randomized Trees models demonstrated high accuracy for testing data at 0.95, 0.96, and 0.95, respectively, indicating their robust performance. Furthermore, the Mean Absolute Error of Gradient Boosting and Extremely Randomized Trees on testing data was the lowest at 36.81 and 35.88 respectively, indicating their precision. This comparative analysis presents a benchmark for understanding the strengths and limitations of each method. These machine learning models have shown the potential to significantly outperform empirical formulations currently used in practice, offering a pathway to more reliable predictions of the ultimate moment capacity of spiral RC columns.

How to cite this article: Al-kharabsheh B, Alqawasmeh H. Developing soft-computing models for simulating the maximum moment of circular reinforced concrete columns. *J Soft Comput Civ Eng* 2024;8(2):46-66. <https://doi.org/10.22115/scce.2023.369096.1556>

2588-2872/ © 2023 The Authors. Published by Pouyan Press.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



## 1. Introduction

Designing reinforced concrete (RC) structures requires a thorough understanding and precise estimation of the moment capacity of the columns. These columns, particularly those with spirals, are believed to possess a higher lateral capacity than their counterparts with ties, due to the greater confinement pressure they can offer [1,2]. Over the years, this hypothesis has been tested and verified through a number of experimental studies, building a robust body of literature on the behavior of these integral structural components [3–6]. Complementing this experimental approach, the last few years have seen an upsurge in the application of machine learning-based methods to predict the behavior of RC columns with interconnecting spirals. This surge has been motivated by the capacity of machine learning techniques to process vast amounts of data, adapt to new information, and provide robust predictions. For instance, studies conducted by Ahmadi et al. [7] and Chang & Zheng [8] have utilized artificial neural network models to evaluate the compressive strength of circular columns. In a similar vein, Koçer et al. [9] examined the ability of an artificial neural network model to predict the moment capacity of spirally reinforced concrete columns. More comprehensive research on this topic was carried out by Naderpour et al. [10], focusing on the accuracy of artificial neural networks in estimating the moment capacity of such columns. Going a step further, other research investigated the effectiveness of several machine learning approaches, including the artificial neural network and ensemble machine learning techniques, in solving other structural engineering problems including design, modeling, assessment, and evaluation of structural behavior and resistance to external forces [11–16]. For instance, Noori and Varaee [14] utilized artificial neural networks to approximate the nonlinear seismic response of steel moment frames, demonstrating the potential of AI-based methods in complex structural analysis. Furthermore, an investigation by Shishegaran et al. [15] offered computational predictions of steel panel shear wall performance under explosive loads, revealing crucial insights into the impacts of extreme conditions. The same research group developed a high correlated variables creator machine to predict concrete's compressive strength, showcasing the power of computational tools in materials testing [16]. Compressive strength, a key parameter of concrete, has also been the focus of numerous AI-based studies. Zhang et al. [17] proposed a novel method combining the Multivariate Adaptive Regression Splines (MARS), the Grasshopper Optimization Algorithm (GOA), and the Monte Carlo Simulation (MCS) to assess the reliability of compressive and splitting tensile strength prediction of roller compacted concrete pavement. Ashrafian et al. [18] introduced an evolutionary Neuro-Fuzzy-Based approach to estimate the compressive strength of eco-friendly concrete containing recycled construction wastes. This trend extends to the estimation of seismic retrofit costs, where Hamzehkolaei and Alizamir [19] evaluated machine learning algorithms' performance using structural parameters. On the optimization front, a hybrid generalized reduced gradient-based particle swarm optimizer was proposed by Varaee et al. [20] for addressing constrained engineering optimization problems. Pour et al. [21] investigated the performance of composite concrete-filled steel tube columns with different cross-section shapes and steel fibers, further highlighting the need for accurate computational modeling in modern construction. Despite this progress, it is apparent that most of the preceding research has been predominantly confined to the use of artificial neural networks, with limited exploration of the potential of other machine learning models. This research project aims to fill this gap by examining the performance of a

variety of machine learning models in determining the ultimate moment capacity of spirally reinforced concrete columns. The study will delve into each machine learning model's performance, analyzing and contrasting them to ascertain the most effective technique. Following this, the outcomes of the selected machine learning models will be juxtaposed with the methodology delineated in the ACI 318 standard by the American Concrete Institute. This comparative analysis will provide valuable insights into the predictive accuracy of the machine learning models relative to conventional structural engineering approaches. In addition to the comparative analysis, the study will conduct an in-depth parametric analysis using the feature importance to underscore the impact of each input variable on the model output. Feature importance is a powerful tool that enables us to identify which variables have the most significant impact on the predictions of the model, thereby shedding light on the most crucial factors that influence the moment capacity of spirally reinforced concrete columns. Accordingly, this article presents novel research focused on using soft-computing techniques to predict the maximum moment capacity of circular RC columns. To address a gap in existing literature, the researchers developed and evaluated various machine learning models for this purpose. By benchmarking these models against experimental and conventional model results, they were able to demonstrate that machine learning can provide highly reliable predictions for the ultimate moment capacity of spiral RC columns. This study further suggests that the developed machine learning models can significantly outperform the empirical formulation introduced by the codes of practice. Thus, the research not only pioneers the application of soft-computing models to this specific area of structural engineering but also highlights their potential for improving accuracy and reliability in predicting the behavior of reinforced concrete structures. This article is structured as follows: Section 2 discusses the materials and methods of the study; Section 3 provides the results and discussions of the study; Section 4 highlights the main conclusions of the paper.

## 2. Materials and methods

According to Koçer et al. [9], evaluating the moment capacity of RC columns with spirals exposed to simultaneous bending and axial movements is estimated in particles using mathematical expressions typically supplied by the local standards of each nation. Such empirical methods have the disadvantage of being limited in ability to suit a variety of column attributes. In order to solve this problem, prior studies advise employing the ANN technique to calculate the moment capacity of RC columns. Nevertheless, the research has been presented in the literature did not highlight the capability and effectiveness of various machine learning models in the estimation. Therefore, this study compares the accuracy of the three most popular neural network models to the empirical methods used in ACI 318 to solve regression problems.

### 2.1. Utilized database

In fact, the Pacific Earthquake Engineering Research Center (PEER) provides an open-source respiratory for experimental data of many circular spirally-reinforced concrete columns. This database is used in this study to develop the selected machine learning models. **Table 1** and **Figure 1** provides the descriptive statistics for this dataset.

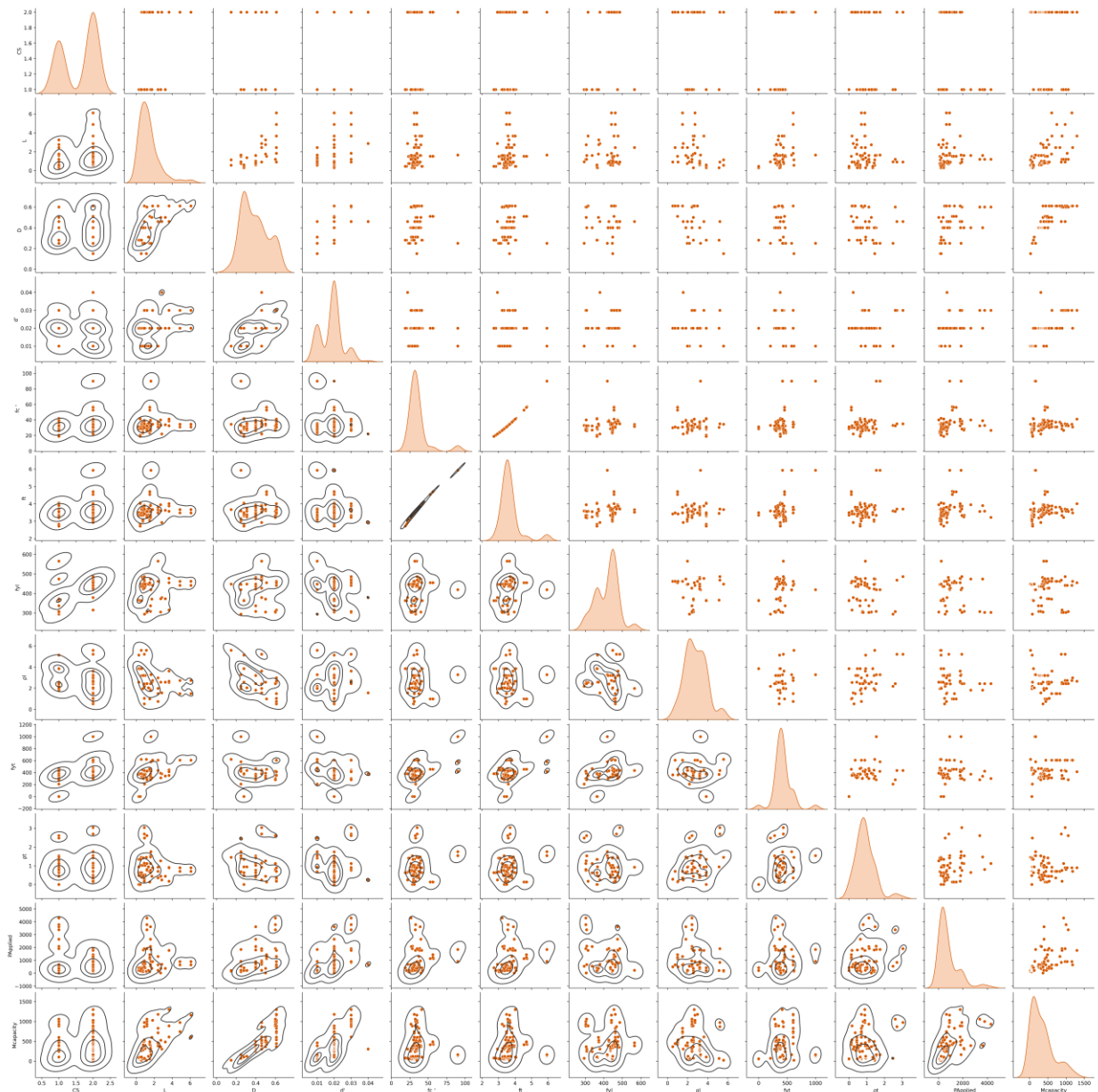
**Table 1**  
Descriptive statistics of the utilized database.

Variable	Sample Size	Mean	Standard Deviation	Variance	Coefficient of variation	Minimum	Q1	Median	Q3	Maximum
Diameter (m)	104	0.390	0.131	0.017	33.440	0.150	0.280	0.400	0.490	0.610
Cross-section	Circular sections or octagonal ones									
Concrete cover (m)	104	0.018	0.006	0.000	34.400	0.010	0.010	0.020	0.020	0.040
Element length (m)	104	1.422	1.154	1.331	81.160	0.300	0.600	1.200	1.650	6.100
Concrete compressive strength (MPa)	104	35.5	13.9	192.1	39.0	18.9	29.4	32.5	35.5	90.0
Concrete tensile strength (MPa)	104	3.676	0.609	0.370	16.560	2.720	3.388	3.560	3.720	5.930
Yield stress of longitudinal reinforcement (MPa)	104	418.2	59.3	3519.9	14.2	294.0	366.0	441.3	453.3	565.4
Longitudinal reinforcement ratio	104	2.796	1.077	1.160	38.510	0.520	2.040	2.570	3.793	5.580
Yield stress of transverse reinforcement (MPa)	104	422.3	152.1	23127.9	36.0	0.0	361.0	401.5	444.0	1000.0
Volumetric transverse reinforcement ratio	104	0.910	0.546	0.298	60.020	0.000	0.630	0.920	1.135	3.040
Axial load applied (kN)	104	798.7	846.4	716313.5	106.0	26.4	220.5	430.5	921.7	4300.0
Moment capacity (kN.m)	104	342.5	313.1	98050.3	91.4	22.0	85.8	295.0	479.3	1300.0

## 2.2. Machine learning models

The present era witnesses a prevalent use of the multiple linear regression (MLR) method in determining the linear relationship between a single dependent parameter and multiple independent factors. This mathematical model provides a compelling way of understanding the interplay between several variables and their impact on an outcome of interest. Currently, the method of stochastic gradient descent (SGD) is widely employed in large-scale MLR modeling. This method shows remarkable efficiency and delivers superior results when dealing with large datasets. SGD essentially operates by computing the gradient of the loss function for each sample in the dataset and subsequently updates the model parameters. Generally, there are a variety of options available when choosing a loss function for SGD, including the squared Euclidean norm, the absolute norm of ElasticNet, or even a combination of these two. These are all techniques aimed at reducing the coefficients of the model to a zero vector, thus minimizing the error and ensuring the best possible model fit. Another significant method utilized in engineering, particularly in regression problems, is the support vector regressor (SVR). This supervised learning algorithm was developed to handle cases where the nature of the problem made traditional methods unsuitable. SVR allows for a variety of kernel functions, which could be linear, polynomial, radial basis function (RBF), or sigmoid, each with distinct advantages. This research aims to evaluate multiple kernels to pinpoint the one that yields the most optimal results. Furthermore, the decision tree (DT) and SVR techniques have gained considerable recognition in the data mining industry. Both are extensively used for solving regression and

classification problems. DTs come with certain unique advantages such as providing an exhaustive analysis of all potential outcomes and the ability to track the path leading to each solution. The DT method is known for conducting an in-depth examination and evaluation of each outcome and its path. This feature is particularly crucial for conducting additional analysis on decision nodes. Also, the DT method operates by dividing the dataset iteratively. Each partition serves as the basis for generating the estimating model, and the decision tree itself is ultimately produced from the integration of these predictive models. In recent times, a myriad of DT methods have been proposed and developed. This study will employ the classification and regression tree (CART) technique under the DT model. This approach is known for its efficiency in handling numerical target parameters, making it a robust choice for various applications.



**Fig. 1.** Visualized descriptive statistics of the utilized database (Parameters' units are the same as that defined in Table 1).

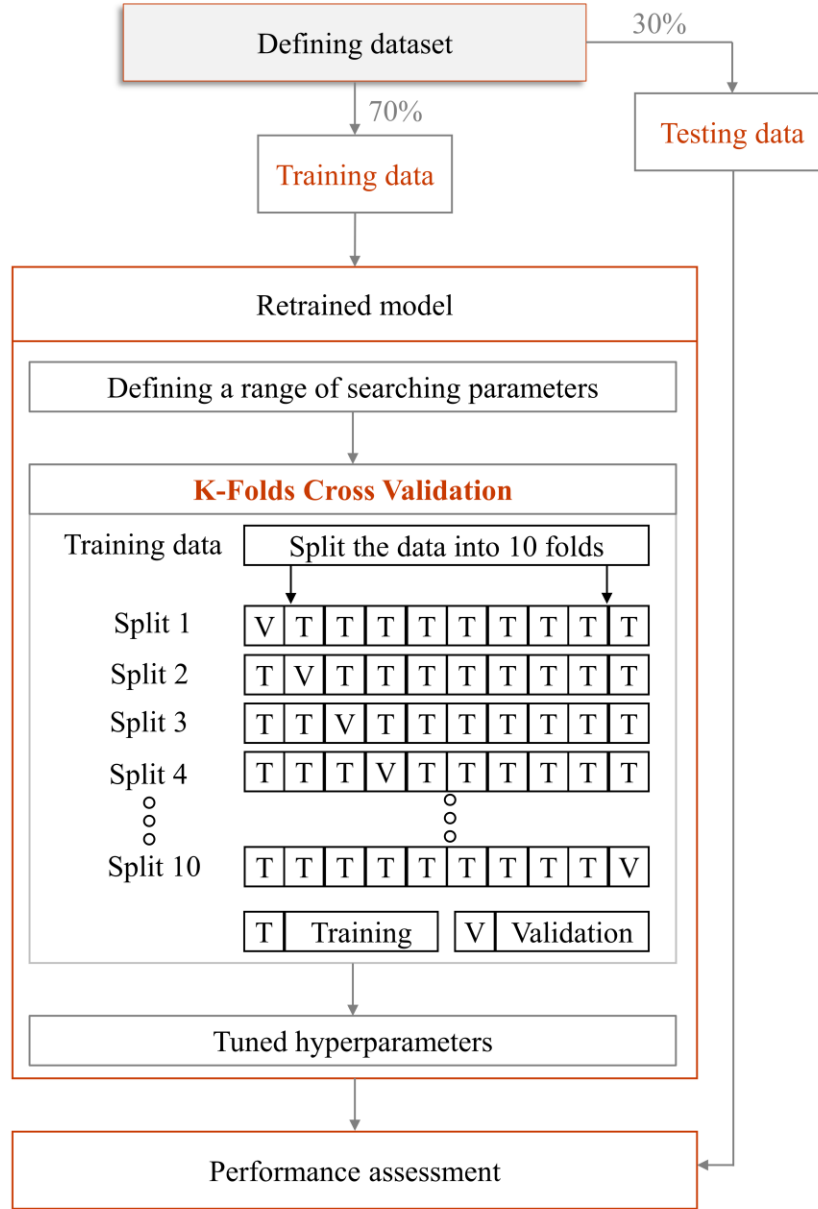
Random forest (RF) has emerged as one of the most sought-after machine learning algorithms in the literature. Known for its precision and adaptability with various datasets, RF has gained significant traction in the field of civil engineering, wherein it has been utilized to construct functional models that help in decision-making and prediction. An RF model comprises numerous decision trees (DT), each controlled by a random subset of the training data, distinct from the singular tree structure of the DT model. RF employs the concept of bootstrapping and aggregation, creating multiple Classification And Regression Trees (CARTs), contributing to the robustness of the model. A key distinction arises when comparing the RF and DT methodologies. While RF works based on critical thresholds, an extremely randomized trees (ETR) approach hinges on random determination of thresholds for each potential feature. In the ETR model, the splits within the tree's nodes are also randomly determined to identify the most suitable candidate for the splitting criterion. The ETR method significantly reduces the model's variance but tends to increase the bias. Unlike the RF, which uses bootstrap replicas (substituting input data through down-sampling), ETR examines the entire original sample. Adaptive boosting (AdaBoost or Ada) is a versatile meta-algorithm designed to enhance prediction performance. It has been implemented in a myriad of learning methodologies cited in literature. This algorithm is typically based on an iterative process, adjusting the weights when a previous experiment doesn't yield the desired results. AdaBoost uses the real training dataset and regressor to fit multiple instances of the regression model, adhering to Drucker's rules for handling complex situations. In this study, the Ada weak model is a single-level DT regressor. Stochastic gradient boosting (GB) is an enhancement of the traditional gradient boosting technique, used extensively for regression and classification tasks. The primary difference between Ada and GB lies in their weak learner models. While Ada's weak learner is a single-level regressor, GB's weak learner has a deeper decision tree. Another distinctive feature of GB is that it does not require training on the entire dataset, thereby reducing the risk of overfitting and lowering the correlation of the trees. Histogram-based gradient boosting (HGB) distinguishes itself from other machine learning techniques by assigning fixed data points into bins to create a histogram during the training phase. This strategy facilitates faster training stages, rapidly improves model quality, and reduces the model's memory requirements. Therefore, the HGB approach promises superior machine learning applications in shorter timeframes. Extreme gradient boosting (XGB), a highly flexible and effective machine learning technique, produces a series of decision trees by allocating a weight classification to each independent parameter. These weight classifications are then assigned to the decision tree to predict outcomes. In subsequent decision trees, classifications are performed for the incorrectly predicted parameters, but with higher weight allocations. Primarily, the XGB method relies on the effect of the weight, merging several predictions and classifiers to create an accurate and robust model. The XGB model is particularly effective at mitigating overfitting problems and consistently delivering superior results. A key similarity between XGB and gradient boosting is their reliance on the gradient boosting concept, which contributes to unique modeling features. These various methodologies are each powerful tools within the machine learning repertoire. However, they all have their unique characteristics and suitable

applications. Understanding these distinctions is crucial in selecting the appropriate model for a given task. Random forests, for example, excel in dealing with high-dimensional data and are immune to overfitting due to their ensemble nature. Extremely randomized trees, on the other hand, offer an even more randomized variant of the decision tree, which can prove beneficial in certain cases where an additional level of randomness can help improve generalization. Adaptive boosting, in contrast, works by adjusting weights of the instances in the data set based on the previous classification. If an instance was incorrectly classified, the model increases the weight of this instance and the next classifier in the sequence is forced to focus on this instance. This approach makes AdaBoost very responsive to changes and can therefore quickly adapt to shifting data patterns. On the other hand, stochastic gradient boosting and histogram-based gradient boosting are both extensions of the traditional gradient boosting method. Both methods iteratively add new models into the ensemble and use gradient descent to minimize the loss function. However, the way they handle data and train models is fundamentally different, which results in different strengths and weaknesses for these two methods. Lastly, extreme gradient boosting, as its name suggests, pushes the boosting technique to its limit. It is known for its efficiency and performance, often outperforming other methods on many machine learning benchmarks. While it is similar to gradient boosting in that it relies on the concept of boosting, it comes with several enhancements that reduce overfitting, improve computational speed, and allow for better model interpretation. It represents one of the most advanced machine learning techniques available today, capable of handling a wide variety of tasks with high accuracy and robustness.

### 2.3. Model development and hyperparameters tuning

During the training process, the hyperparameters are optimized using the grid search approach with 10 folds cross-validation. As a result, each model was trained extensively to reach an optimal parameter that provides the best accuracy. Figure 2 illustrates the flowchart and the models created for this method. Overall, the method starts by defining a set of search parameters for each model. After that, the dataset is split into training and testing at 70% and 30% ratios, respectively. Then the training dataset is used to train the model using the k-fold cross-validation method to overcome the overfitting issue. Once many setups are built, the best model is used to predict the testing dataset the model performance is captured for later comparison in this study.

The coefficient of determination ( $R^2$ ) was used to assess the goodness of fit for the linear regression model, where the numerator of the  $R^2$  fraction depends on the unidentified dissimilarities by the response of the independent parameters and the denominator of the  $R^2$  fraction depends on the total dissimilarities in the response, Eq. 1 [12]. The range of  $R^2$  where 1 represents the strongest linear relationship and values are between 0 and 1. Therefore, the measurement tool used for the variations between evaluated and observed values is root-mean-square error (RMSE), an error analysis method, Eq. 2. Furthermore, the variations between the absolute evaluated and observed values are the mean absolute error (MAE) a type of error analysis, Eq. 3.



**Fig. 2.** Flowchart of the approach used for developing the machine learning models' performance assessment.

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x}_i)^2} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \tag{2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}_i| \tag{3}$$

where  $x_i$  is the measured value,  $\bar{x}_i$  is the mean of the measured values,  $y_i$  is the predicted value,  $\bar{y}_i$  is the mean of the predicted values, and  $n$  is the number of observations.



## 2.4. Feature importance analysis

The inference of the model is an essential element for attaining the areas of enhancing various models, developing trust among trained algorithms, and formulating strategies for decision-making. The use of machine learning models has advanced in several ways, including the ability to do parametric evaluations by assessing the importance of each input characteristic on the prediction outcomes. Therefore, the parametric evaluation using machine learning features the importance of analysis, which has been used in many articles. Feature selection and feature extraction are two key techniques employed in machine learning to identify the most relevant features and transform them into meaningful representations, respectively. These techniques have a significant impact on feature importance analysis, which aims to understand the relative contribution of different features to the predictive power of the model. By incorporating feature selection or feature extraction into the analysis, it becomes possible to identify the most influential features that significantly impact the model's performance. This knowledge not only improves model interpretability but also guides feature engineering efforts and enables better decision-making regarding feature inclusion or exclusion. In the present study, feature importance analysis is performed to assess the effect of each input parameter on the moment capacity of circular spirally reinforced concrete columns. By understanding the importance of different input variables, it becomes possible to determine their impact on the performance of the anticipated model, specifically the dependent variable. This analysis provides valuable insights for developing more accurate and reliable models, enhancing their predictive capabilities, and formulating effective strategies for decision-making.

## 3. Results and discussions

The ERT, Ada, GB, and XGB models were found to be the top-performing machine learning models for predicting maximum moment capacity in RC structural circular columns. It should be highlighted that these models were exceptionally robust and reliable, showcasing promising precision in their predictions compared to the rest. They displayed lesser deviations from the equity line, representing a minimal disparity between their predicted and actual values. In contrast, the SGD case was the underperformer amongst all the models examined. Its predictions demonstrated a substantial deviation from the equity line, thus indicating a lack of precision and reliability.

Table 2 presents the optimized hyperparameters for seven models utilized in the study, namely DT, RF, ETR, Ada, GB, HGB, and XGB. For the DT model, the 'criterion' hyperparameter is optimized to 'absolute\_error', suggesting that the model minimizes the total absolute error when making splits. The 'max\_depth' is unrestricted, allowing the tree to expand freely, and 'splitter' is set to 'best', meaning the decision tree utilizes the optimal split at each node. The RF model utilizes 'squared\_error' as the 'criterion' to minimize total squared error when creating splits. The 'max\_depth' hyperparameter is capped at 10, limiting the tree's depth to prevent overfitting. Moreover, the 'n\_estimators' hyperparameter is optimized to 250, building a random forest comprised of 250 decision trees. The ETR model employs similar hyperparameters to the RF model, but the 'max\_depth' remains unrestricted and the model includes 100 estimators ('n\_estimators': 100). For the AdaBoost model, a DecisionTreeRegressor serves as the

'base\_estimator' and the 'learning\_rate' is established at 1.0. The 'loss' function employed is 'linear', and the number of weak learners ('n\_estimators') is set to 100. The GB model uses 'least squares' ('ls') as the loss function and a learning rate of 0.05. It also incorporates an elaborate configuration with 'max\_depth': 100, 'min\_samples\_leaf': 1, 'min\_samples\_split': 10, 'n\_estimators': 200. The 'subsample' is set to 1.0, implying that the entire dataset is utilized for training each tree. The HGB model applies 'least\_squares' as the loss function with an 'l2\_regularization' of 0.25 to prevent overfitting. The 'max\_bins' hyperparameter is set to 180, and 'max\_depth' remains unrestricted. Lastly, for the XGB model, a 'dart' booster is applied, and 'reg:squarederror' serves as the objective. The 'max\_depth' is set to 1, and the 'n\_estimators' to 1000. All 'colsample\_by\*' parameters are set to 1.0, meaning all columns are utilized at every level, node, and tree. Furthermore, specific regularization parameters, 'estimator\_\_gamma': 2 and 'estimator\_\_lambda': 0, are part of the optimized setup.

**Table 2**

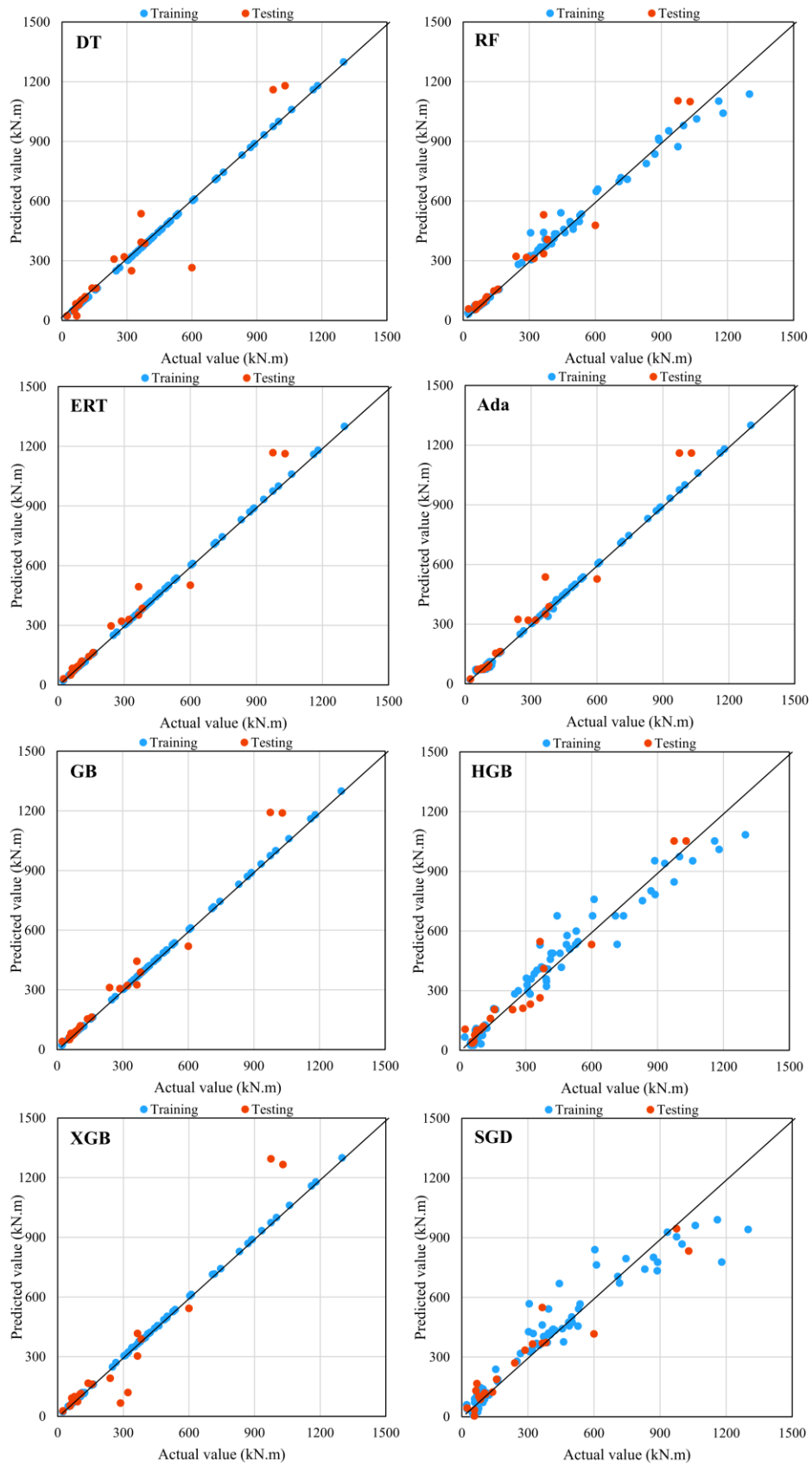
Optimized hyperparameters for the models developed in this study.

Model	Optimized Hyperparameters
DT	{'criterion': 'absolute_error', 'max_depth': None, 'random_state': 0, 'splitter': 'best'}
RF	{'criterion': 'squared_error', 'max_depth': 10, 'n_estimators': 250, 'random_state': 0}
ETR	{'criterion': 'squared_error', 'max_depth': None, 'n_estimators': 100, 'random_state': 0}
Ada	{'base_estimator': DecisionTreeRegressor(), 'learning_rate': 1.0, 'loss': 'linear', 'n_estimators': 100, 'random_state': 0}
GB	{'learning_rate': 0.05, 'loss': 'ls', 'max_depth': 100, 'max_features': None, 'min_samples_leaf': 1, 'min_samples_split': 10, 'n_estimators': 200, 'random_state': 0, 'subsample': 1.0}
HGB	{'l2_regularization': 0.25, 'learning_rate': 0.17, 'loss': 'least_squares', 'max_bins': 180, 'max_depth': None, 'random_state': 0}
XGB	{'booster': 'dart', 'colsample_bylevel': 1.0, 'colsample_bynode': 1.0, 'colsample_bytree': 1.0, 'estimator__gamma': 2, 'estimator__lambda': 0, 'learning_rate': 1.0, 'max_depth': 1, 'n_estimators': 1000, 'objective': 'reg:squarederror', 'random_state': 0}

Furthermore, an essential observation derived from Figure 3 was the performance difference between the training and testing datasets. As anticipated, all the models performed better on the training data, likely due to the direct exposure and adaptation to this dataset during the training phase. However, the difference between these performances was not substantial enough to suggest the presence of overfitting, which is a common concern in machine learning practices. This indicates that all models, including the top performers, were able to generalize their learning effectively to unseen data, an essential attribute for real-world application.

Although this analysis presents a clear comparative view of the models, it should be noted that model selection is often dependent on specific criteria or requirements of a given scenario. Certain situations might favor less computationally expensive models or models that are more interpretable, despite a small trade-off in prediction accuracy. Therefore, while the ERT, Ada, GB, and XGB models outperformed others in this study, the choice of machine learning model for practical application should be a nuanced decision. Further study and validation might also be required to verify these results in different contexts and for a more diverse range of datasets. As defined in Eq. 4, the residual analysis is used in this study to calculate the difference between the observed and predicted values and further investigate the models' performances.

$$e_i = (y_i - \hat{y}_i) \quad (4)$$



**Fig. 3.** Performance of the developed machine learning models.

Figure 4 displays the residuals of several machine learning models as they correspond to each observation in the training dataset. The term 'residual' here represents the difference between the observed and predicted values. From the figure, it's apparent that most residuals are concentrated around zero, implying a high level of accuracy in predictions for most of the models, excluding SGD, RF, and HGB. This finding suggests that these models may not perform as well as others in the given task, as the residuals are larger. Conversely, Figure 5 represents the residual plot for the test dataset. Unlike the training dataset, the residuals of the test dataset show considerable scatter for all the models. This is an expected phenomenon as testing data usually contains unseen scenarios that can be challenging to predict precisely. The DT model shows the highest negative residual at approximately -300 kN.m, which implies it has the highest under-prediction, while the XGB model recorded the maximum positive residual, around 300 kN.m, showing the highest over-prediction. It's crucial to note that both under-predictions and over-predictions can be problematic depending on the application, and a balance is often desirable. Interestingly, the HGB model demonstrated a consistent range of residuals between the training and testing datasets. This is an important observation as it suggests that the HGB model is controlling overfitting issues effectively. Overfitting occurs when a model learns the training data too well, leading to poor performance on unseen data.

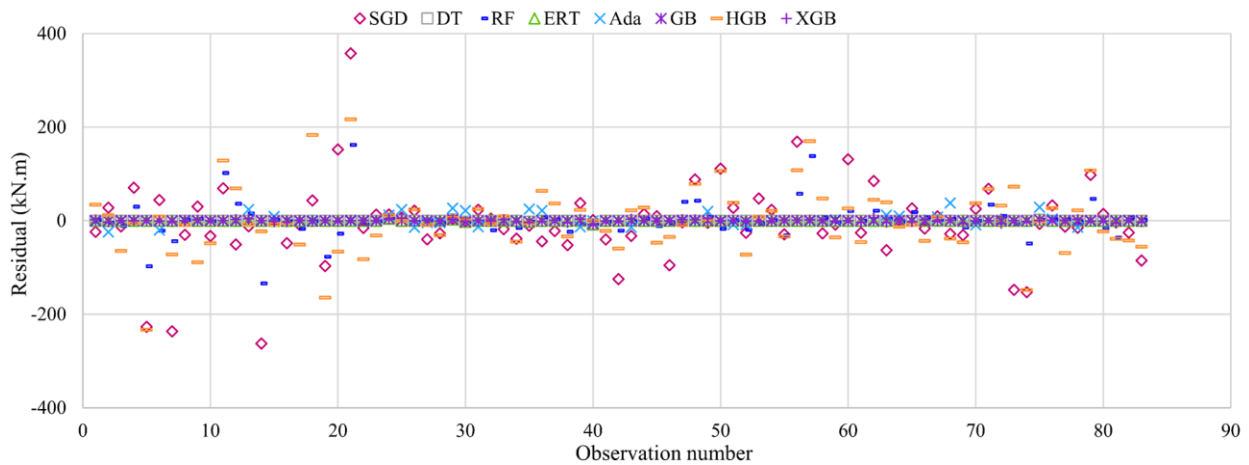


Fig. 4. Residuals of the developed machine learning models for the training dataset.

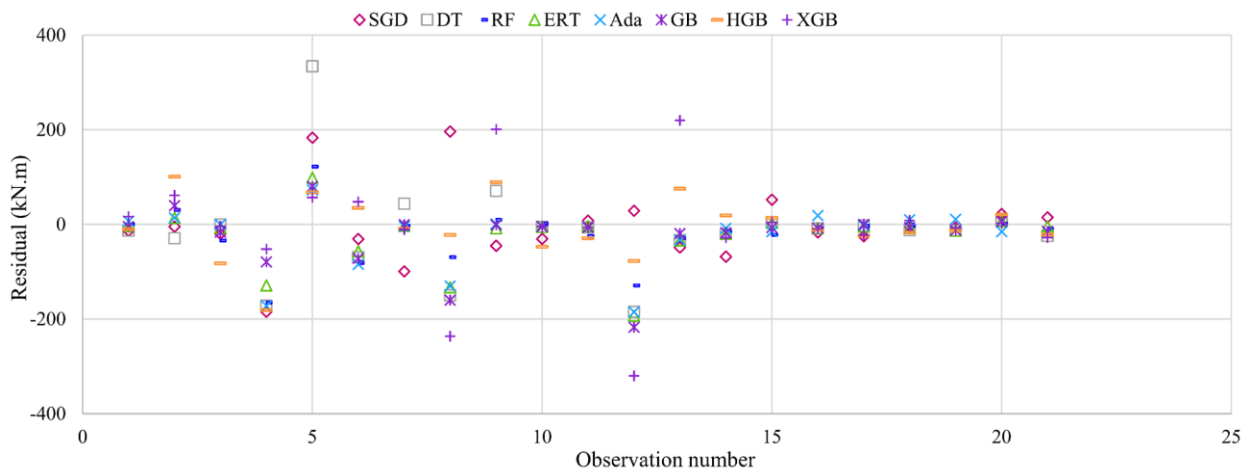
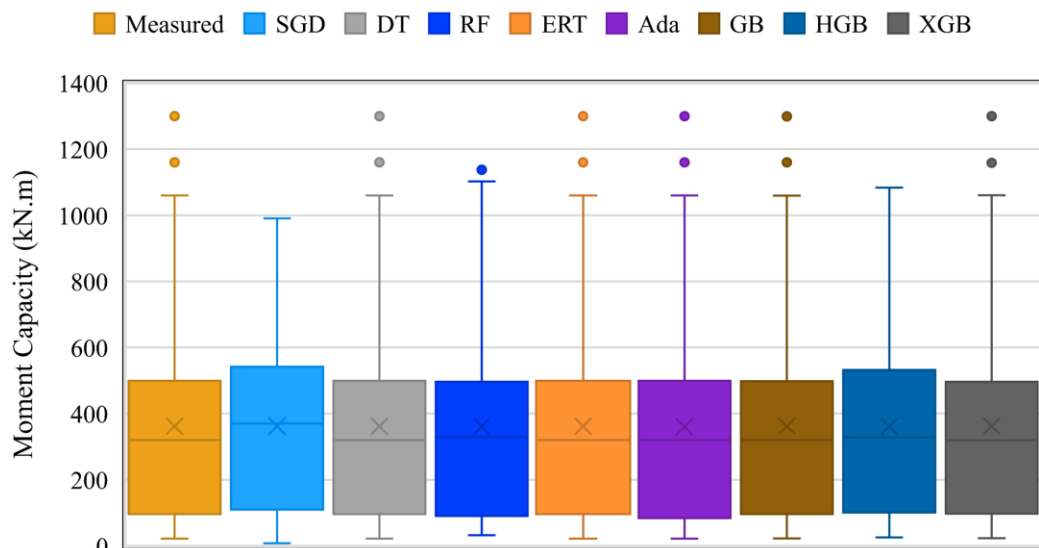


Fig. 5. Residuals of the developed machine learning models for the testing dataset.

The ability to produce a similar range of residuals in both datasets indicates that the HGB model has achieved a good balance between learning the data and generalizing to unseen scenarios. This is an important quality in a machine learning model and shows the robustness of HGB in this context. However, these observations should be considered in conjunction with other performance metrics to get a comprehensive understanding of model performance. The importance of residual magnitude can vary based on the context and practical implications of prediction errors in a particular application.

Figures 6 and 7 present the comparison between the predicted values from the machine learning models and the actual measurements for both the training and testing datasets, respectively. This comparison aims to give us a deeper understanding of the predictive capabilities of these models and how well they align with the real-world findings. Continuing the trend from previous observations, the SGD, RF, and HGB models did not perform satisfactorily in the training dataset. This was highlighted by the mismatch in the statistical distributions of the predicted and measured values. Such a divergence suggests a poor fitting of these models to the training data, leading to unsatisfactory predictions. It could be due to the models not being able to capture the complexities of the data, or perhaps the models are not suitable for this particular task. Further investigation would be required to diagnose the exact issue. In the case of the testing dataset, the XGB model, which previously displayed strong performance, was noted to have two significant outliers that adversely affected its performance. An 'outlier' in this context means a data point that deviates significantly from the expected range or pattern. These outliers can be problematic as they can cause the model to produce predictions that are inconsistent with the overall pattern of the data. In other words, the presence of such outliers may distort the model's ability to accurately predict maximum moment capacity in RC structural circular columns. However, it's worth noting that the presence of outliers is not always an indication of poor model performance. Sometimes, outliers can arise due to exceptional cases in the data that are not representative of the typical scenario. Thus, it would be beneficial to investigate the source of these outliers in the XGB model's predictions and whether they are due to unusual data points or some characteristic of the model itself.



**Fig. 6.** Boxplots for the developed machine learning models in the training dataset.

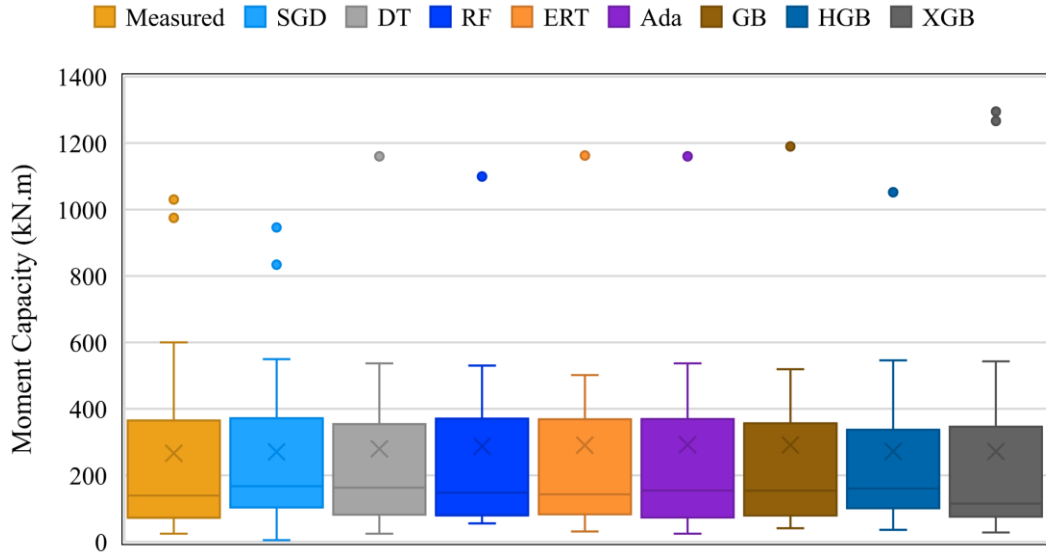


Fig. 7. Boxplots for the developed machine learning models in the testing dataset.

Figure 8 presents a quantitative evaluation of the machine learning models using three metrics for goodness-of-fit and error assessment. These metrics provide a more objective measure of model performance, as they capture the degree of agreement between the predicted and actual values, as well as the magnitude of errors made by the models. As per this evaluation, the XGB model demonstrated poor performance in the testing dataset due to the presence of two significant outliers that adversely affected its results. This supports the earlier observation of the model's vulnerability to extreme cases, which could potentially limit its effectiveness in making precise predictions in certain situations. Conversely, the SGD and RF models recorded the poorest overall performance, combining both training and testing datasets. This finding further substantiates the previous indications of these models' deficiencies in adequately capturing the intricacies of the data and the task at hand. In contrast, the Histogram-based HGB model was found to have the most robust performance. It demonstrated exceptional control over overfitting issues, as evidenced by the consistency in the range of residuals between training and testing datasets and the corresponding goodness-of-fit metrics. This illustrates the HGB model's potential as a suitable choice for predicting the maximum moment capacity of RC structural circular columns.

In general, the ACI 318 provides engineers with the theoretical frameworks and practical design procedures necessary for creating safe, durable, and efficient structures. One of the core components of ACI 318 is the approach it offers for estimating the maximum moment capacity of spiral RC columns. These types of columns, typically characterized by a continuous helical reinforcement wrapped around the longitudinal reinforcement, are widely recognized for their high load-bearing capacity and ductility. The ACI 318 approach to estimating the maximum moment capacity of these columns combines various mechanical properties of the constituent materials and geometrical characteristics of the column. The calculation procedure is initiated by determining the axial load capacity ( $P_n$ ) of the column. This is done by adding the contribution from the concrete ( $0.85 f_c A_g$ ) and the steel ( $f_y A_s$ ), where  $f_c$  is the specified compressive strength of the concrete,  $A_g$  is the gross area of the column,  $f_y$  is the yield strength of the steel, and  $A_s$  is the area of the steel.

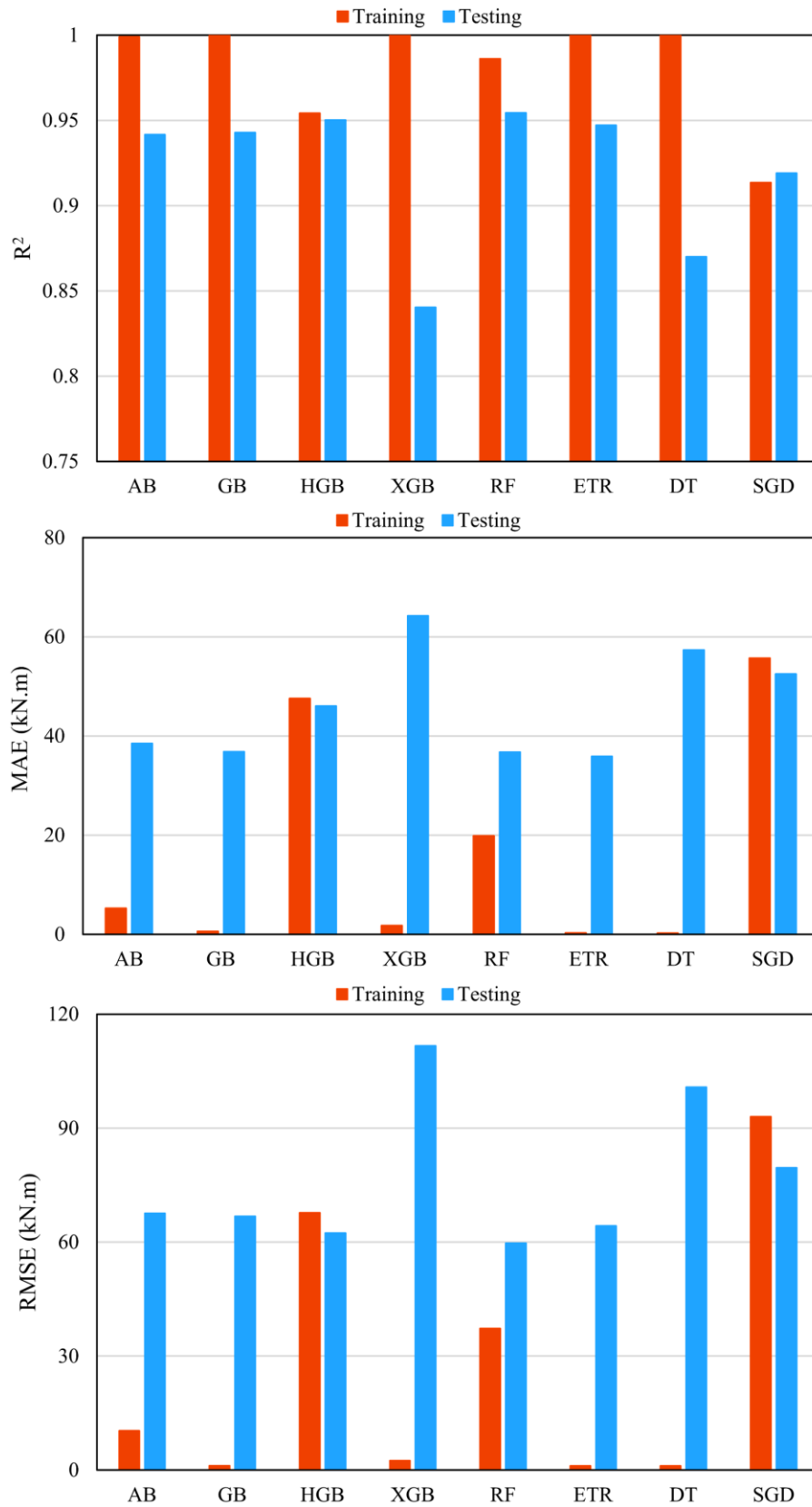


Fig. 8. Goodness of fit and errors of the developed machine learning models.

For spiral RC columns, ACI 318 specifies an enhancement factor due to the confinement provided by the spiral reinforcement. The concrete strength,  $f_c$ , is multiplied by this enhancement factor. This factor is higher for spirally reinforced columns, acknowledging the additional confinement and subsequent strength the spiral provides to the concrete core. Next, the nominal moment capacity ( $M_n$ ) is calculated using the equation  $M_n = P_n \cdot (h/2 - d')$ , where  $h$  is the overall dimension of the column, and  $d'$  is the distance from the extreme compression fiber to the centroid of the longitudinal reinforcement. This equation reflects the understanding that moment is the product of force and distance (lever arm). To ensure the structure's safety under service loads, ACI 318 requires that the design strength ( $\phi P_n$  for axial force and  $\phi M_n$  for moment) must be greater than the factored loads ( $P_u, M_u$ ), where  $\phi$  is the strength reduction factor reflecting the variability of material strengths, workmanship, and the level of strain in the failure mode. The values of  $\phi$  are specified in the code based on the type of failure mode, with lower values for brittle failure modes and higher for ductile modes. To consider the biaxial bending, ACI 318 introduces the interaction diagram for the column, which is a graphical representation of the axial load capacity and the moment capacity in two orthogonal directions. It is a plot of the axial force versus bending moment for various eccentricities. For a specific combination of axial force and bending moment, the design is considered safe if it falls within the boundary of the interaction diagram.

Figures 9 and 10 further delve into a performance comparison, where the machine learning models are benchmarked against the conventional ACI 318 approach. The ACI 318 is a widely used method for estimating the maximum moment capacity of spiral RC columns. Comparatively, the machine learning models, with the exception of SGD, achieved superior performance, indicating their potential for enhanced precision in structural capacity predictions. The SGD model, however, was the outlier with the highest residual value, reiterating its unsatisfactory performance in this context. The comparative superiority of machine learning models over the traditional ACI 318 equation demonstrates the significant potential of machine learning in enhancing the precision of maximum moment capacity prediction in spiral RC columns. However, model selection and calibration remain pivotal in realizing this potential, as the performance can significantly vary based on the characteristics of the model and the data.

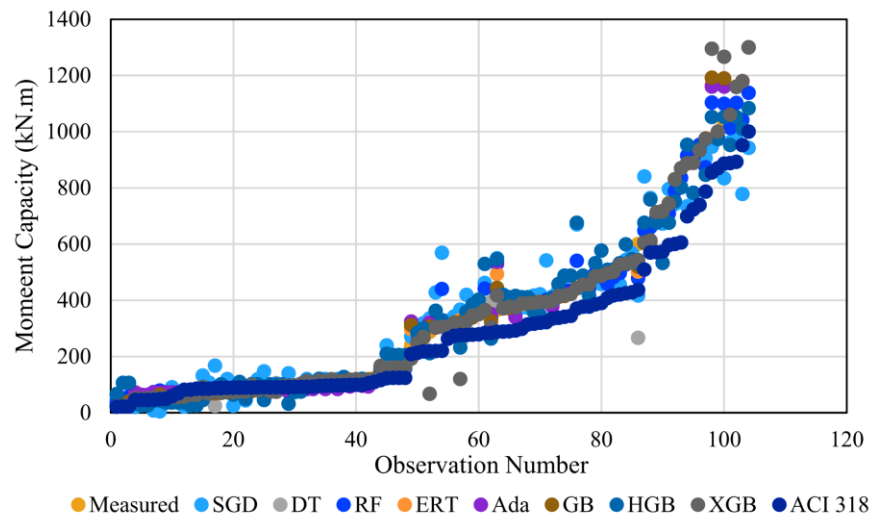
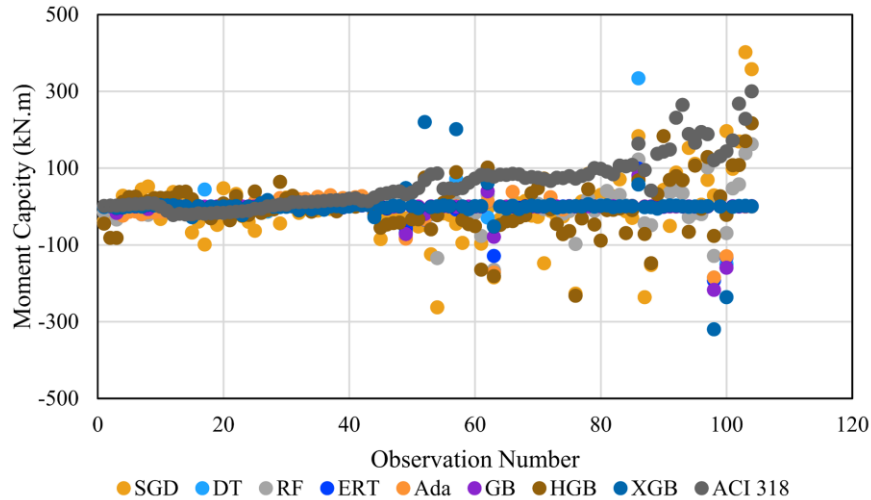


Fig. 9. Overall performance of the developed machine learning models against the ACI 318.





**Fig. 10.** Overall residuals of the developed machine learning models against the ACI 318.

Figure 11 gives an interesting perspective by comparing the statistical distribution of the results from the machine learning models and the ACI 318 approach. The statistical distribution is an essential tool for understanding data as it provides information on the variability, central tendency, and skewness in the data, which are important characteristics for interpreting and predicting outcomes. The American Concrete Institute's ACI 318 approach, a traditionally employed method for estimating the maximum moment capacity of RC columns, shows a significant underestimation of the moment capacity when compared to the measured values. Such a tendency towards underestimation can be problematic as it might lead to overly conservative designs and higher costs. Contrarily, the ERT and Ada machine learning models showed promising results, effectively replicating the overall distribution of the observed values. This is a desirable quality in predictive models, as it implies these models can better capture the variability in the data and, therefore, can provide more accurate and reliable predictions. These findings serve to underline the potential of machine learning models in enhancing the accuracy of structural capacity predictions. By mimicking the actual distribution of the outcomes more closely, these models can provide more reliable and cost-effective solutions. However, as always in machine learning, a single measure such as the statistical distribution should not be the only criterion for evaluating model performance. Other aspects such as the magnitude of residuals, handling of outliers, and computational efficiency should also be considered to ensure a comprehensive assessment.

In general, the success of employed algorithms such as XGB can be attributed to several factors. Firstly, XGB is known for its strong performance due to its ensemble approach, which combines multiple weak learners to create a powerful model. XGB leverages gradient boosting, which sequentially adds new models that focus on the errors of the previous models, leading to better overall predictions. Moreover, XGB handles complex datasets effectively, thanks to its ability to handle missing values, feature interactions, and nonlinear relationships. On the other hand, the weak results observed with SGD could be due to various reasons. SGD is a powerful optimization algorithm for large-scale datasets, but it may struggle with certain characteristics.

For instance, SGD's performance can suffer when dealing with sparse data or when the learning rate is not properly tuned. Additionally, SGD might require careful hyperparameter tuning and regularization to prevent overfitting. Overall, the success of XGB lies in its ensemble approach and ability to handle complex data, while the weaker results with SGD could be due to issues related to the dataset or parameter settings.

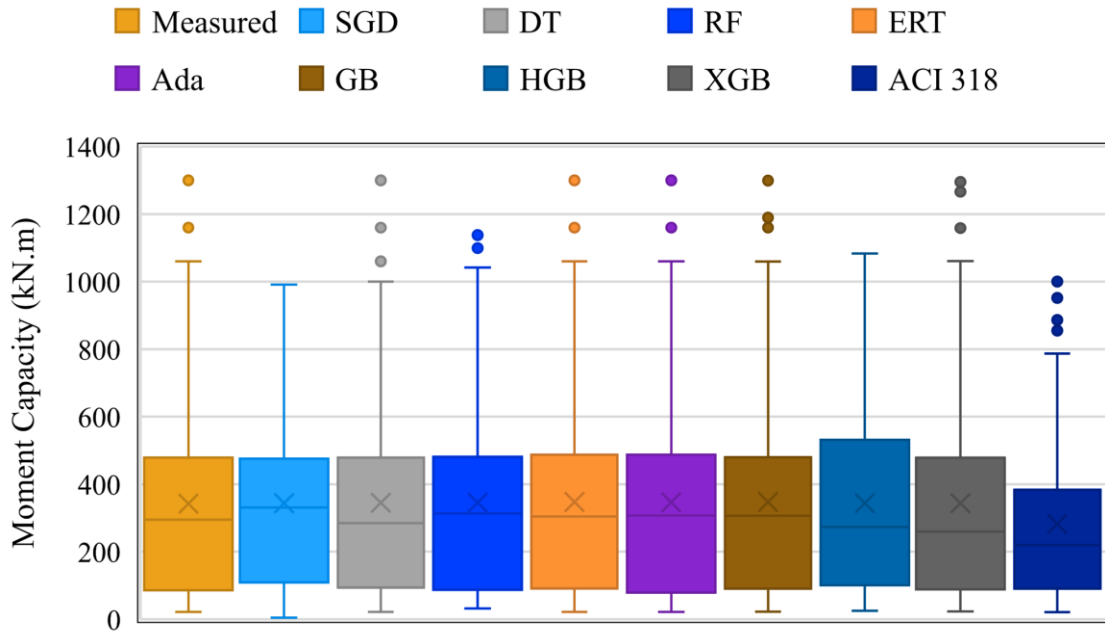
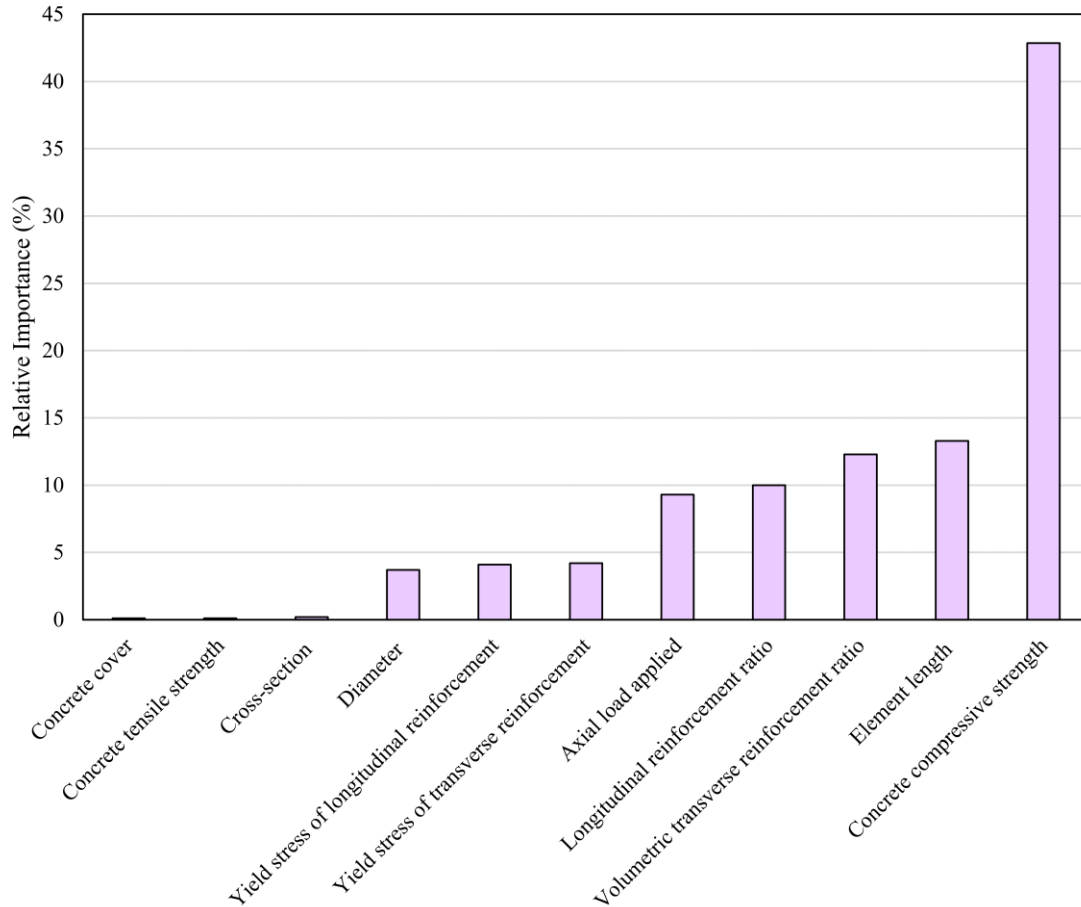


Fig. 11. Overall reliability of the developed machine learning models against the ACI 318.

Figure 12 delivers an essential perspective on the machine learning models used in this study by assessing the influence of input parameters on the outcomes of the models. This parameter evaluation, often referred to as feature importance, is vital as it provides insight into which factors significantly contribute to the predictive capabilities of the models. From the evaluation, it appears that concrete compressive strength is the most impactful parameter in the prediction decision across the machine learning models, indicating its central role in determining the maximum moment capacity of RC structural circular columns. This is followed by the element length, which is also deemed significantly important in driving the predictions. On the contrary, some parameters such as the concrete cover, concrete tensile strength, and the cross-section of the element exhibited lower relative importance in the prediction process. Although these parameters are less influential in isolation, it's important to note that they still contribute to the overall predictive power of the models. In machine learning, even seemingly less important features can add value by enhancing the models' ability to capture the complexity and diversity of the data. They may interact with other features in ways that improve the models' accuracy and ability to generalize to new data. It's also worth mentioning that the importance of a feature can depend on the specific model being used. Different models might assign different levels of importance to the same feature, depending on how they learn from the data and make predictions.



**Fig. 12.** Relative importance of each input variable on the machine learning model evaluated using the XGB method.

## 4. Conclusions

In summary, this study aimed to evaluate the effectiveness and potential of various machine learning models for predicting the maximum moment capacity of spirally RC columns. The performance of the developed machine learning models was compared to the conventional ACI 318 method for estimating the maximum moment capacity. Further, a parametric evaluation was carried out to assess the impact of different input variables on the machine learning models' outputs. Based on the analysis and observations, the following conclusions can be drawn:

1. The machine learning models demonstrated high accuracy in predicting the maximum moment capacity of spiral RC columns, as reflected in the  $R^2$  testing values. Particularly, the HGB, RF, and ETR models achieved  $R^2$  values of 0.95, 0.96, and 0.95, respectively, indicating their superior accuracy.
2. The HGB model was identified as the most robust, showing excellent control over overfitting, high goodness-of-fit, and lower RMSE (62.37) and MAE (46.05) on testing data compared to other models.
3. The performance of the machine learning models, in general, surpassed that of the conventional ACI 318 approach. The ACI 318 method tends to underestimate the

maximum moment capacity of spiral RC columns, which could potentially lead to overly conservative and costly designs.

4. From the feature importance analysis, it was determined that the concrete compressive strength and element length have the highest influence on the predictions of the machine learning models. This highlights the significance of these factors in the structural capacity of RC columns and validates the role of these parameters in the design process.

It's important to note that while these findings provide promising indications of the potential of machine learning in structural engineering, further research could be beneficial to enhance the predictive capabilities of these models and to investigate their performance in other scenarios or with different datasets. This could include exploring different machine learning techniques, optimizing model parameters, and considering additional input features that could improve the models' accuracy.

## **Funding**

This research received no external funding.

## **Conflicts of interest**

The authors declare no conflict of interest.

## **Authors contribution statement**

BA: Conceptualization; BA: Formal analysis; HA: Investigation; BA and HA: Methodology; BA: Writing – original draft; HA: Reviewing/editing –final draft.

## **References**

- [1] Tanaka H. Effect of lateral confining reinforcement on the ductile behaviour of reinforced concrete columns. 1990.
- [2] Coffman HL, Marsh ML, Brown CB. Seismic Durability of Retrofitted Reinforced-Concrete Columns. *J Struct Eng* 1993;119:1643–61. [https://doi.org/10.1061/\(asce\)0733-9445\(1993\)119:5\(1643\)](https://doi.org/10.1061/(asce)0733-9445(1993)119:5(1643)).
- [3] Desayi P, Sundara Raja Iyengar KT, Sanjeeva Reddy T. Equation for stress-strain curve of concrete confined in circular steel spiral. *Matériaux Constr* 1978;11:339–45. <https://doi.org/10.1007/BF02473875>.
- [4] Mander JB, Priestley MJN, Park R. Observed Stress-Strain Behavior of Confined Concrete. *J Struct Eng* 1988;114:1827–49. [https://doi.org/10.1061/\(ASCE\)0733-9445\(1988\)114:8\(1827\)](https://doi.org/10.1061/(ASCE)0733-9445(1988)114:8(1827)).
- [5] Li B, Park R, Tanaka H. Stress-Strain Behaviour of High Strength Concrete Confined by Ultra High and Normal Strength Transverse Reinforcements. *Aci Struct J* 2001;98:395–406.
- [6] Zhou J-K, Lin W-K, Guo S-X, Zeng J-J, Bai Y-L. Behavior of FRP-confined FRP spiral reinforced concrete square columns (FCFRCs) under axial compression. *J Build Eng* 2022;45:103452. <https://doi.org/10.1016/j.job.2021.103452>.

- [7] Ahmadi M, Naderpour H, Kheyroddin A. ANN Model for Predicting the Compressive Strength of Circular Steel-Confined Concrete. *Int J Civ Eng* 2017;15:213–21. <https://doi.org/10.1007/s40999-016-0096-0>.
- [8] Chang W, Zheng W. Estimation of compressive strength of stirrup-confined circular columns using artificial neural networks. *Struct Concr* 2019;20:1328–39. <https://doi.org/10.1002/suco.201800259>.
- [9] Koçer M, Öztürk M, Hakan Arslan M. Determination of moment, shear and ductility capacities of spiral columns using an artificial neural network. *J Build Eng* 2019;26:100878. <https://doi.org/10.1016/j.jobbe.2019.100878>.
- [10] Naderpour H, Parsa P, Mirrashid M. Innovative Approach for Moment Capacity Estimation of Spirally Reinforced Concrete Columns Using Swarm Intelligence–Based Algorithms and Neural Network. *Pract Period Struct Des Constr* 2021;26. [https://doi.org/10.1061/\(ASCE\)SC.1943-5576.0000612](https://doi.org/10.1061/(ASCE)SC.1943-5576.0000612).
- [11] Naderpour H, Mirrashid M, Parsa P. Failure mode prediction of reinforced concrete columns using machine learning methods. *Eng Struct* 2021;248. <https://doi.org/10.1016/j.engstruct.2021.113263>.
- [12] Renaud O, Victoria-Feser M-P. A robust coefficient of determination for regression. *J Stat Plan Inference* 2010;140:1852–62. <https://doi.org/10.1016/j.jspi.2010.01.008>.
- [13] Ramanauskas R, Kaklauskas G, Sokolov A. Estimating the primary crack spacing of reinforced concrete structures: Predictions by neural network versus the innovative strain compliance approach. *Mech Adv Mater Struct* 2020;29:1–17. <https://doi.org/10.1080/15376494.2020.1751352>.
- [14] Noori F, Varaee H. Nonlinear Seismic Response Approximation of Steel Moment Frames Using Artificial Neural Networks. *Jordan J Civ Eng* 2022;16:108–38.
- [15] Shishegaran A, Karami B, Danalou E, Varaee H, Rabczuk T. Computational predictions for predicting the performance of steel 1 panel shear wall under explosive loads. *Eng Comput* 2021;ahead-of-print. <https://doi.org/10.1108/EC-09-2020-0492>.
- [16] Shishegaran A, Varaee H, Rabczuk T, Shishegaran G. High correlated variables creator machine: Prediction of the compressive strength of concrete. *Comput Struct* 2021;247. <https://doi.org/10.1016/j.compstruc.2021.106479>.
- [17] Zhang G, Safaeian N, Rashnoozadeh H, S. Band S, Mosavi A. Reliability assessment of compressive and splitting tensile strength prediction of roller compacted concrete pavement: introducing MARS-GOA-MCS. *Int J Pavement Eng* 2021;1–18. <https://doi.org/10.1080/10298436.2021.1990920>.
- [18] Ashrafian A, Safaeian Hamzehkolaei N, Dwijendra NKA, Yazdani M. An Evolutionary Neuro-Fuzzy-Based Approach to Estimate the Compressive Strength of Eco-Friendly Concrete Containing Recycled Construction Wastes. *Buildings* 2022;12:1280. <https://doi.org/10.3390/buildings12081280>.
- [19] Safaeian Hamzehkolaei N, Alizamir M. Performance Evaluation of Machine Learning Algorithms for Seismic Retrofit Cost Estimation Using Structural Parameters. *J Soft Comput Civ Eng* 2021;5:32–57. <https://doi.org/10.22115/scce.2021.284630.1312>.
- [20] Varaee H, Safaeian Hamzehkolaei N, Safari M. A Hybrid Generalized Reduced Gradient-Based Particle Swarm Optimizer for Constrained Engineering Optimization Problems. *J Soft Comput Civ Eng* 2021;5:86–119. <https://doi.org/10.22115/scce.2021.282360.1304>.
- [21] Karimi Pour A, Shirkhani A, Safaeian N, Zhuge Y, Farsangi E. Performance evaluation of composite concrete-filled steel tube columns by steel fibers and different cross-section shapes: Experimental and numerical investigations. *J Constr Steel Res* 2023;200. <https://doi.org/10.1016/j.jcsr.2022.107656>.