

Article

Compressive Strength Estimation of Steel-Fiber-Reinforced Concrete and Raw Material Interactions Using Advanced Algorithms

Kaffayatullah Khan ^{1,*}, Waqas Ahmad ², Muhammad Nasir Amin ¹, Ayaz Ahmad ³, Sohaib Nazar ² and Anas Abdulalim Alabdullah ¹

- ¹ Department of Civil and Environmental Engineering, College of Engineering, King Faisal University, Al-Ahsa 31982, Saudi Arabia; mgadir@kfu.edu.sa (M.N.A.); 218038024@student.kfu.edu.sa (A.A.A.)
² Department of Civil Engineering, COMSATS University Islamabad, Abbottabad 22060, Pakistan; waqasahmad@cuiatd.edu.pk (W.A.); sohaibnazar@cuiatd.edu.pk (S.N.)
³ MaREI Centre, Ryan Institute and School of Engineering, College of Science and Engineering, National University of Ireland Galway, H91 HX31 Galway, Ireland; a.ahmad8@nuigalway.ie
* Correspondence: kkhan@kfu.edu.sa

Abstract: Steel-fiber-reinforced concrete (SFRC) has been introduced as an effective alternative to conventional concrete in the construction sector. The incorporation of steel fibers into concrete provides a bridging mechanism to arrest cracks, improve the post-cracking behavior of concrete, and transfer stresses in concrete. Artificial intelligence (AI) approaches are in use nowadays to predict concrete properties to conserve time and money in the construction industry. Accordingly, this study aims to apply advanced and sophisticated machine-learning (ML) algorithms to predict SFRC compressive strength. In the current work, the applied ML approaches were gradient boosting, random forest, and XGBoost. The considered input variables were cement, fine aggregates (sand), coarse aggregates, water, silica fume, super-plasticizer, fly ash, steel fiber, fiber diameter, and fiber length. Previous studies have not addressed the effects of raw materials on compressive strength in considerable detail, leaving a research gap. The integration of a SHAP analysis with ML algorithms was also performed in this paper, addressing a current research need. A SHAP analysis is intended to provide an in-depth understanding of the SFRC mix design in terms of its strength factors via complicated, nonlinear behavior and the description of input factor contributions by assigning a weighing factor to each input component. The performances of all the algorithms were evaluated by applying statistical checks such as the determination coefficient (R^2), the root mean square error (RMSE), and the mean absolute error (MAE). The random forest ML approach had a higher, i.e., 0.96, R^2 value with fewer errors, producing higher precision than other models with lesser R^2 values. The SFRC compressive strength could be anticipated by applying the random forest ML approach. Further, it was revealed from the SHapley Additive exPlanations (SHAP) analysis that cement content had the highest positive influence on the compressive strength of SFRC. In this way, the current study is beneficial for researchers to effectively and quickly evaluate SFRC compressive strength.

Keywords: SFRC; building material; compressive strength; steel fiber; concrete



Citation: Khan, K.; Ahmad, W.; Amin, M.N.; Ahmad, A.; Nazar, S.; Alabdullah, A.A. Compressive Strength Estimation of Steel-Fiber-Reinforced Concrete and Raw Material Interactions Using Advanced Algorithms. *Polymers* **2022**, *14*, 3065. <https://doi.org/10.3390/polym14153065>

Academic Editor: Junfeng Guan

Received: 31 May 2022

Accepted: 6 July 2022

Published: 29 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The mechanical properties, toughness, ductility, fatigue resistance, and crack-arresting of concrete can be improved by adding fibers into it [1–9]. Specifically, the addition of steel fibers to cementitious concrete enhances the post-cracking behavior and toughness [10–15]. The addition of adequate steel fiber content (i.e., 0–1.5%) to concrete can improve its properties [1]. Steel, artificial, and natural fibers are incorporated into concrete to enhance the mechanical properties and resistance against cracks of cementitious concrete composites [2–8]. Different studies have been conducted on models for regular concrete mechanical properties depending

on a wide database [16], although there are additional predicting parameters such as fiber type, aspect ratio, and volumetric content for SFRC compared to normal concrete. However, the development of appropriate predictive models is still new. Subsequently, the scuffling of conventional nonlinear and linear regression models is used to determine the compressive strength of SFRC. ML techniques may assist in resolving the issue of difficulty for the strength prediction of SFRC [9–16]. Although multiple experimental studies have been conducted for this purpose, as reported in the literature, the prediction of SFRC properties having different mix design components is still quite hard. Therefore, in the current study, an effort is made to predict SFRC compressive properties by employing ML approaches.

The employment of ML techniques may effectively resolve complex problems in various engineering fields. Based on an input database, ML techniques may estimate outcomes. To predict concrete properties, two ML approaches, a standalone procedure (a single-model-based decision tree) and ensemble techniques (i.e., random forest, gradient boosting, and XGBoost) are used. As reported in the literature, the performance of ensemble models has been better than individual ML models. Chaabene, et al. [17] evaluated in detail the employment of ML techniques for the prediction of concrete mechanical properties. Furthermore, several types of research have been conducted for the anticipation of mechanical properties of different concrete types, such as phase-change-material-integrated concrete [18], self-healing concrete [19], high-performance concrete (HPC) [20], recycled aggregate concrete (RAC) [21], etc. The employment of ML techniques was performed by Han, et al. [22] to predict HPC compressive strength. Fine aggregates, coarse aggregates, cement, water, GGBFS, age, and five additional variable combinations were all considered as the database for input factors. The developed model provided HPC compressive strength prediction with high precision.

SFRC mechanical properties have extensively been determined in various studies [23–25]. However, the procedures of specimen-casting in the laboratory, curing, and testing consume a lot of time, effort, cost, and labor. Therefore, the employment of ML modern techniques for the assessment of SFRC mechanical properties may resolve such issues and reduce expenses for experimentation. Moreover, the effects of raw ingredients on compressive strength have still not been addressed considerably in recent research, providing a research gap. Accordingly, the effects of raw materials, i.e., input factors, on the outcome factor, i.e., compressive strength, are also determined and explained by performing SHapley Additive exPlanations. In the current study, different ensemble ML techniques are applied to predict SFRC compressive strength. Random forest, gradient boosting, and XGBoost are employed as ensemble ML models. Furthermore, a statistical analysis is also performed to evaluate all the models, and different ML models are compared. Based on the performances, a superior model is proposed to predict SFRC compressive strength. In addition, a SHapley Additive exPlanations (SHAP) analysis, i.e., a post hoc model agnostic approach, is employed to gain insight into the ML models [26,27]. SHAP integration with the ML algorithms is also performed in this study, which is still a research gap. The purpose of a SHAP analysis is to give an in-depth understanding of the SFRC mix design in terms of its strength factors via complex, nonlinear behavior and the description of input factor contributions by allocating a weighing factor to every input factor. It assists in the development of sustainable and durable concrete mixes.

2. Methodology

2.1. Machine-Learning Techniques

This ensemble method for classification and regression was proposed by Friedman [28]. The gradient-boosting method is the same as other boosting techniques but is limited to regression only. In this technique, each training set iteration is selected randomly and is validated by the base model, as represented in Figure 1. The execution accuracy and speed of gradient boosting can be enhanced by randomly subsampling the training data, which ultimately helps to avoid overfitting. The smaller the training data fraction, the higher the regression speed to fit smaller model data at each iteration. A shrinkage rate and an n-tree tuning factor are needed in gradient-boosting regression, where n-tree denotes the number

of grown trees. Here, the n -tree value should not be too small, and the shrinkage factor, usually named as the learning rate, is applicable for each expansion tree.

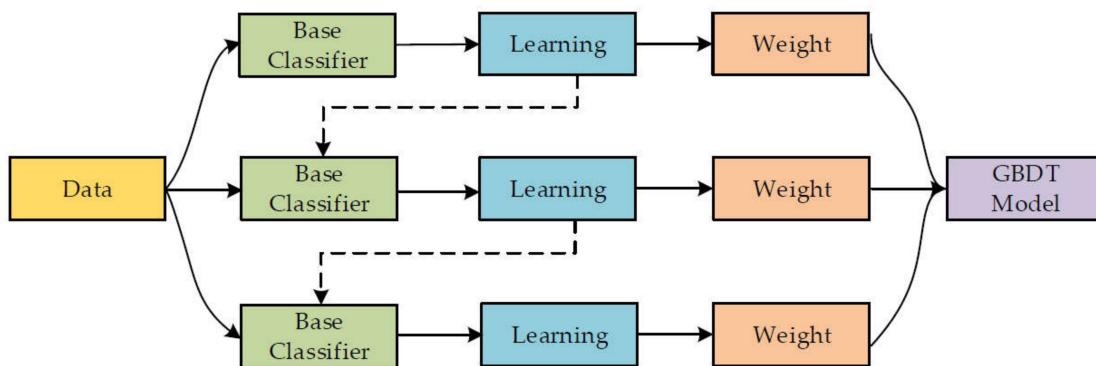


Figure 1. Gradient-boosting training process [29].

Chen and Guestrin [30] proposed an extreme gradient-boosting (extreme gradient boosting) algorithm, which is considered an authentic tool for researchers in the data science field due to the effective tree-based ensemble learning algorithm. Gradient-boosting architecture, i.e., applying different functions for result estimation using Equation (1), is the basis of extreme gradient boosting [28]:

$$\bar{y}_i = y_i^0 + \eta \sum_{k=1}^n f_k(U_i) \quad (1)$$

where the predicted output is shown by \bar{y}_i , using i th data with U_i as a parameter vector; n shows the estimator quantity in correspondence with independent tree structures against every f_k , where the range of k is from 1 to n ; and y_i^0 is the main hypothesis (mean of the original factors in the training dataset). η depicts the learning rate to enhance the model performance, along with the connection of additional trees to avoid overfitting. One major conflict in ML is developing a model with the least amount of overfitting. The training phase is complementarily evaluated in the extreme gradient-boosting model.

As per Equation (1), at the k^{th} level, the k^{th} estimator is in connection with the model, the forecasting of the k^{th} y_i^{-k} is determined through the predicted output $y_i^{-(k-1)}$ in a further step, and the respective developed f_k against the k^{th} complementary estimator is provided in Equation (2):

$$y_i^{-k} = y_i^{-(k-1)} + \eta f_k \quad (2)$$

where f_k depicts the weight of the leaves and is developed by minimizing the k^{th} tree objective function (Equation (3)):

$$f_{\text{obj}} = \gamma Z + \sum_{a=1}^Z \left[g_a \omega_a + \frac{1}{2} (h_a + \lambda) \omega_a^2 \right] \quad (3)$$

where the leaf node quantum is denoted by Z , the complexity factor by c , the constant coefficient by λ , and the weight (i.e., $1-Z$) of the leaf by ω_a^2 . λ and c are controlling factors applied to improve the model in terms of avoiding overfitting. h_a and g_a are the summed factors for the whole dataset linked with the previous and initial loss function gradient leaves, respectively. For building the k^{th} tree, a leaf is further bifurcated into multiple leaves. Gain parameters are used to apply such a system, as given in Equation (4):

$$G = \frac{1}{2} \left[\frac{O_L^2}{P_L + \lambda} + \frac{O_R^2}{P_R + \lambda} + \frac{(O_L + O_R)^2}{P_L + P_R + \lambda} \right] \quad (4)$$

where the gain parameters are denoted by G , and the right and left leaves are P_R and O_R , as well as P_L and O_L , respectively. The division criteria are generally assumed when

approximating the gain parameter at zero. λ and c are controlling factors that are dependent indirectly on the gain parameters. For instance, the gain parameter can considerably be decreased by a larger regularization parameter, ultimately preventing the leaf convolution process. However, the model performance for adopting training data is also be reduced by this. The basic, level-wise structure of the extreme gradient boosting tree model is shown in Figure 2.

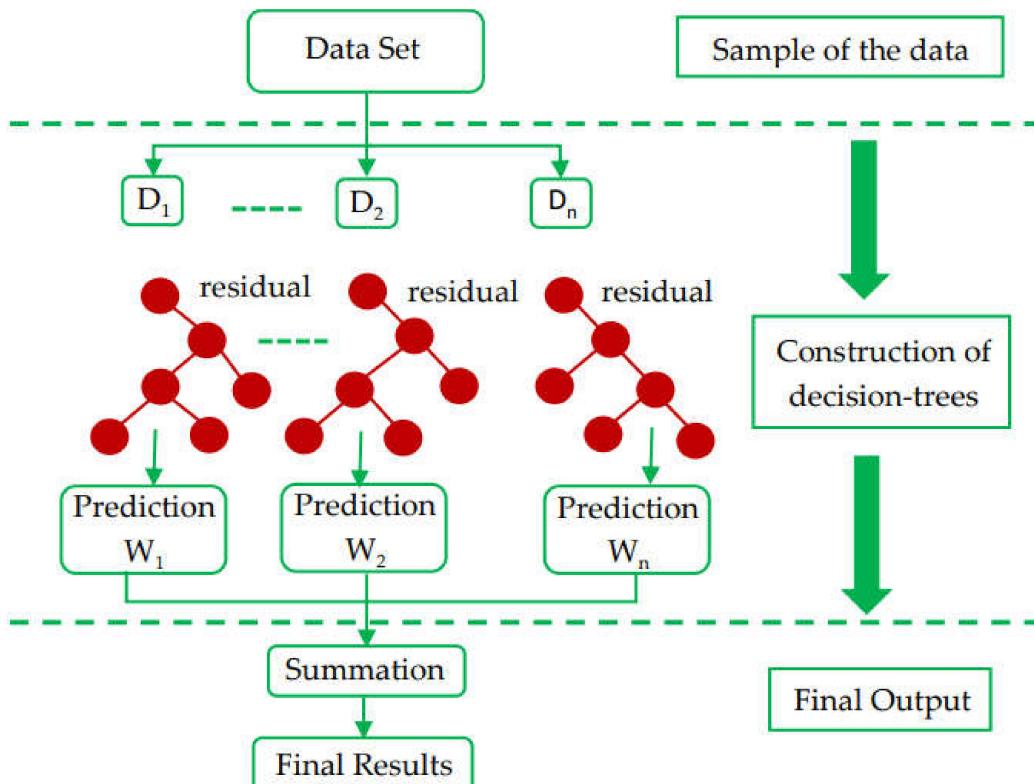


Figure 2. XGBoost (extreme gradient-boosting) algorithm structure [31].

The random forest model is a regression- and classification-based approach that has been studied by various researchers [22,32]. The compressive strength of concrete is predicted using a random forest model, as performed by Shaqadan [33]. The prime difference between random forest and DT is the number of trees. A single tree is developed in DT; however, in random forest, multiple trees are built, which is known as a forest. The dissimilar data are selected arbitrarily and are accordingly allocated to respective trees. Each tree has data in rows and columns, and different dimensions for the rows and columns are determined. The following steps are carried out for the growth of each tree: The data frame comprises 2/3 of the whole dataset that is randomly selected for each tree. This method is known as random forest. Random selection is made for the prediction variables, and the node splitting is achieved by finely splitting these variables. For all the trees, the remaining data are utilized to estimate the out-of-bag error. Accordingly, the final out-of-bag error rate is assessed by combining errors from each tree. Each tree provides regression, and among all the forest trees, the forest with the greatest amount of votes is selected for the model. The value of the votes can either be 1 or 0. The obtained proportion of 1 specifies the prediction probability. Among all the ensemble algorithms, random forest is the most sophisticated one. It includes desirable features for variable importance measures (VIMs) with robust overfitting resistance and fewer model parameters. DT is used as a base predictor for random forest. Acceptable results can be produced by random forest models with default parameter settings [34]. As allowed by random forest, combinations

of parameter settings and base predictors can be reduced to one. The basic, level-wise structure of the random forest model is presented in Figure 3.

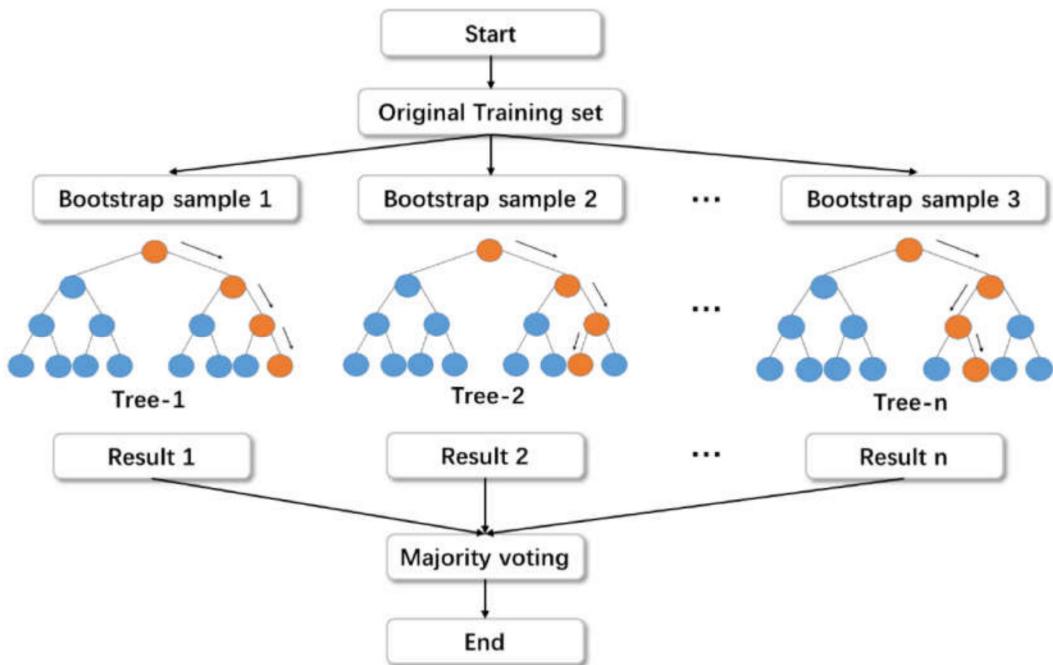


Figure 3. Random forest algorithm structure [35].

Furthermore, the current study identified influences for a global feature, as well as the interactions and dependencies of the considered feature on SFRC, depending on a game theory technique named SHapley Additive exPlanations (SHAP) [36] to enhance the explainability of the proposed model. In this technique, the prediction of every instance was explained by the computation of contributions for all the considered features for forecasting by applying Shapley values from the game theory coalition. The contribution for each individual feature value over all the possible combinations was marginally averaged to produce the Shapley value. The more influential features had higher absolute Shapley values. The Shapley values against each feature from the database were averaged to attain global feature influences. Afterwards, the sorting of these values in a decreasing manner in terms of importance was performed, followed by their plotting. A single point on the plot represented a Shapley value against individual instances and features. Feature importance and Shapley values determined the y and x axis positions, respectively. The higher influence of a feature on SFRC was depicted from its higher position on the y-axis, and its importance was depicted from a low-to-high color scale. The interactions of features and the corresponding impacts on SFRC were depicted from the SHAP feature dependence plots, in which interactions with other features were colored. This process provides better information than conventional plots of partial dependence [37]. In SHAP, more specifically, the feature importance (j) for the output of the model $f \phi^j(f)$, is allocated weightage for the summation of feature contributions towards the model outcome $f(x_i)$ to gain the overall possible feature combinations [38]. The $\phi^j(f)$ is expressed by Equation (5), as provided below:

$$\phi^j(f) = \sum_{S \subseteq \{x^1, \dots, x^p\} / \{x^j\}} \frac{|S|!(p - |S| - 1)!}{p!} \left(f(S \cup \{x^j\}) - f(S) \right) \quad (5)$$

where

S is the feature subset;

x_j is the feature j ;

p is the feature number in the model.

In the SHAP technique, the feature importance is determined by quantifying prediction errors while disturbing a specific feature value. The sensitivity of the prediction error is taken to allocate weightage to the significance of the feature while perturbing its value. The performance of the trained machine-learning model was also explained with the help of SHAP. SHAP employs an additional feature attribution technique, i.e., the addition of linear input factors, to demonstrate an interpretable model, which is taken by the model's outcome. For instance, a model with input parameters x_i (in which i ranges between 1 and k , and k depicts the input parameter number) and $h(x_s)$ depicts the explanation model with x_s as a simple input. However, Equation (6) is deployed to portray the original model $f(x)$:

$$f(x) = h(x_s) = \phi_0 + \sum_{i=1}^p \phi_i x_s^i \quad (6)$$

where

p is the input feature number;

ϕ_0 is the constant without any information (i.e., no input).

$x = m_x(x_s)$ indicates that the mapping function has a relationship with both the x and x_s input parameters. Lundberg and Lee [39] presented Equation (6) in which ($h()$), i.e., the prediction value, was enhanced by the ϕ_0 , ϕ_1 , and ϕ_3 terms, and a decrease of ϕ_4 in the $h()$ value was also observed (Figure 4). A single-value key to Equation (6) is the inclusion of three desired characteristics, such as consistency, missingness, and local accuracy. Consistency ensures no reduction in the attribution and is assigned to the respective feature as a change in a feature of more impact. Missingness ensures no value for importance is assigned to the missing features, i.e., $\phi_i = 0$ is employed by $x_s^i = 0$. Local accuracy ensured that the summation of feature attribution is taken as a function for the outcome, which includes a requirement of the model for matching the outcome f with x_s as a simplified input. $x = m_x x_s$ represents the attainment of local accuracy.

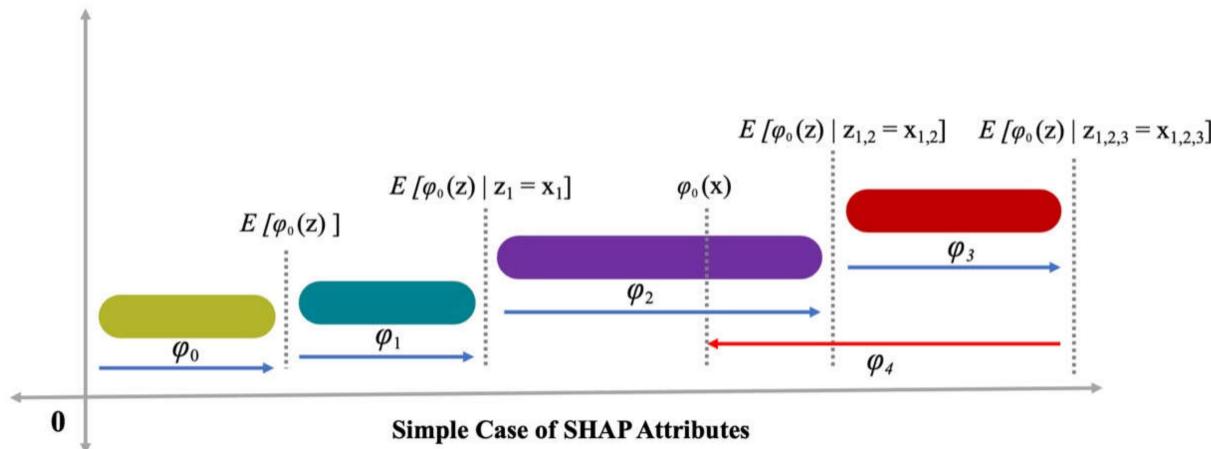


Figure 4. SHAP attributes [40].

2.2. Dataset Description

The database that was employed for the prediction of SFRC compressive strength is shown in Figure 5. Data regarding the compressive strength of SFRC were extracted from the literature [41–57]. These included cement, water, sand, coarse aggregates, superplasticizer, silica fume, fly ash, steel fiber, fiber length, and fiber diameter as inputs. These input factors were considered as compressive strength predictor variables. All these input and output parameters were collected within a compressive strength range of 20–100 MPa. These studies were selected because of the similarities between their input parameters. Figure 5 depicts the range for every variable and the minimum and maximum values. The compressive strength of SFRC was estimated using the Python and Spyder scripting of Anaconda software. The compressive strength histogram taken in the current study is presented in Figure 6.

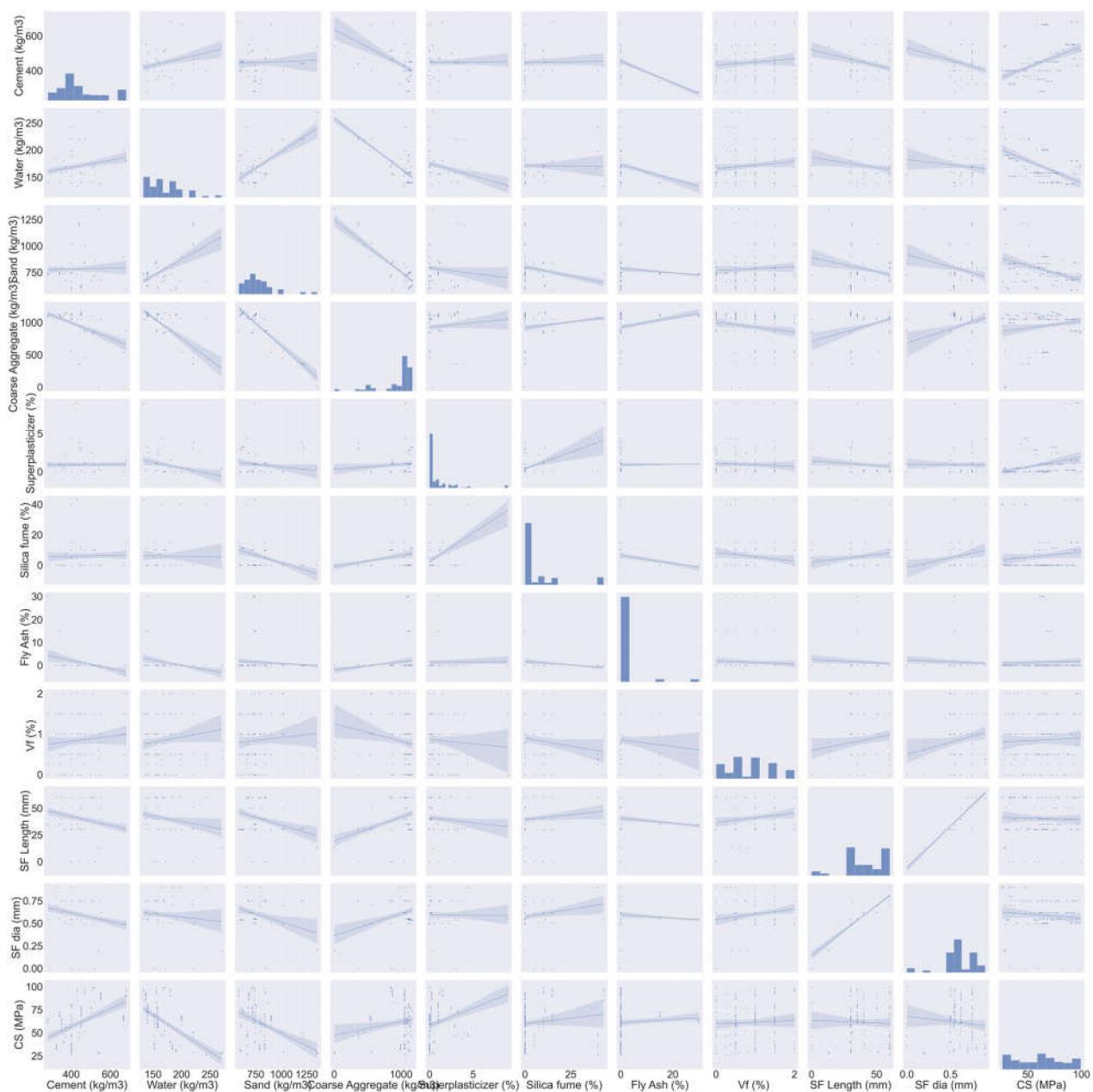


Figure 5. Description of parameter data.

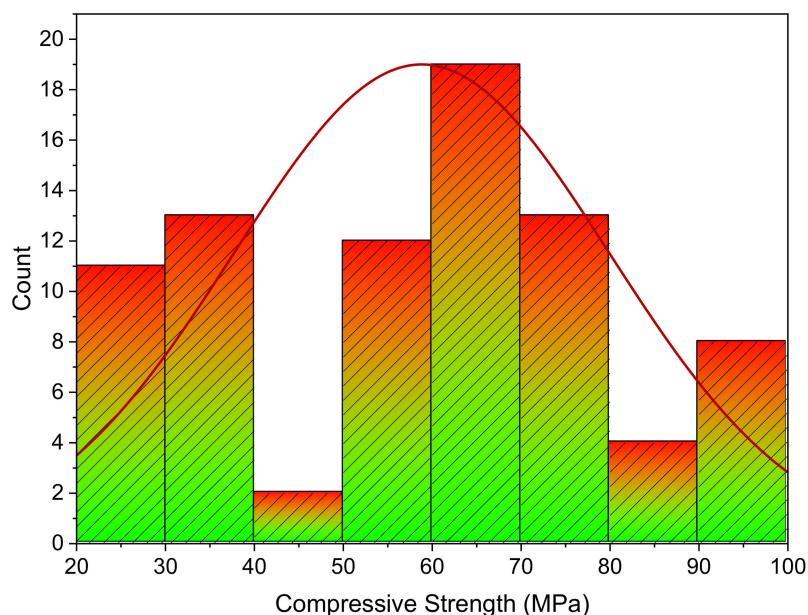


Figure 6. SFRC compressive strength distribution.

3. Results and Discussion

3.1. XGBoost

Figure 7 depicts the predicted and experimental value comparison for SFRC compressive strength using the XGBoost algorithm. A highly accurate outcome prediction for SFRC compressive strength was provided by the XGBoost algorithm. The adequacy of the XGBoost model was specified with the satisfactory R^2 value of 0.90. The error distribution of XGBoost was predicted, and experimental values for SFRC compressive strength are illustrated in Figure 8. The average error value for SFRC compressive strength was 4.63 MPa. A total of 70% of the total error values were below 5 MPa; 16% of these values ranged between 5 and 10 MPa, and 14% were above 10 MPa.

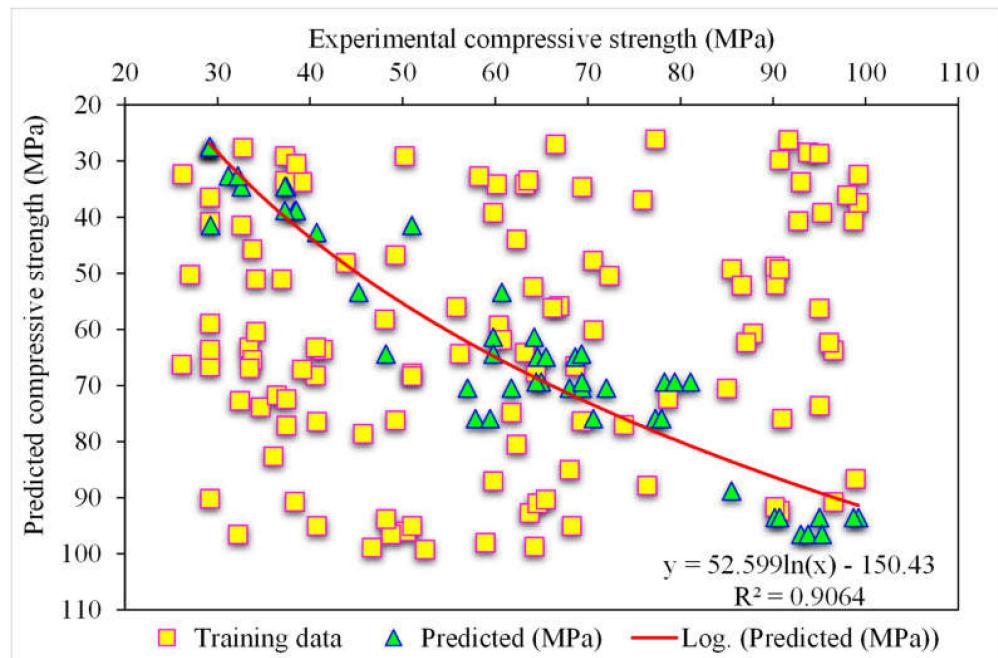


Figure 7. Experimental and XGBoost predicted results.

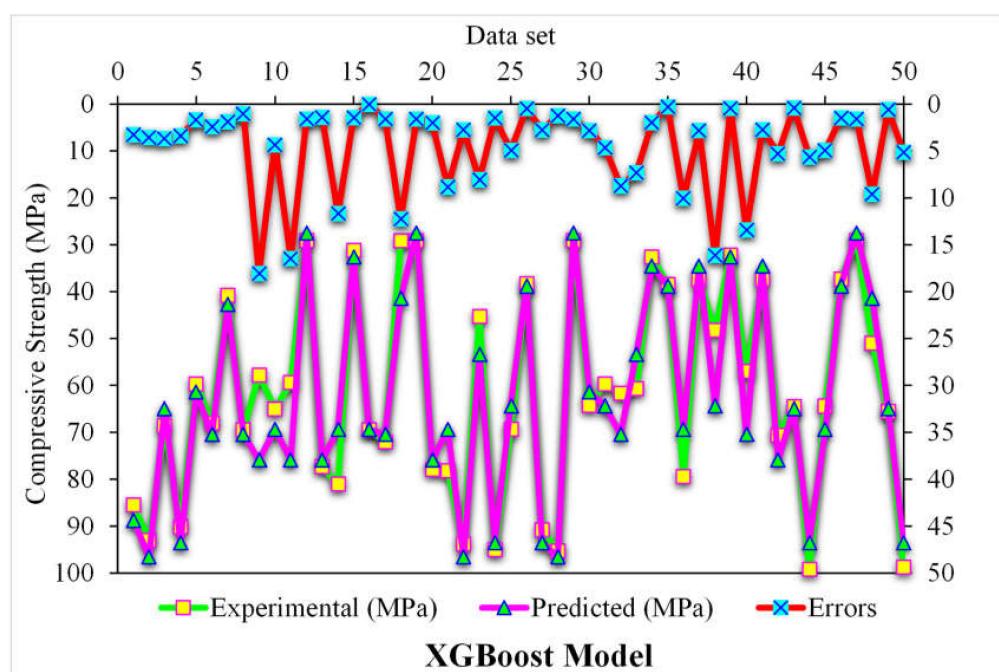


Figure 8. Experimental and XGBoost predicted values with errors.

3.2. Gradient Boosting

Figure 9 depicts the estimated gradient-boosting model and experimental outcome values for SFRC compressive strength. The $0.95 R^2$ value in the case of gradient boosting showed outcomes with higher accuracies than the XGBoost model. Figure 10 shows the error distribution for the gradient boosting estimated and experimental values in the case of SFRC compressive strength. It can be observed that 86% of the values were less than 5 MPa, 10% were between 5 and 10 MPa, and the remaining 4% of the values were above 10 MPa. The higher R^2 and lesser error values represented the higher precision of gradient boosting than the XGBoost model.

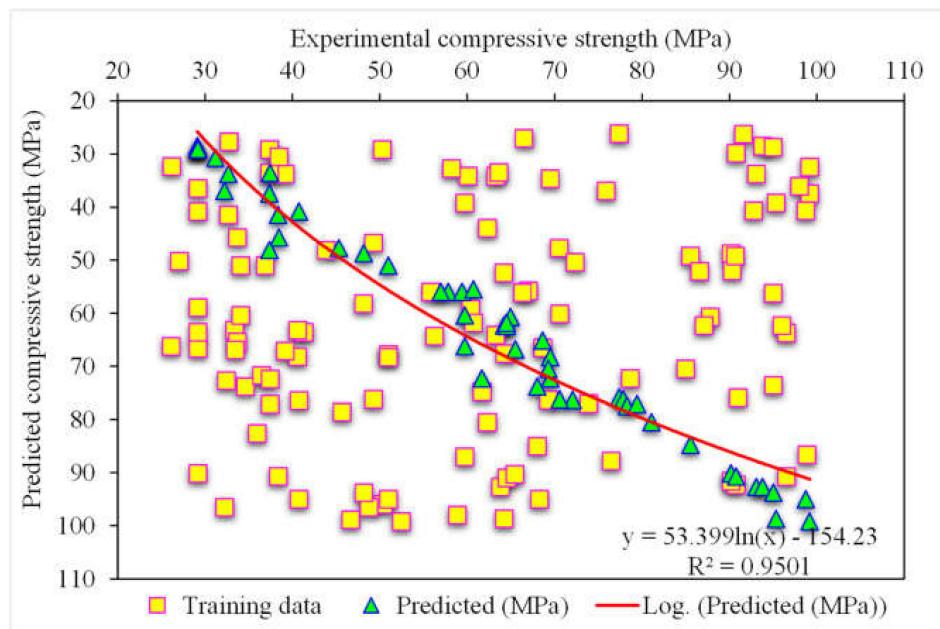


Figure 9. Experimental and gradient boosting predicted results.

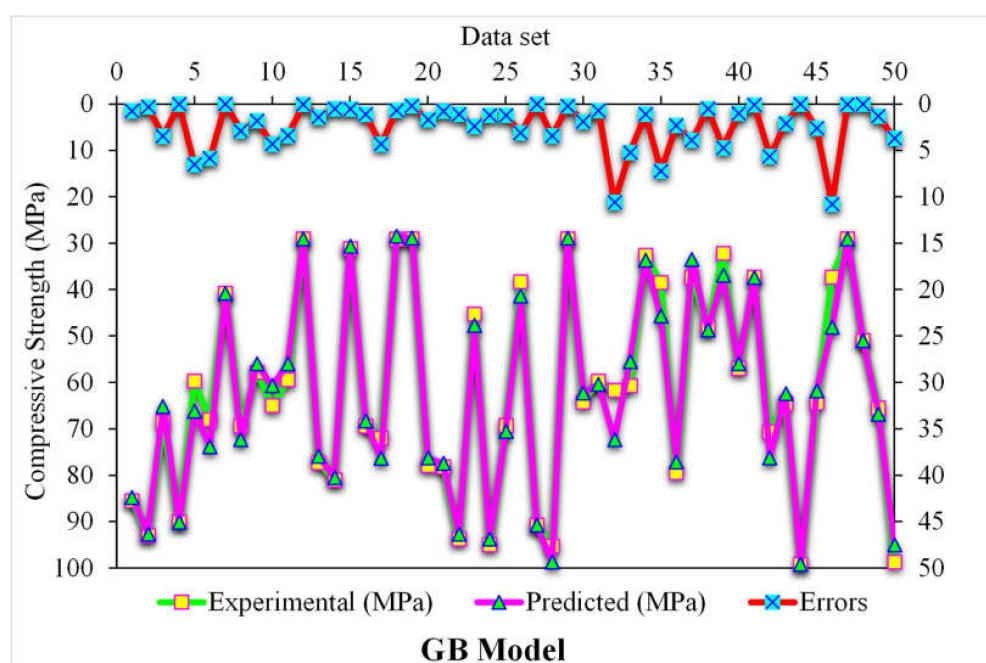


Figure 10. Experimental and gradient boosting predicted values with errors.

3.3. Random Forest

Figure 11 demonstrates the random forest estimated outcomes and experimental values for SFRC compressive strength. The 0.96 R^2 value for the random forest model represented more adequate results than the other two models. At the same time, the estimated SFRC compressive strength outcomes in the case of random forest were preferable to all the other considered ensemble models. Figure 12 reveals the distribution of random forest predicted and experimental values with errors for SFRC compressive strength. Here, 90% of the error values were less than 5 MPa, and the remaining 10% of the values were between 5 and 10 MPa. At the same time, not a single error value was more than 10 MPa. The R^2 and error values for SFRC compressive strength in the case of random forest were more precise and acceptable. Therefore, this outcome indicated that high-precision results could be predicted using random forest compared to other models.

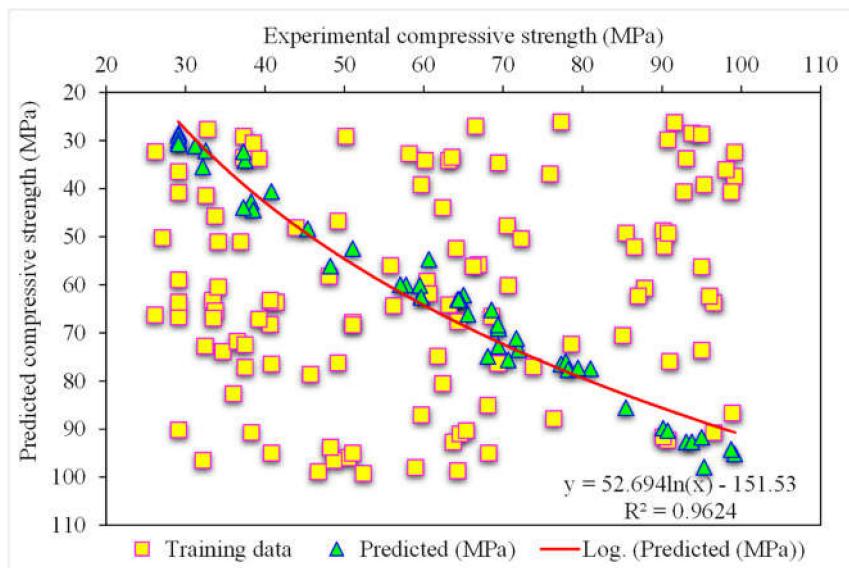


Figure 11. Experimental and random forest predicted results.

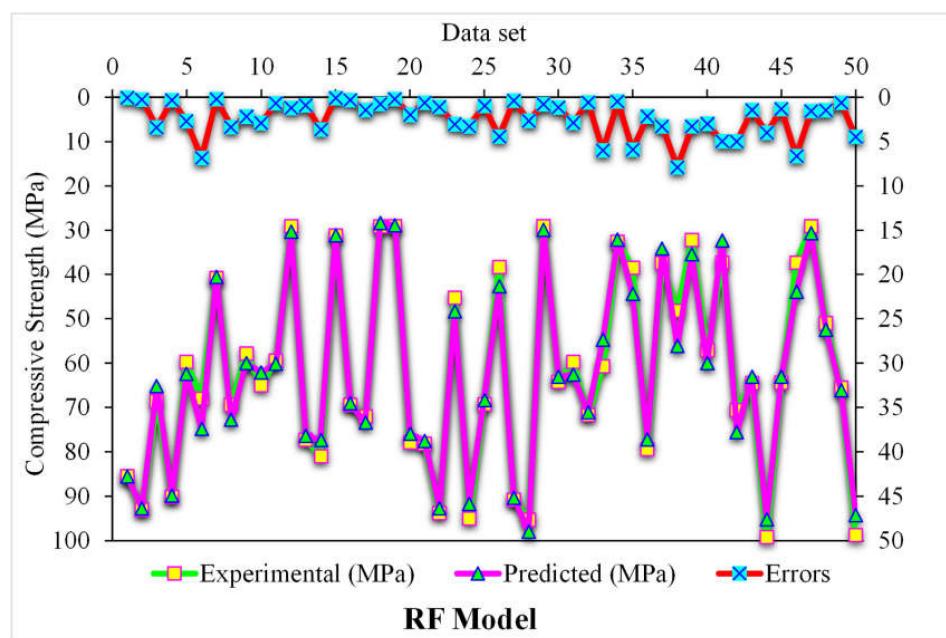


Figure 12. Distribution of experimental and BSA predicted values with errors.

3.4. Comparison of All Models

A k-fold cross-validation technique was adapted to validate the executing model. The model's performance was assessed by employing statistical checks [58–61]. Generally, the data were split into ten groups with random dispersion to perform k-fold cross-validation, and the repetition of this method was made ten times to attain results within an acceptable range, as presented in Figure 13. The statistical checks are listed in Table 1 for all the models. The R^2 values for the random forest, gradient-boosting, and XGBoost models were 0.96, 0.95, and 0.90, respectively, as illustrated in Figure 14a–c. The R^2 value for random forest was more than the other considered models, having less error values for SFRC compressive strength.

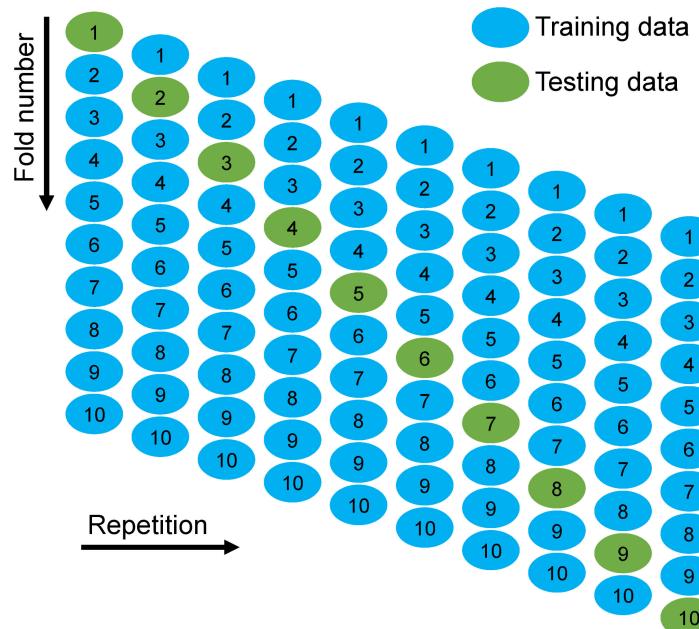
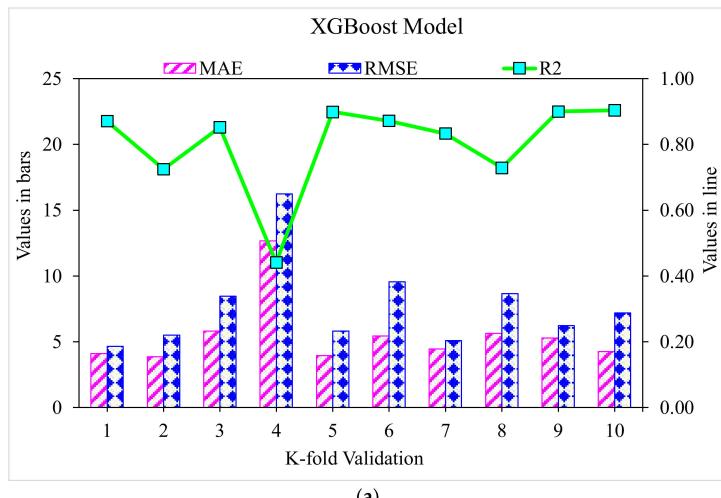


Figure 13. K-fold cross-validation procedure [62].

Table 1. Statistical checks of XGBoost, gradient-boosting, and random forest models.

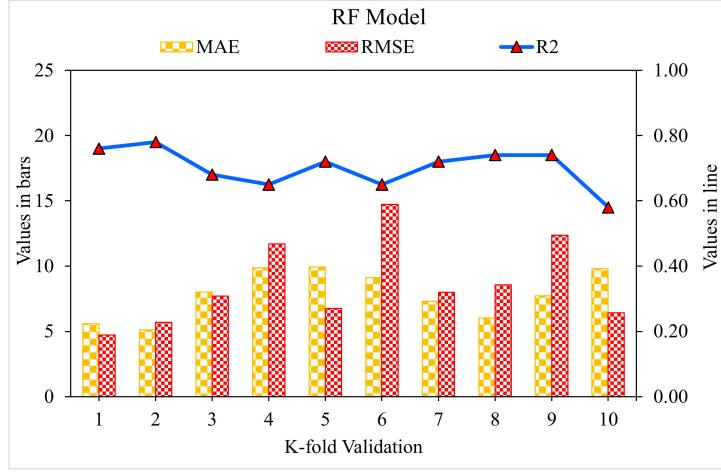
Techniques	MAE (MPa)	RMSE (MPa)	R ²
XGBoost	4.6	6.5	0.90
Gradient boosting	2.4	3.5	0.95
Random forest	2.4	3.1	0.96



(a)



(b)



(c)

Figure 14. Statistical representation: (a) XGBoost; (b) gradient boosting; and (c) random forest.

The SFRC compressive strength was estimated by employing the ensemble ML techniques in the current study to offer reliable and efficient results. The 0.96 R^2 value in the case of the random forest outcome showed a precise estimation of SFRC compressive strength. The superiority of the ensemble random forest ML algorithms for the compressive strength prediction of SFRC utilizing a single, optimized model out of twenty submodels is depicted in Figure 15a–c. Hence, it can be summarized that random forest showed higher precision and lower errors than the other considered models.

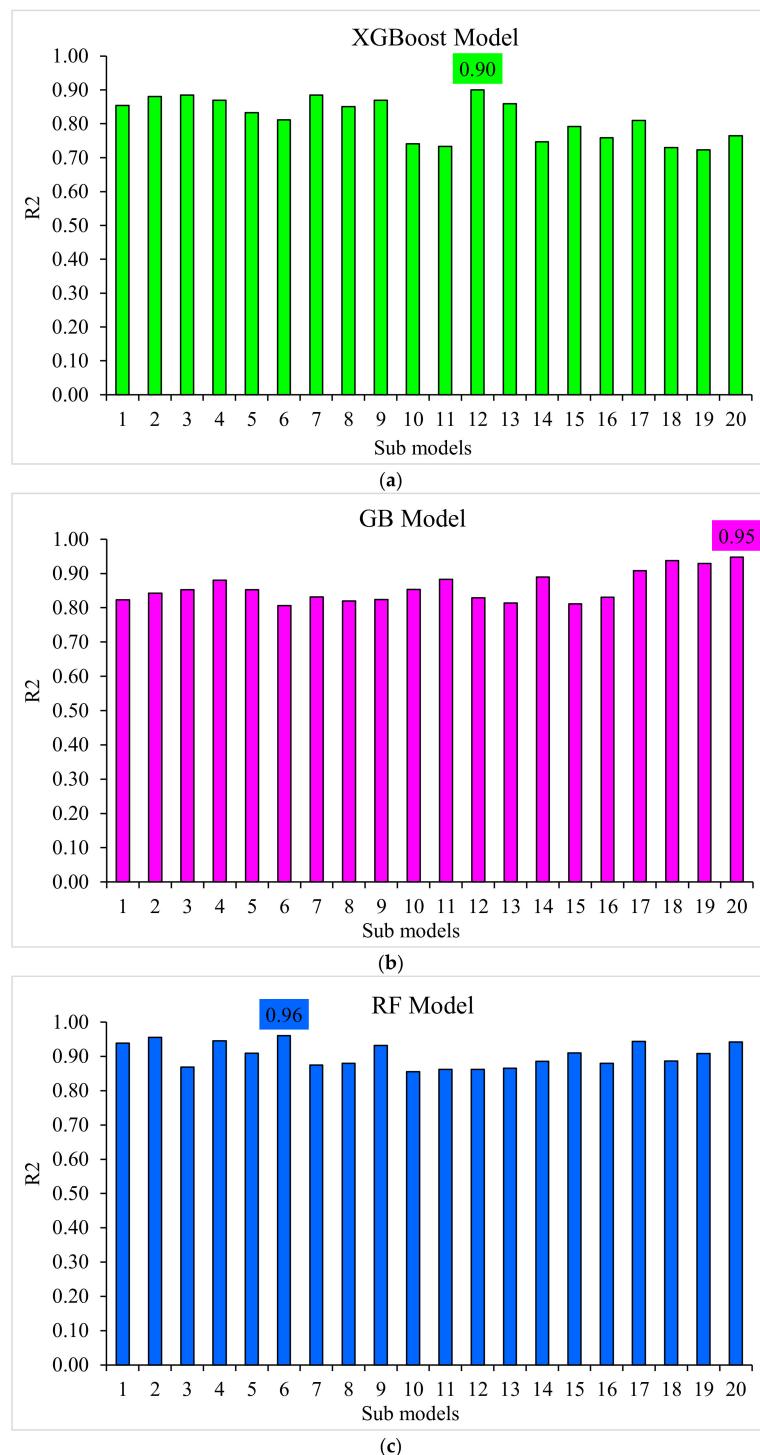


Figure 15. Submodel results: (a) XGBoost; (b) gradient boosting; and (c) random forest.

3.5. Enhanced Explainability of ML Models

This study also presents a detailed explanation of machine-learning models, as well as the interactions and dependencies of all the input features. By employing a SHAP tree explainer for the whole database, an enhanced feature that influences global representation is presented by merging local SHAP explanations. TreeExplainer, i.e., a tree-like SHAP approximation approach, was applied [63]. In this process, tree-based models' internal structures were assessed, which is the summation of the calculations linked with a tree model leaf node, leading to low-order complexity [63]. As the XGBoost model provided highly precise SFRC compressive strength prediction, this section interprets the model for SFRC compressive strength with the help of SHAP analysis. The different features were correlated with SHAP values for the SFRC compressive strength (as acquired from ensemble XGBoost modeling), as illustrated in Figure 16. It may be noted that the cement feature had the highest, i.e., approximately 20, SHAP value for SFRC compressive strength prediction. The cement feature positively influenced the SFRC compressive strength, which means that, by increasing the cement content, its strength was enhanced. The second-highest SHAP value was for water against SFRC compressive strength; however, it negatively influenced it. Enhancing the water content resulted in reduced SFRC compressive strength and vice versa. Afterwards, silica fume, the main factor for SFRC, had a SHAP value of approximately 5 (Figure 17). The silica fume content positively influenced the compressive strength of SFRC. Increasing the content of silica fume turned into increased SFRC compressive strength. Then, coarse aggregates were the next in terms of SHAP value. However, in this scenario, enhancement in the SFRC compressive strength resulted in the optimum content of coarse aggregates. After the optimized content, with any further addition of coarse aggregate content, the compressive strength of the SFRC decreased. In both ways, i.e., positive and negative, the influence of coarse aggregates on the compressive strength of SFRC was demonstrated by this behavior, whereas, in the case of fine aggregates, a negative influence on the compressive strength of SFRC was observed. SFRC particle packing density was difficult to attain in the case of an enhanced content of sand. In the same way, the SHAP value for fiber volume was next, followed by super-plasticizer, and silica fume, as well as steel fiber length and diameter. More or less the same SHAP values near zero were observed for all these features, depicting their lesser influence on SFRC compressive strength.

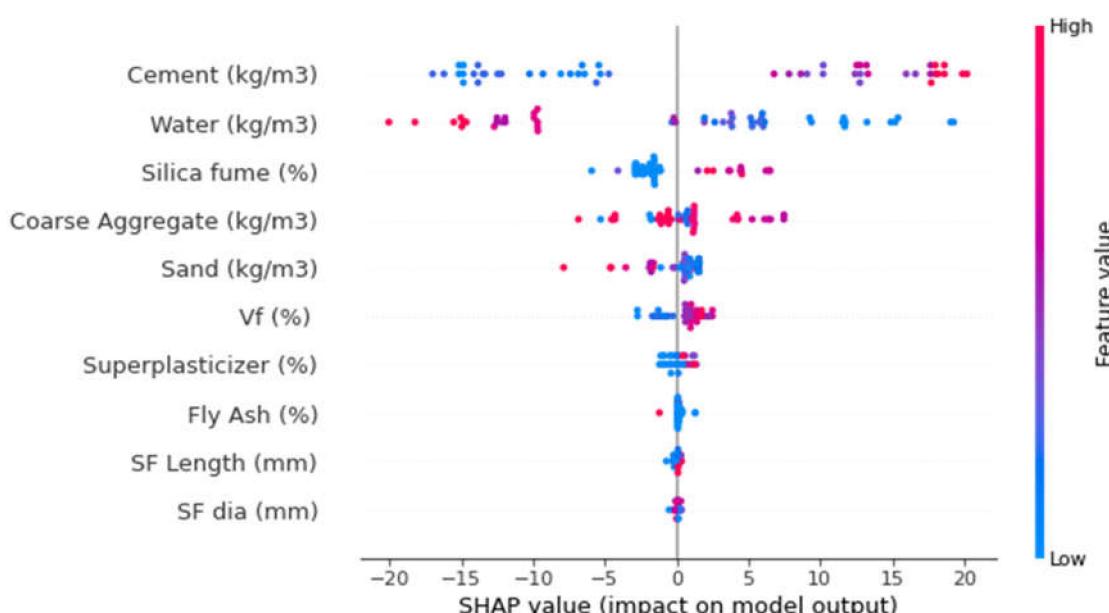


Figure 16. SHAP plot.

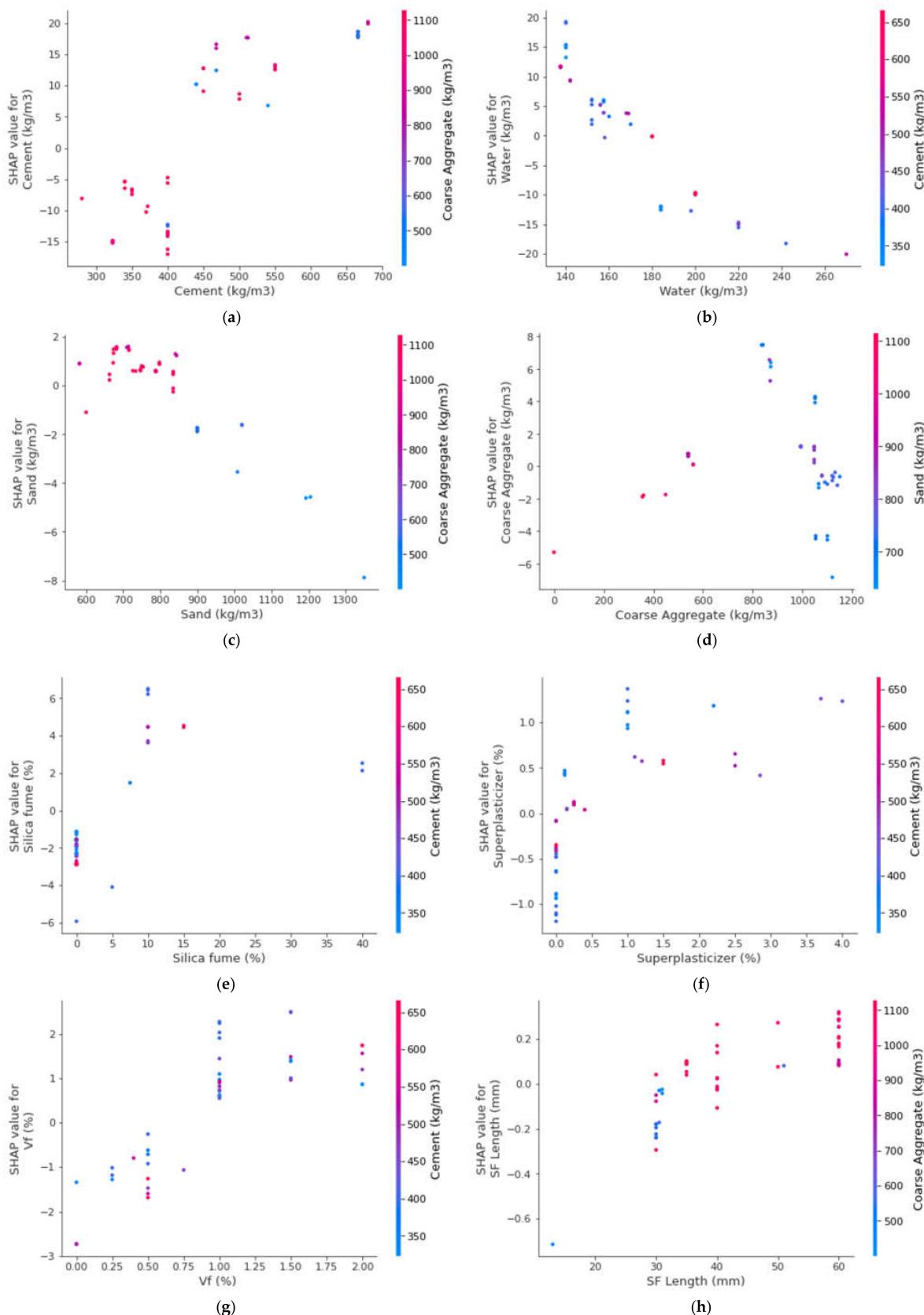


Figure 17. Interaction plots of various parameters: (a) cement; (b) water; (c) sand; (d) coarse aggregates; (e) silica fume; (f) super-plasticizer; (g) V_f ; and (h) fiber length.

The different features' interactions with SFRC compressive strength are illustrated in Figure 17. The interaction of the cement feature is presented in Figure 17a. It may be noted from the plot that cement influenced the SFRC compressive strength and was in a direct relationship with it. In Figure 17b, the negative influence of water content on SFRC compressive strength can be observed. The inverse relation of water content with SFRC compressive strength was observed. The interaction of the sand feature is presented in Figure 17c. Due to its effect on silica fume, the sand content depicted a negative influence and resulted in reduced SFRC compressive strength. Then, the coarse aggregates feature depicted both positive and negative interactions, depending upon the content (Figure 17d). The range up to the optimum content resulted in positive interactions, while interactions above that were negative. The silica fume interaction plot is shown in Figure 17e. It depicted an increasing trend of up to 20% content. However, it indicated negative influence at a considerably high content, i.e., 40%. Similarly, super-plasticizer showed a positive influence up to 2% content; however, further content negatively influenced the SFRC compressive strength. Likewise, the steel fiber volumetric content interacted positively and influenced SFRC compressive strength (Figure 17g). Figure 17h shows the interaction of steel fiber length with coarse aggregate content. It depicts a positive influence on SFRC compressive strength with enhancement in the steel fiber length.

4. Conclusions

The employment of machine-learning (ML) approaches to predict the mechanical properties of concrete in the construction industry is gaining attention nowadays. The main aim of the current study was to assess the precision of ML techniques to predict the compressive strength of steel-fiber-reinforced concrete (SFRC). The ten input parameters for prediction were cement, fine aggregates (sand), coarse aggregates, water, super-plasticizer, fly ash, silica fume, steel fiber length, and fiber diameter. The following outcomes were drawn from the conducted study:

- The 0.96 R^2 value in the case of the random forest model showed its accuracy in predicting SFRC compressive strength. In the case of ensemble gradient-boosting and XGBoost ML models having 0.95 and 0.90 R^2 values, respectively, the predicted SFRC compressive strength had less accuracy.
- The predicted SFRC compressive strength was optimized using twenty submodels with a range of 10 to 200 predictors. The ensemble random forest model produced a comparatively more precise prediction of SFRC compressive strength than all the other considered models.
- As revealed from the k-fold cross-validation outcomes, the gradient-boosting and random forest models had higher R^2 and lesser RMSE and MAE values for SFRC compressive strength than the other considered models, where the random forest model displayed the best accuracy for SFRC compressive strength prediction.
- Statistical checks such as RMSE and MAE were employed to evaluate the performances of the models. However, the higher determination coefficient and lower error value showed the superiority of the random forest model in the prediction of SFRC compressive strength.
- Among all the ML techniques, the random forest was the best approach to estimate SFRC compressive strength.
- The cement feature had the highest influence on the prediction of SFRC compressive strength, followed by water content, silica fume, coarse aggregates, sand, volumetric fiber content, and content of super-plasticizer, as revealed from SHAP analysis. However, the SFRC compressive strength was least influenced by the diameter of the steel fibers.
- SFRC compressive strength was positively influenced by cement content, as well as steel fiber volumetric content and length, as depicted from the feature interaction plots.

Indeed, a proper relational database and testing are important for engineering applications. This study was limited to the prediction of compressive strength with ten input parameters and did not include any other factors. However, a large database with an increased number

of experiments and more input parameters, such as specimen size, curing age, etc., must be developed in the future for the utilized models to provide more accurate results.

Author Contributions: K.K.: conceptualization, funding acquisition, resources, project administration, supervision, and writing—reviewing and editing. W.A.: conceptualization, data curation, software, methodology, investigation, validation, and writing—original draft. M.N.A.: methodology, investigation, and writing—reviewing and editing. A.A.: resources, visualization, and writing—reviewing and editing. S.N.: data curation, methodology, and writing—reviewing and editing. A.A.A.: visualization and writing—reviewing and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Project No. GRANT708).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this research are properly cited and reported in the main text.

Acknowledgments: The authors would like to acknowledge the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Project No. GRANT645).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cao, M.; Mao, Y.; Khan, M.; Si, W.; Shen, S. Different testing methods for assessing the synthetic fiber distribution in cement-based composites. *Constr. Build. Mater.* **2018**, *184*, 128–142. [[CrossRef](#)]
2. Khan, M.; Cao, M.; Hussain, A.; Chu, S. Effect of silica-fume content on performance of CaCO_3 whisker and basalt fiber at matrix interface in cement-based composites. *Constr. Build. Mater.* **2021**, *300*, 124046. [[CrossRef](#)]
3. Arshad, S.; Sharif, M.B.; Irfan-ul-Hassan, M.; Khan, M.; Zhang, J.-L. Efficiency of supplementary cementitious materials and natural fiber on mechanical performance of concrete. *Arab. J. Sci. Eng.* **2020**, *45*, 8577–8589. [[CrossRef](#)]
4. Xie, C.; Cao, M.; Guan, J.; Liu, Z.; Khan, M. Improvement of boundary effect model in multi-scale hybrid fibers reinforced cementitious composite and prediction of its structural failure behavior. *Compos. Part B Eng.* **2021**, *224*, 109219. [[CrossRef](#)]
5. Cao, M.; Khan, M. Effectiveness of multiscale hybrid fiber reinforced cementitious composites under single degree of freedom hydraulic shaking table. *Struct. Concr.* **2021**, *22*, 535–549. [[CrossRef](#)]
6. Khan, U.A.; Jahanzaib, H.M.; Khan, M.; Ali, M. Improving the tensile energy absorption of high strength natural fiber reinforced concrete with fly-ash for bridge girders. In *Key Engineering Materials*; Trans Tech Publications: Freienbach, Switzerland, 2018; pp. 335–342.
7. Khan, M.; Cao, M.; Ai, H.; Hussain, A. Basalt Fibers in Modified Whisker Reinforced Cementitious Composites. *Period. Polytech. Civ. Eng.* **2022**, *66*, 344–354. [[CrossRef](#)]
8. Zhang, N.; Yan, C.; Li, L.; Khan, M. Assessment of fiber factor for the fracture toughness of polyethylene fiber reinforced geopolymer. *Constr. Build. Mater.* **2022**, *319*, 126130. [[CrossRef](#)]
9. Khan, M.; Ali, M. Improvement in concrete behavior with fly ash, silica-fume and coconut fibres. *Constr. Build. Mater.* **2019**, *203*, 174–187. [[CrossRef](#)]
10. Khan, M.; Cao, M.; Chu, S.; Ali, M. Properties of hybrid steel-basalt fiber reinforced concrete exposed to different surrounding conditions. *Constr. Build. Mater.* **2022**, *322*, 126340. [[CrossRef](#)]
11. Li, L.; Khan, M.; Bai, C.; Shi, K. Uniaxial tensile behavior, flexural properties, empirical calculation and microstructure of multi-scale fiber reinforced cement-based material at elevated temperature. *Materials* **2021**, *14*, 1827. [[CrossRef](#)]
12. Khan, M.; Cao, M.; Xie, C.; Ali, M. Hybrid fiber concrete with different basalt fiber length and content. *Struct. Concr.* **2022**, *23*, 346–364. [[CrossRef](#)]
13. Khan, M.; Cao, M.; Xie, C.; Ali, M. Effectiveness of hybrid steel-basalt fiber reinforced concrete under compression. *Case Stud. Constr. Mater.* **2022**, *16*, e00941. [[CrossRef](#)]
14. Wang, J.; Dai, Q.; Si, R. Experimental and Numerical Investigation of Fracture Behaviors of Steel Fiber-Reinforced Rubber Self-Compacting Concrete. *J. Mater. Civ. Eng.* **2022**, *34*, 04021379. [[CrossRef](#)]
15. Su, P.; Dai, Q.; Li, M.; Ma, Y.; Wang, J. Investigation of the mechanical and shrinkage properties of plastic-rubber compound modified cement mortar with recycled tire steel fiber. *Constr. Build. Mater.* **2022**, *334*, 127391. [[CrossRef](#)]
16. Khan, M.; Lao, J.; Dai, J.-G. Comparative study of advanced computational techniques for estimating the compressive strength of UHPC. *J. Asian Concr. Fed.* **2022**, *8*, 51–68. [[CrossRef](#)]

17. Chaabene, W.B.; Flah, M.; Nehdi, M.L. Machine learning prediction of mechanical properties of concrete: Critical review. *Constr. Build. Mater.* **2020**, *260*, 119889. [\[CrossRef\]](#)
18. Marani, A.; Nehdi, M.L. Machine learning prediction of compressive strength for phase change materials integrated cementitious composites. *Constr. Build. Mater.* **2020**, *265*, 120286. [\[CrossRef\]](#)
19. Ramadan Suleiman, A.; Nehdi, M.L. Modeling self-healing of concrete using hybrid genetic algorithm–artificial neural network. *Materials* **2017**, *10*, 135. [\[CrossRef\]](#)
20. Castelli, M.; Vanneschi, L.; Silva, S. Prediction of high performance concrete strength using genetic programming with geometric semantic genetic operators. *Expert Syst. Appl.* **2013**, *40*, 6856–6862. [\[CrossRef\]](#)
21. Zhang, J.; Huang, Y.; Aslani, F.; Ma, G.; Nener, B. A hybrid intelligent system for designing optimal proportions of recycled aggregate concrete. *J. Clean. Prod.* **2020**, *273*, 122922. [\[CrossRef\]](#)
22. Han, Q.; Gui, C.; Xu, J.; Lacidogna, G. A generalized method to predict the compressive strength of high-performance concrete by improved random forest algorithm. *Constr. Build. Mater.* **2019**, *226*, 734–742. [\[CrossRef\]](#)
23. Ali, B.; Kurda, R.; Ahmed, H.; Alyousef, R. Effect of recycled tyre steel fiber on flexural toughness, residual strength, and chloride permeability of high-performance concrete (HPC). *J. Sustain. Cem.-Based Mater.* **2022**, *1*, 1–17. [\[CrossRef\]](#)
24. Ali, B.; Raza, S.S.; Hussain, I.; Iqbal, M. Influence of different fibers on mechanical and durability performance of concrete with silica fume. *Struct. Concr.* **2021**, *22*, 318–333. [\[CrossRef\]](#)
25. Ali, B. Development of environment-friendly and ductile recycled aggregate concrete through synergetic use of hybrid fibers. *Environ. Sci. Pollut. Res.* **2022**, *29*, 34452–34463. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Lauritsen, S.M.; Kristensen, M.; Olsen, M.V.; Larsen, M.S.; Lauritsen, K.M.; Jørgensen, M.J.; Lange, J.; Thiesson, B. Explainable artificial intelligence model to predict acute critical illness from electronic health records. *Nat. Commun.* **2020**, *11*, 3852. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Johnsen, P.V.; Riemer-Sørensen, S.; DeWan, A.T.; Cahill, M.E.; Langaas, M. A new method for exploring gene–gene and gene–environment interactions in GWAS with tree ensemble methods and SHAP values. *BMC Bioinform.* **2021**, *22*, 230. [\[CrossRef\]](#)
28. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [\[CrossRef\]](#)
29. Yao, M.; Zhu, Y.; Li, J.; Wei, H.; He, P. Research on predicting line loss rate in low voltage distribution network based on gradient boosting decision tree. *Energies* **2019**, *12*, 2522. [\[CrossRef\]](#)
30. Chen, T.; Guestrin, C. Xgboost: A scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
31. Amjad, M.; Ahmad, I.; Ahmad, M.; Wróblewski, P.; Kamiński, P.; Amjad, U. Prediction of Pile Bearing Capacity Using XGBoost Algorithm: Modeling and Performance Evaluation. *Appl. Sci.* **2022**, *12*, 2126. [\[CrossRef\]](#)
32. Zhang, J.; Ma, G.; Huang, Y.; Aslani, F.; Nener, B. Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression. *Constr. Build. Mater.* **2019**, *210*, 713–719. [\[CrossRef\]](#)
33. Shaqadan, A. Prediction of concrete mix strength using random forest model. *Int. J. Appl. Eng. Res.* **2016**, *11*, 11024–11029.
34. Xu, Y.; Ahmad, W.; Ahmad, A.; Ostrowski, K.A.; Dudek, M.; Aslam, F.; Joyklad, P. Computation of High-Performance Concrete Compressive Strength Using Standalone and Ensembled Machine Learning Techniques. *Materials* **2021**, *14*, 7034. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Guo, K.; Wan, X.; Liu, L.; Gao, Z.; Yang, M. Fault Diagnosis of Intelligent Production Line Based on Digital Twin and Improved Random Forest. *Appl. Sci.* **2021**, *11*, 7733. [\[CrossRef\]](#)
36. Lundberg, S. A game theoretic approach to explain the output of any machine learning model. *Github* **2021**.
37. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Wang, Q.; Hussain, A.; Farooqi, M.U.; Deifalla, A.F. Artificial intelligence-based estimation of ultra-high-strength concrete’s flexural property. *Case Stud. Constr. Mater.* **2022**, e01243. [\[CrossRef\]](#)
39. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 4–9 December 2017; Volume 30.
40. Shen, Z.; Deifalla, A.F.; Kamiński, P.; Dyczko, A. Compressive Strength Evaluation of Ultra-High-Strength Concrete by Machine Learning. *Materials* **2022**, *15*, 3523. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Soulioti, D.; Barkoula, N.; Paipetis, A.; Matikas, T. Effects of fibre geometry and volume fraction on the flexural behaviour of steel-fibre reinforced concrete. *Strain* **2011**, *47*, e535–e541. [\[CrossRef\]](#)
42. Yoo, D.-Y.; Yoon, Y.-S.; Banthia, N. Flexural response of steel-fiber-reinforced concrete beams: Effects of strength, fiber content, and strain-rate. *Cem. Concr. Compos.* **2015**, *64*, 84–92. [\[CrossRef\]](#)
43. Lee, J.-H.; Cho, B.; Choi, E. Flexural capacity of fiber reinforced concrete with a consideration of concrete strength and fiber content. *Constr. Build. Mater.* **2017**, *138*, 222–231. [\[CrossRef\]](#)
44. Köksal, F.; Altun, F.; Yiğit, İ.; Şahin, Y. Combined effect of silica fume and steel fiber on the mechanical properties of high strength concretes. *Constr. Build. Mater.* **2008**, *22*, 1874–1880. [\[CrossRef\]](#)
45. Yoon, E.-S.; Park, S.-B. An experimental study on the mechanical properties and long-term deformations of high-strength steel fiber reinforced concrete. *KSCE J. Civ. Environ. Eng. Res.* **2006**, *26*, 401–409.
46. Abbass, W.; Khan, M.I.; Mourad, S. Evaluation of mechanical properties of steel fiber reinforced concrete with different strengths of concrete. *Constr. Build. Mater.* **2018**, *168*, 556–569. [\[CrossRef\]](#)

47. Yoo, D.-Y.; Yoon, Y.-S.; Banthia, N. Predicting the post-cracking behavior of normal-and high-strength steel-fiber-reinforced concrete beams. *Constr. Build. Mater.* **2015**, *93*, 477–485. [[CrossRef](#)]
48. Lee, H.-H.; Lee, H.-J. Characteristic strength and deformation of SFRC considering steel fiber factor and volume fraction. *J. Korea Concr. Inst.* **2004**, *16*, 759–766. [[CrossRef](#)]
49. Oh, Y.-H. Evaluation of flexural strength for normal and high strength concrete with hooked steel fibers. *J. Korea Concr. Inst.* **2008**, *20*, 531–539.
50. Song, P.; Hwang, S. Mechanical properties of high-strength steel fiber-reinforced concrete. *Constr. Build. Mater.* **2004**, *18*, 669–673. [[CrossRef](#)]
51. Jang, S.-J.; Yun, H.-D. Combined effects of steel fiber and coarse aggregate size on the compressive and flexural toughness of high-strength concrete. *Compos. Struct.* **2018**, *185*, 203–211. [[CrossRef](#)]
52. Aldossari, K.; Elsaigh, W.; Shannag, M. Effect of steel fibers on flexural behavior of normal and high strength concrete. *Int. J. Civ. Environ. Eng.* **2014**, *8*, 22–26.
53. Dinh, N.H.; Park, S.-H.; Choi, K.-K. Effect of dispersed micro-fibers on tensile behavior of uncoated carbon textile-reinforced cementitious mortar after high-temperature exposure. *Cem. Concr. Compos.* **2021**, *118*, 103949. [[CrossRef](#)]
54. Thomas, J.; Ramaswamy, A. Mechanical properties of steel fiber-reinforced concrete. *J. Mater. Civ. Eng.* **2007**, *19*, 385–392. [[CrossRef](#)]
55. Sivakumar, A.; Santhanam, M. Mechanical properties of high strength concrete reinforced with metallic and non-metallic fibres. *Cem. Concr. Compos.* **2007**, *29*, 603–608. [[CrossRef](#)]
56. Afroughsabet, V.; Ozbakkaloglu, T. Mechanical and durability properties of high-strength concrete containing steel and polypropylene fibers. *Constr. Build. Mater.* **2015**, *94*, 73–82. [[CrossRef](#)]
57. Atiç, C.D.; Karahan, O. Properties of steel fiber reinforced fly ash concrete. *Constr. Build. Mater.* **2009**, *23*, 392–399. [[CrossRef](#)]
58. Ahmad, W.; Ahmad, A.; Ostrowski, K.A.; Aslam, F.; Joyklad, P.; Zajdel, P. Application of advanced machine learning approaches to predict the compressive strength of concrete containing supplementary cementitious materials. *Materials* **2021**, *14*, 5762. [[CrossRef](#)]
59. Ahmad, A.; Ahmad, W.; Aslam, F.; Joyklad, P. Compressive strength prediction of fly ash-based geopolymmer concrete via advanced machine learning techniques. *Case Stud. Constr. Mater.* **2021**, *16*, e00840. [[CrossRef](#)]
60. Yuan, X.; Tian, Y.; Ahmad, W.; Ahmad, A.; Usanova, K.I.; Mohamed, A.M.; Khallaf, R. Machine Learning Prediction Models to Evaluate the Strength of Recycled Aggregate Concrete. *Materials* **2022**, *15*, 2823. [[CrossRef](#)]
61. Shang, M.; Li, H.; Ahmad, A.; Ahmad, W.; Ostrowski, K.A.; Aslam, F.; Joyklad, P.; Majka, T.M. Predicting the Mechanical Properties of RCA-Based Concrete Using Supervised Machine Learning Algorithms. *Materials* **2022**, *15*, 647. [[CrossRef](#)]
62. Wang, Q.; Ahmad, W.; Ahmad, A.; Aslam, F.; Mohamed, A.; Vatin, N.I. Application of Soft Computing Techniques to Predict the Strength of Geopolymer Composites. *Polymers* **2022**, *14*, 1074. [[CrossRef](#)]
63. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. Explainable AI for trees: From local explanations to global understanding. *arXiv* **2019**, arXiv:1905.04610. [[CrossRef](#)]