

Article

Prediction of Ecofriendly Concrete Compressive Strength Using Gradient Boosting Regression Tree Combined with GridSearchCV Hyperparameter-Optimization Techniques

Zaineb M. Alhakeem¹, Yasir Mohammed Jebur² , Sadiq N. Henedy³, Hamza Imran⁴ , Luís F. A. Bernardo^{5,*} 
and Hussein M. Hussein⁶ 

¹ Computer Engineering Department, Iraq University College, Basrah 61004, Iraq

² Building and Construction Techniques Engineering Department, Al-Mustaqbal University College, Hillah 51001, Iraq

³ Department of Civil Engineering, Mazaya University College, Nasiriya City 64001, Iraq

⁴ Department of Environmental Science, College of Energy and Environmental Science, Alkarkh University of Science, Baghdad 10081, Iraq

⁵ Centre of Materials and Building Technologies (C-MADE), Department of Civil Engineering and Architecture, University of Beira Interior, 6201-001 Covilhã, Portugal

⁶ Medical Physics Department, Hilla University College, Babylon 51002, Iraq

* Correspondence: lfb@ubi.pt



Citation: Alhakeem, Z.M.; Jebur, Y.M.; Henedy, S.N.; Imran, H.; Bernardo, L.F.A.; Hussein, H.M. Prediction of Ecofriendly Concrete Compressive Strength Using Gradient Boosting Regression Tree Combined with GridSearchCV Hyperparameter-Optimization Techniques. *Materials* **2022**, *15*, 7432. <https://doi.org/10.3390/ma15217432>

Academic Editor: Miguel Ángel Sanjuán

Received: 4 October 2022

Accepted: 19 October 2022

Published: 23 October 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/>).

Abstract: A crucial factor in the efficient design of concrete sustainable buildings is the compressive strength (Cs) of eco-friendly concrete. In this work, a hybrid model of Gradient Boosting Regression Tree (GBRT) with grid search cross-validation (GridSearchCV) optimization technique was used to predict the compressive strength, which allowed us to increase the precision of the prediction models. In addition, to build the proposed models, 164 experiments on eco-friendly concrete compressive strength were gathered for previous researches. The dataset included the water/binder ratio (W/B), curing time (age), the recycled aggregate percentage from the total aggregate in the mixture (RA%), ground granulated blast-furnace slag (GGBFS) material percentage from the total binder used in the mixture (GGBFS%), and superplasticizer (kg). The root mean square error (RMSE) and coefficient of determination (R^2) between the observed and forecast strengths were used to evaluate the accuracy of the predictive models. The obtained results indicated that—when compared to the default GBRT model—the GridSearchCV approach can capture more hyperparameters for the GBRT prediction model. Furthermore, the robustness and generalization of the GSC-GBRT model produced notable results, with RMSE and R^2 values (for the testing phase) of 2.3214 and 0.9612, respectively. The outcomes proved that the suggested GSC-GBRT model is advantageous. Additionally, the significance and contribution of the input factors that affect the compressive strength were explained using the Shapley additive explanation (SHAP) approach.

Keywords: gradient boosting regression tree; machine learning; SHAP; eco-friendly concrete; compressive strength; prediction

1. Introduction

In recent years, the design and construction of sustainable buildings has become a major goal. One of the most important factors influencing this criterion is the use of Portland cement substitutes. For instance, cement production consumes vast energy and significantly contributes to greenhouse gas emissions and environmental degradation. Therefore, a reduction in cement incorporated in concrete mixes is necessary to protect the environment. In addition, to avoid the negative impacts of burying hazardous industrial wastes, using recycled concrete aggregate (RCA) and ground granulated blast-furnace slag (GGBFS) as waste materials in concrete mixes would also allow us to reduce greenhouse gas emissions, since less cement and natural aggregate (NA) will be produced [1].

Some researchers investigated the behavior of RCA in terms of durability and mechanical performance. They found that RCA has lower properties than natural concrete aggregate (NCA) [2–7].

Hence, concrete scientists must increase the qualities of RCA in order to improve the products produced with it due to the numerous environmental advantages of using RCA in the construction sector. Additionally, improving RCA makes it possible to utilize it to prepare concrete for a structure incorporating load-bearing structural members. However, old connected mortar, crushing techniques, and a poor interfacial transition zone (ITZ) are the critical causes for RCA's inferior quality. Previous studies have shown that whereas RCA has five phases—NA, old ITZ, old cement paste, new ITZ, and new cement paste—NCA has three phases, which are NA, ITZ, and cement mixture. As a result, the rheological, mechanical, and durability characteristics of RCA are gradually decreased because of the higher significant porosity and water absorption capacity caused by higher number of phases. Various approaches have been reported in the literature to reduce the lower properties of RCA and make it similar to NCA. Incorporating supplemental cement-based materials such as ground granulated blast-furnace slag (GGBFS) was one of the solutions to improve the RCA properties and eliminate its lower properties [8].

Slag cement, commonly known as GGBFS, is a by-product of the iron production sector. A hydraulic cementitious substance with pozzolanic properties is a slag cement. Any reactive aluminosilicate substance is referred to as pozzolan. Pozzolans interact with the cement hydration process by-products to generate concrete. Pozzolans do not often provide enough strength when combined with water alone; thus, cement must be added to the mixture. Because of this, it is widely believed in the concrete industry that cement, at the very least in a reasonable proportion, is necessary to provide the necessary strength. When combined with water, slag cement, which has cementitious qualities similar to cement, will hydrate and provide strength.

Studies on the impact of GGBFS on several RCA properties have been performed. When comparing with NCA, Çakır [9] reported that concrete with 30% GGBFS and 50% RCA shows good mechanical properties. Similarly, according to Majhi and Nayak [10], concrete containing 40% GGBFS and 50% RCA has good mechanical and physical properties. It was also found that this concrete is as durable as NCA. Rashad [11] looked into how GGBFS content affects the compressive strength (Cs) and found that a mix with 50% GGBFS can have similar characteristics as concrete mix with ordinary Portland cement (OPC) content. In their study on the effects of 65% GGBFS on the properties of RCA, Ann et al. [12] discovered that GGBFS enhances the resistance of RCA to permeability and corrosion when used with 100% RCA. Afroughsabet et al. [13] studied the impact of adding 30% GGBFS to RCA. They found that substituting 30% OPC with GGBFS increased both the concrete splitting tensile and flexural strength while maintaining its Cs. Slag was shown to reduce both the water absorption and workability of concrete. RC beams using recycled aggregate from electric arc furnaces (EAF) were the subject of a study to determine how they would behave [14]. According to the experimental results, EAF concrete beams have enhanced material qualities, resulting in higher ultimate flexural and shear capacities and smaller crack widths than equivalent standard RC beams. Another study [15] concluded that utilizing ternary cements, which have the right proportions of blast-furnace slag and limestone filler, can minimize gaseous emissions without sacrificing the cement mechanical qualities and promote the effective use of by-products and natural resources.

One of the most crucial aspects to consider while building and engineering structures is testing the concrete Cs. Typically, the compression test for the specimens at the specified curing age is used in the lab to determine the Cs. However, in addition to time and money, a significant amount of material is needed to carry out this test. As a result, some researchers have developed a technique to estimate the Cs of eco-friendly concrete utilizing RAC as a replacement material for natural aggregate or GGBFS as a replacement material for ordinary cement.

2. Related Works

In civil engineering, supervised machine learning models have acquired a lot of momentum, especially for predicting eco-friendly concrete properties, because they can forecast the results with high accuracy. Researchers have presented many machine learning (ML)-based prediction models for forecasting the mechanical and physical properties of RCA. The suggested approaches have been validated as a reliable alternative to expensive and time-consuming testing in laboratory usually used to assess concrete qualities. For example, Han et al. [16] presented an ensemble ML model for concrete derived from RCA to estimate the concrete modulus of elasticity. It is demonstrated that when compared to stand-alone models, the ensemble ML model consistently generates more precise predictions. Additionally, prediction models for the C_s of concrete incorporating recycled materials were suggested using random forest, linear regression, and nonlinear regression techniques [17]. The models based on random forest and nonlinear regression were found to be more accurate than linear regression. Furthermore, Liu et al. [18] used ML algorithms to predict the carbonation depth in RCA. According to the results, the random forest model outperforms the stand-alone artificial neural network (ANN) model and Gaussian process regression model. Last but not least, 10 ML algorithms were tested using a dataset containing 962 experimental C_s of NCA and RCA [19]. As they outperform existing methods, the ML models established for this work can be suggested as a valuable tool for the prediction of the C_s . In term of GGBFS, numerous studies established an ML model to forecast the mechanical characteristics of eco-friendly concrete that includes GGBFS material as a substitute for regular cement. The NN and ANFIS models for predicting chloride permeability in concrete were introduced by Boğa et al. [20]. Their study used concrete specimens containing solely calcium nitrite-based corrosion inhibitor (CNI), GGBFS, and a mixture of these ingredients in various ratios. The evaluation of the results demonstrated that both models estimate the permeability of chloride with good precision. In addition, random forest was used by Mai et al. [21] to forecast the C_s of GGBFS-containing concrete. RF performances in terms of R^2 , RMSE, and MAE were 0.9729, 4.9585, and 3.9423, respectively. Han et al. [22] created an innovative hybrid model for calculating the C_s of GGBFS concrete and validated the synergistic benefits of the hybrid algorithm over a single algorithm. The new PSO-BP hybrid neural network model outperformed basic ANNs trained by a single method and was shown to be suited for estimating the C_s of GGBFS concrete. Table 1 provides a comprehensive summary of relevant prior work related to predicting the C_s of eco-friendly concrete.

Table 1. Various algorithms to predict the Cs of eco-friendly concrete.

Algorithm Used	Data Points	Reference	Year	R ²	RMSE (MPa)	Replacement Material Used	Limitation
Gene Expression Programming (GEP)	251	[23]	2017	–	7.9	RCA	No GGBFS material included
M5 model tree (M5)	156	[24]	2020	–	8.3	RCA	No GGBFS material included
Multivariate Adaptive Regression Splines (MARS)	156	[24]	2020	–	9.1	RCA	No GGBFS material included
Least Squares Support Vector Regression (LSSVR)	156	[24]	2020	–	7.7	RCA	No GGBFS material included
convolutional neural network (CNN)	74	[25]	2018	–	–	RCA	No GGBFS material included
Multiple nonlinear regression (MNR)	650	[26]	2019	0.027	11.94	RCA	No GGBFS material included
imperialist competitive algorithm-extreme gradient boosting (ICA-XGBoost)	209	[27]	2021	0.983	1.147	RCA	No GGBFS material included
imperialist competitive algorithm—adaptive network-based fuzzy inference system (ICA-ANFIS)	209	[27]	2021	0.940	2.770	RCA	No GGBFS material included
imperialist competitive algorithm-Artificial neural network (ICA-ANN)	209	[27]	2021	0.960	2.225	RCA	No GGBFS material included
imperialist competitive algorithm-Support Vector Regression (ICA-SVR)	209	[27]	2021	0.962	2.149	RCA	No GGBFS material included
Backpropagation neural network models (BPNN)	344	[28]	2020	0.828	6.639	RCA	No GGBFS material included
Random Forest (RF)	453	[21]	2021	0.946	4.958	GGBFS	No RCA material included
ANN	284	[29]	2009	0.981	2.511	GGBFS	No RCA material included
ANFIS	284	[29]	2009	0.968	3.379	GGBFS	No RCA material included
Hybridized multiobjective ANN and a multiobjective slap swarm algorithm (MOSSA)	624	[30]	2020	0.941	2.39	GGBFS	No RCA material included
M5P model tree algorithm	624	[30]	2020	0.883	4.60	GGBFS	No RCA material included
ANN model	269	[22]	2019	0.961	3.332	GGBFS	No RCA material included

3. Research Significance

A powerful ensemble learning approach built on a gradient boosting system is called a gradient boosting regression tree (GBRT) [31]. To be more precise, GBRT is a robust data-mining technique that has been extensively tested and shown to be successful in various classification and regression problems [32–43]. As a result, in this study the prediction of the Cs of eco-friendly concrete was chosen to demonstrate the potential of the GBRT technique. Additionally, adjusting the hyperparameters of GBRT models for eco-friendly concrete datasets is beneficial. An essential step in the ML process is hyperparameter tuning (optimization). A wise selection of hyperparameters may either help a model to achieve the intended metric value or, on the other hand, cause it to enter an endless loop

of iterative training and optimization. Thus, the hyperparameters of GBRT are optimized using the GridSearchCV technique. Since the prediction of the Cs of eco-friendly concrete that contains recycled aggregate (RA) as a replacement for NA and GGBFS material as a replacement for OPC has never been studied in the way that this study does, this is a novel piece of research. This article has the following format. The research technique for this work is outlined in Section 2, and Section 3 displays a description of the dataset. The outcomes from the model prediction and the comparative analysis are described in Section 4. Section 5 provides a thorough analysis of the significance and contribution of each input variable to the final Cs. Finally, Section 6 provides a summary of the main conclusion.

4. Materials and Methods

4.1. Research Methodology

The methodology of the study is displayed in Figure 1. The data were first gathered, then divided into training and testing datasets: 80% and 20% of the total data, respectively. The GBRT technique for generating models and the GridSearchCV approach for determining precise model parameters were then introduced. The evaluation and interpretation of the GSC-GBRT model made up the last phase. A performance assessment was carried out based on the statistical performance measurement metrics. Related metrics were used to quantify the evaluation results, including RMSE and R^2 measurements. The notion of SHAP was finally introduced throughout the interpretation process, and global and local assessments were carried out. The Python 3.7 Scikit-learn software [44] was used to model and tune the GBRT in order to produce GSC-GBRT.

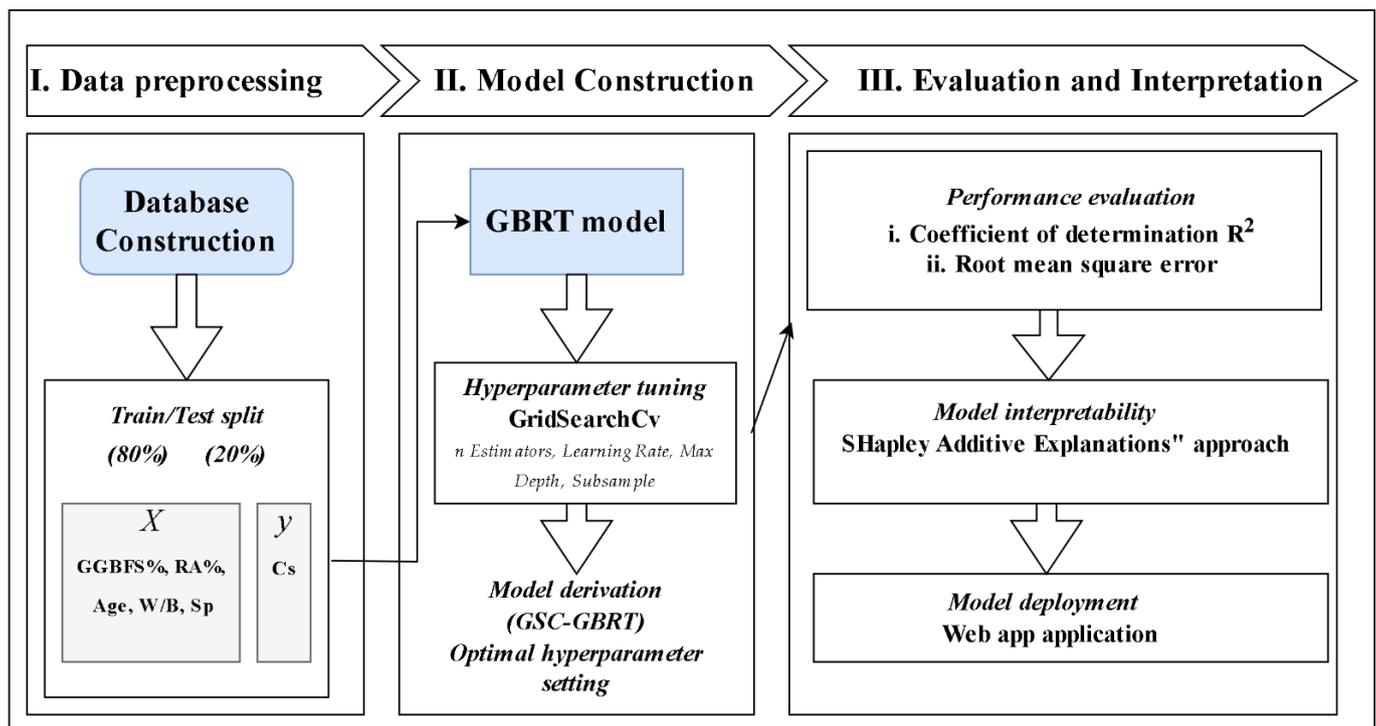


Figure 1. Research methodology.

4.2. Gradient Boosting Algorithm

Let us consider x as a collection of random inputs variable $x = \{x_1, x_2, \dots, x_n\}$ and y as response variable. Using a training data in the form of $\{(x_i, y_i)\}$ for $i = 1, 2, \dots, N$ with $x_i \in R^n$ and $y_i \in R$, the goal is finding an approximation $\tilde{F}(x)$ of the function $F(x)$ mapping x to y , to minimize loss function $L(y, F(x))$, see Equation (1). Errors are inevitable when it is expected to seek function $\tilde{F}(x)$. Each weak learner model seeks to correct errors produced by earlier weak learner models as the gradient boosting approach fits weak learners to the

loss functions. Because of this, the performance of the prediction model may be improved, and prediction error can be decreased.

$$\tilde{F}(x) = \operatorname{argmin}_{F(x)} L_{y,x}(y, F(x)) \quad (1)$$

The estimation of the approximation function as $L(y, F(x)) = (y - F(x))^2$ needs the squared error function to be used as the loss function. The gradient boosting technique uses the steepest descent step to minimize the loss function after establishing an initial base learner $F_0(x)$, which is typically a constant function in step 1. Step 2 requires defining the iteration space for $m = 1, \dots, M$. Finding the local minimum requires the steepest descent, which makes steps proportional to the negative gradient of the loss function.

Specifically, the following equation can be used to determine the gradient of the loss function $L(y, F(x))$ (step 3):

$$\hat{y}_i = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}, i = 1, \dots, N \quad (2)$$

When regression trees $h(x_i; a)$ are employed with parameter a as weak learners, it can generalize the range of the gradient calculation. The a parameters define it as a parameterized function of the input variables x [45]. The equation below can be solved to produce the tree (step 4):

$$a_m = \operatorname{argmin}_{a,\beta} \sum_{i=1}^N [\hat{y}_i - \beta h(x_i; a)]^2 \quad (3)$$

where β is the weight value, commonly known as the expansion coefficient of each weak learner, and a_m is the parameters discovered at iteration m . The current negative gradient is fitted to each regression tree. The model $F_m(x)$ is then updated at step 6 at each iteration m , with $m = 1, \dots, M$, after step 5 determines the ideal length p_m . Algorithm 1 from Friedman formalizes the gradient boosting algorithm [29], see Algorithm 1.

Algorithm 1: gradient boosting.

- 1: $F_0(x) = \operatorname{argmin}_{\rho} \sum_{i=1}^N L(y_i, \rho)$
 - 2: For $m = 1$ to M do;
 - 3: $\tilde{y}_i = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}, i = 1, \dots, N$
 - 4: $a_m = \operatorname{argmin}_{a,\beta} \sum_{i=1}^N [\tilde{y}_i - \beta h(x_i; a)]^2$
 - 5: $\rho_m = \operatorname{argmin}_{\rho} \sum_{i=1}^N L(y_i - F_{m-1}(x_i) + \rho h(x_i; a_m))$
 - 6: $F_m(x) = F_{m-1}(x) + \rho_m h(x; a_m)$
 - 7: End for
 - 8: End algorithm
-

4.3. Gradient Boosting Regression Tree Algorithm (GBRT)

Lie et al. [46] introduced classification and regression trees (CARTs) in 1984. CARTs can be used for regression and classification models [31–35]. The trees utilized in these two models are known as decision trees, and the development of decision trees involves using recursive techniques to produce binary trees. The technique that creates regression trees using the square error minimization criteria is mainly discussed here since the aim is to research concrete Cs predictions. It is noted that He et al. [47] introduced the GBRT method, which combines the CART algorithm with the GB algorithm. Because it can represent non-linear interactions without requiring previous knowledge of the probability distribution of variables, it is noted that the CART has highest excellent prediction performance than most artificial intelligence models [47]. As mentioned before, the gradient boosting algorithm combines weak and robust learners. Regression trees produced by the CART method serve as weak learners in this investigation. In order to further minimize the prediction error and

raise the model accuracy, the weak learners are added to the model to correct the prediction errors created by previous models.

The formalization of GBRT Algorithm 2 is presented in Algorithm 2.

Algorithm 2: GBRT.

1: $F_0(x) = \operatorname{argmin}_c \sum_{i=1}^N L(y_i, c)$
 2: $m = 1$ to M do;
 3: $r_{m,i} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}, i = 1, \dots, N$
 4: $c_{m,j} = \operatorname{argmin}_c \sum_{x_i \in R_{m,j}} L(y_i, F_{m-1}(x) + c)$
 5: $F_m(x) = F_{m-1}(x) + \sum_{j=1}^{J_m} c_{m,j} I(x \in R_{m,j})$
 6: $F_M(x) = F_0(x) + \sum_{m=1}^M \sum_{j=1}^{J_m} c_{m,j} I(x \in R_{m,j})$
 7: End for
 8: End algorithm

The initial value of $F_0(x)$ is set by the GBRT algorithm using the following equation (step 1).

$$F_0(x) = \operatorname{argmin}_c \sum_{i=1}^N L(y_i, \rho_0) \quad (4)$$

$L(\cdot)$ in the previous equation is a loss function. The following equation is used to determine the value of the current loss function negative gradient in the model for the residual approximation at iteration m , where $m = 1, 2, \dots, M$. (step 3).

$$r_{m,i} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}, i = 1, \dots, N. \quad (5)$$

Each regression tree divides the input space J_m into disjoint regions $R_{m,1}, \dots, R_{m,J_m}$ and predicts a value $c_{m,j}$ for region $R_{m,j}$ based on the assumption that there are J_m splits. The following equation can be minimized to get the value of $c_{m,j}$ (step 4).

$$c_{m,j} = \operatorname{argmin}_c \sum_{x_i \in R_{m,j}} L(y_i, F_{m-1}(x) + c) \quad (6)$$

The updated model, or the m -th regression tree $F_m(x)$, whose corresponding leaf node area is $R_{m,j}$ $j = 1, 2, \dots, J_m$, may be calculated as follows (step 5).

$$F_m(x) = F_{m-1}(x) + \sum_{j=1}^{J_m} c_{m,j} I(x \in R_{m,j}) \quad (7)$$

where J_m is a representation of the m -th regression tree number of leaf nodes and $I = 1$ if $x \in R_{m,j}$ and $I = 0$ otherwise. At step 6, the model is lastly updated.

4.4. Hyperparameter Tunning with GridSearchCV

In this research, many models were trained on the dataset for almost every ML project before choosing the one that performs the best. However, there is still potential for improvement because there is no certainty that this specific model is the best for the issue at hand. As a result, the aim is to make the model better in whatever manner. These models' hyperparameters play a crucial role in how well they function: if the correct values are chosen for these hyperparameters, the performance of the model performance can advance considerably. Grid search cross-validation (GridSearchCV) was used to choose the best model for each ML approach. With the parameters that produced the better cross-validation performance, a new model is automatically fitted using this method to the whole training dataset. This method aids in obtaining a more accurate generalization performance estimate. For example, with $k = 5$, the K-fold cross-validation procedure was used. A portion of the data is used in the K-fold cross-validation to test the model and

another portion to fit it. The prediction error from Equation (8) is then estimated using cross-validation as follows:

$$CV(f) = \frac{1}{n} \sum_{i=1}^n T(y_i, f^{-k(i)}(x_i)) \tag{8}$$

where k is the number of subsets, n is the size of the dataset, T is the loss function, and $f^{-k(i)}$ is the fitted function. The GridSearchCV methodology utilized in this study for model training and hyperparameter selection is shown in Figure 2 as demonstrated in [48].

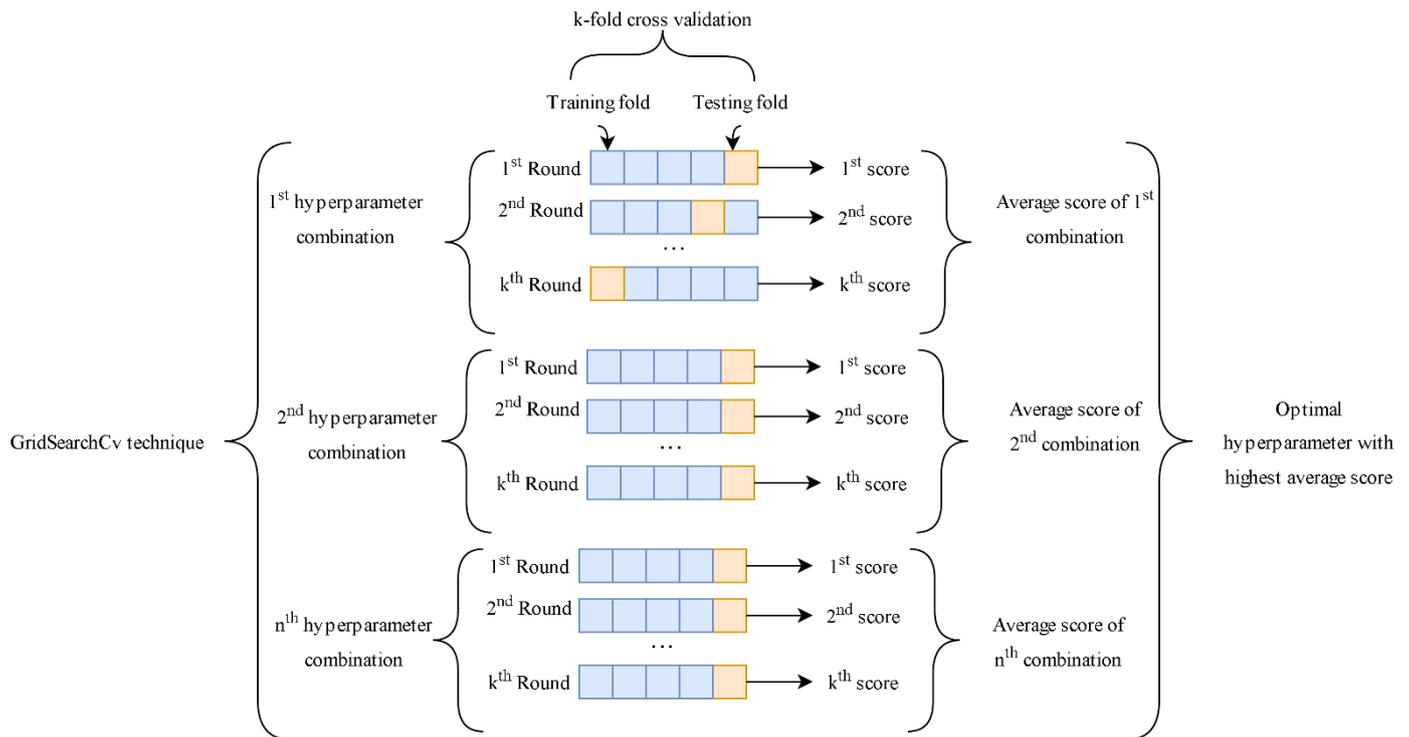


Figure 2. Tuning of hyperparameters using fivefold cross-validation (GridSearchCV).

4.5. Model Interpretation with the SHAP Method

In this research, SHAP was applied to the GSC-GBRT model output of the predicted values to explain them. This technique is known as the decoupling of each input parameter’s effect on the Cs of a particular mixture sample. A model of explanation developed by SHAP is expressed as follows:

$$g(z') = \phi_0 + \sum_{j=1}^K \phi_j z'_j \tag{9}$$

where ϕ_0 is the constant if all inputs are missing, ϕ_j stands for the i -th feature contribution value, and $z' \in$ and K represent the number of input features. A mechanism is created to determine how much each input information adds to the value that the model generates. The Shaply value is the one that most closely resembles human intuition and fits the three criteria that the additive feature attribution approach should meet (local accuracy, missingness, and consistency) [49]. Shaply values indicate the extent that each predictor (feature) contributes to a machine learning model.

Two models are trained: f_{S_U} when a particular feature i is included and f_S when it is not, to investigate the impact of that feature on the model. The variance in the results obtained from these two models for a given input x_S reveals how feature i affected the model. This theoretical idea serves as the foundation for the final calculation of the Shapely

value, which represents the contribution of each feature, as the weighted average of all potential differences, as indicated in the following equation:

$$\phi_j = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} \left[f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_S(x_S) \right] \quad (10)$$

where S represents the set of all features excluding i , F represents the set of all features, and f is the prediction/estimation model.

4.6. Performance Metrics

The coefficient of determination (R^2), as indicated by Equation (11), was used to evaluate the accuracy of the training and testing datasets for each model.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i^{obs} - y_i^{pre})^2}{\sum_{i=1}^n (y_i^{obs} - y_i^{-obs})^2} \quad (11)$$

where y_i^{pre} and y_i^{obs} represent the predicted output and the actual outcome (Cs), respectively; n denotes the number of data used in the Cs modeling, and y_i^{-obs} is the mean value of the real outcome.

The root mean square error (RMSE) of the algorithm prediction was computed from Equation (12). The comparison between RMSE values can be used to assess the optimization process since a more accurate model has a comparatively lower RMSE value.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^{obs} - y_i^{pre})^2} \quad (12)$$

Another performance metric used in this study is the variance accounted for (VAF). The following equation was used to compute this metric.

$$VAF = \left[1 - \frac{\text{VAR}(y_i^{obs} - y_i^{pre})}{\text{VAR}(y_i^{obs})} \right] \times 100 \quad (13)$$

When the RMSE, R^2 , and VAR become closer to 0, 1, and 100, respectively, the accuracy of the model prediction increases.

5. Dataset Used

Various studies have investigated the Cs of environmentally friendly concrete. As a result, an extensive dataset with 164 experiments on the Cs of eco-friendly concrete incorporating both RA and GGBFS material was recently assembled in reference [50]. The ML model for predicting the concrete Cs was trained and tested using this dataset. Details can be consulted in [50]. According to [39], the data collected from the collected experimental studies consider the effect of both GGBFS and RCA on the Cs of concrete. To unify their results, when collecting the data, specimens with only cube shapes of either 100 mm or 150 mm in length were considered. The authors used Rashid and Mansur's equation to transform the Cs from a 100 mm cube specimen to the equivalent Cs from a 150 mm cube specimen. In addition, five of the data records had no RA and/or GGBFS in their combination proportions, indicating standard concrete blends.

After being randomly sorted for the aim of constructing models, the gathered data records were divided into two fragments. First, the data records were split into training and testing data with 80% and 20%, respectively. Some relevant key parameters of the dataset utilized in this work is shown in Table 2. The primary input factors that have a significant impact on the concrete Cs were the following ones: water/binder ratio (W/B), curing time (Age), the recycled aggregate percentage from total aggregate in the mixture

(RA%), GGBFS material percentage from total binder used in the mixture (GGBFS%), and superplasticizer content (kg). Table 2 presents the lowest, average, median, standard deviation, and maximum values of the input variables for the training and testing sets.

Table 2. Input and output key parameters for the dataset.

Data Category	Statistics	Sp (kg)	RA%	Age (days)	W/B	GGBFS%	CS (MPa)
Training data	Standard deviation	2.12	38.22	29.52	0.10	26.14	11.52
	Mean	1.57	58.70	34.54	0.48	33.47	33.81
	Median	0.76	50.00	28.00	0.50	30.00	33.30
	Maximum	7.80	100.00	90.00	0.75	90.00	65.00
	Minimum	0.00	0.00	7.00	0.25	0.00	11.00
	Kurtosis	2.56	−1.48	−0.30	1.34	−1.14	−0.49
Testing data	Standard deviation	2.27	38.80	31.27	0.10	24.91	11.89
	Mean	1.71	70.91	35.85	0.46	31.21	35.38
	Median	0.76	100.00	28.00	0.50	40.00	35.28
	Maximum	7.80	100.00	90.00	0.75	80.00	68.00
	Minimum	0.00	0.00	7.00	0.25	0.00	12.00
	Kurtosis	2.73	−0.80	−0.62	1.14	−1.31	0.53

Figure 3 displays the relationships and statistical distributions of the Cs and the concrete components. Notably, none of the components are correlated, meaning that all the variables considered to forecast the concrete Cs are independent. Additionally, it can be seen that there is only a weak correlation between the W/B ratio and the Cs.

A heatmap, as shown in Figure 4, clearly summarizes the association of the entire dataset (containing all of the ratios, weight, percentages, age, and Cs of concrete). It is possible to plot the heatmap in Python by using the Seaborn library. A correlation index near 1 indicates that the features are highly connected. A negative correlation value close to −1 indicates a perfect correlation of two features but moving in opposing directions. In contrast, two uncorrelated features have a correlation index near 0. As it can be noticed in the heatmap, W/B, RA%, and GGBFS% have a negative correlation with the output, while Age and Sp have a positive correlation.

The parameters in Table 2 are utilized to produce environmentally friendly concrete. For completeness, a brief description of each variable is provided below to justify why they were considered concrete components in this study.

The W/B ratio has a significant influence on the concrete strength. Researchers noted that reducing the W/B ratio enhanced the Cs, splitting tensile strength, net flexural strength, and elastic modulus of self-compacting concretes fabricated using RA [51]. Moreover, the 28-day Cs of RCA concretes with a W/B ratio of 0.40 could be increased to be higher than the 28-day Cs of NA concrete with a W/B ratio of 0.50 and be closer to the 28-day Cs of NA concrete with a W/B ratio of 0.45 through the use of fly ash at replacement rates of 15 and 25% by weight of binder in RCA concretes [52]. In addition, the test findings demonstrated that the fly ash contributed more for the Cs at lower W/B ratios than it did in mixes made at higher W/B ratios.

The use of demolition waste as RA to create RAC has been investigated by several researchers [53,54]. However, RAC characteristics are inferior to NCA when porous mortar binds the RA [15]. Furthermore, comparing RCA to NCA, a 30–40% reduction in Cs was noticed [16]. Regardless, Aliabdo et al. [55] found that while replacing coarse aggregate with RA results in a negligible drop in concrete strength, substituting fine recycled aggregate results in a considerable reduction. Therefore, RA% was added as a necessary concrete element in the dataset used in this study.

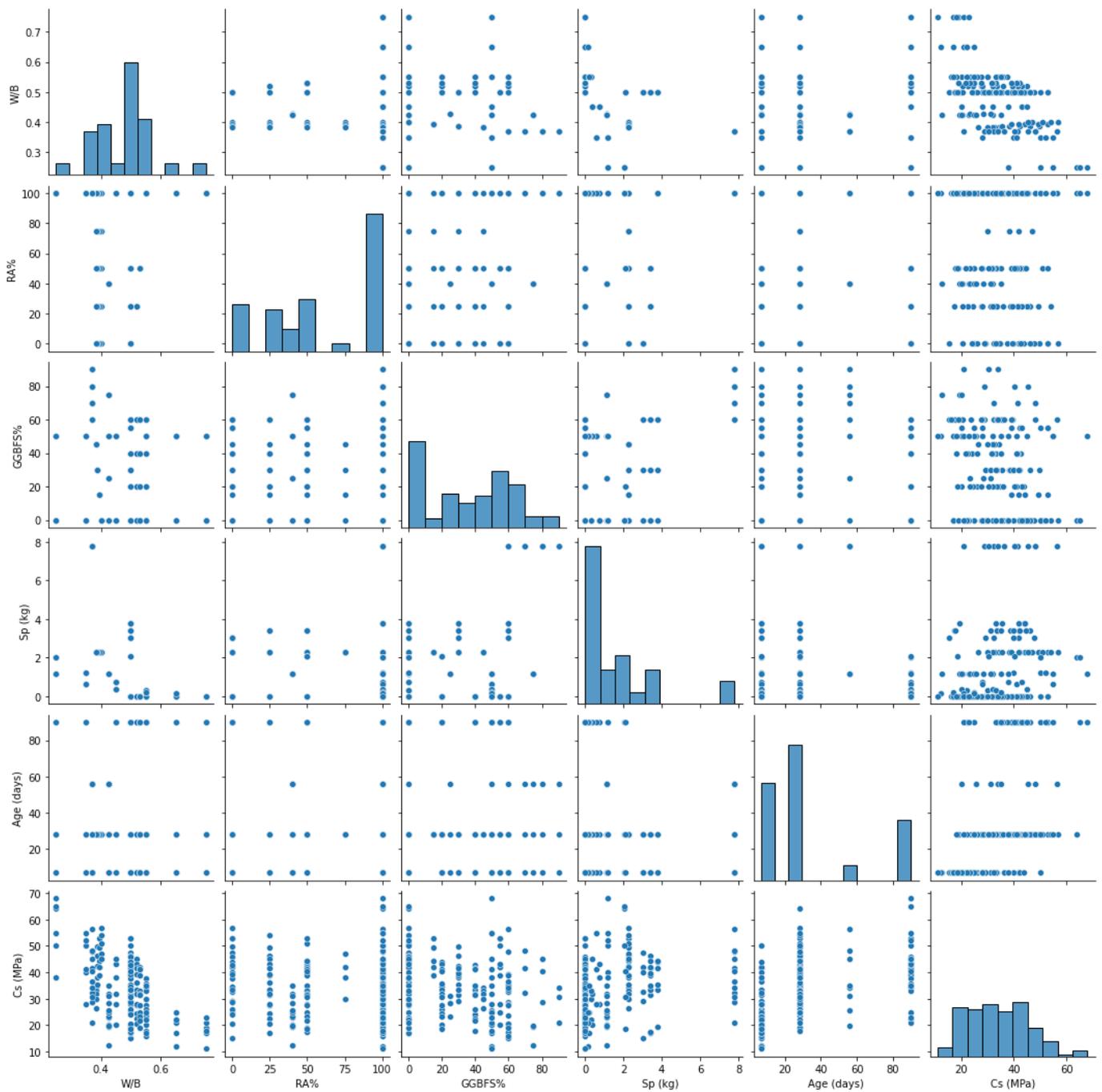


Figure 3. Correlations between the Cs and the primary concrete components in the dataset, as well as their statistical distributions.

With the incorporation of RA, Cs tends to drop; however, adding superplasticizers (Sp) can improve the mix compactness, recovering up most of the strength loss [56]. Cs increased when Sp dose was raised, and it was even more significant for Sp with higher water reduction capacities. With the proportion of added Sp, the mix density followed a similar pattern to Cs and slightly decreased towards a higher dosage [56].

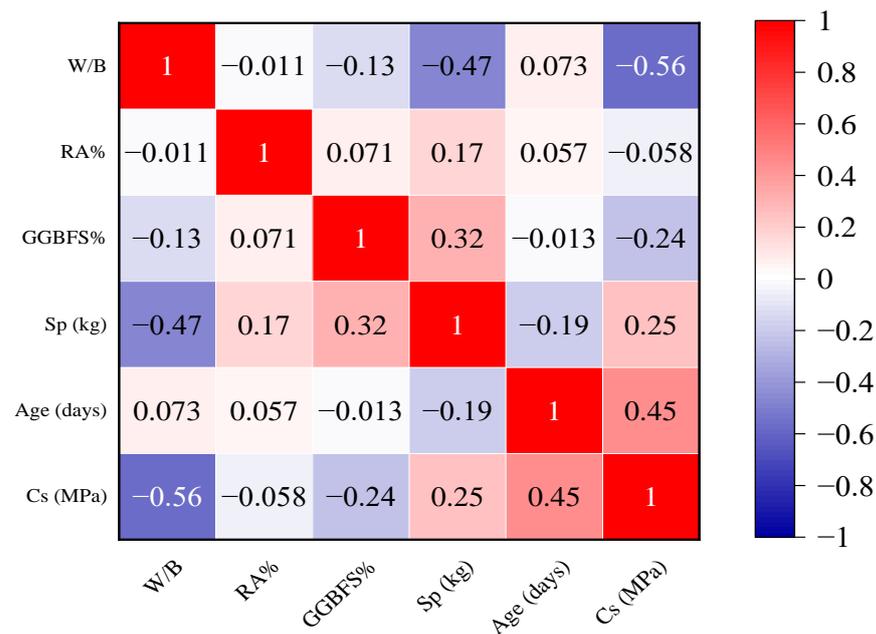


Figure 4. Heatmap showing the correlation between the dataset's Cs, age, and each eco-friendly concrete component.

Some researchers discovered that the Cs of all self-compacting concrete (SCC) mixes dropped with an increase in RCA for SCC incorporating GGBFS and RCA [57]. GGBFS as a cement replacement reduced the Cs at an early age compared to reference concrete, but at later ages (56 and 90 days), the Cs was equal or higher. The maximum Cs was achieved for SCC mixtures containing 15% GGBFS [57].

The development of durable concrete is determined by its age. When both RCA and GGBFS content are increased, the short-term and long-term Cs of concrete fall [10]. However, the strength growth rate with age increases when both RCA and GGBFS rise. This increase in strength growth rate results from GGBFS latent hydraulic activity and increased hydration of the unhydrated cement particles in RCA [10].

6. Model Results

6.1. Hyperparameter Optimization: GridSearchCV

Throughout applying the exhaustive parameter search GridSearchCV method, it was possible to find the parameters that matched the predicted model characteristics. Finding parameters with the optimal model estimation accuracy is easier using the Python GridSearchCV function [58]. The ideal GBRT hyperparameter combination was obtained by defining the values and ranges of the GBRT estimation model hyperparameter and using the GridSearchCV function. The learning rate and the n estimators, which stand for the weights given to each estimator and the number of the model weak learners, respectively, are the most crucial hyperparameters for the GBRT model. Additionally, the subsample parameter, which denotes the percentage of data to be utilized for fitting the individual base learners, and the max depth parameter, which defines the complexity of each tree, can both significantly impact the GBRT model ability to predict outcomes. When the following parameters are set: learning_rate = 0.05, max_depth = 4, n _estimators = 1000, subsample = 0.5, the GBRT using the GridSearchCV approach yields high estimate accuracy. The tuning parameters taken into account by GridSearchCV are listed in Table 3.

Table 3. Hyperparameters for the GBRT model.

Tuning Hyperparameter	Values and Ranges	Optimal Hyperparameters
n Estimators	[100, 500, 1000]	1000
Learning Rate	[0.01, 0.05, 0.1, 0.2]	0.05
Max Depth	[4, 6, 8, 10]	4
Subsample	[0.9, 0.5, 0.2, 0.1]	0.5

6.2. Comparison of the Prediction Results of the Two Models

The observed and predicted C_s values of the testing and training datasets are shown in Figure 5 to understand better the performance of the GSC-GBRT. Figure 5 demonstrates that, when compared to the default GBRT predictive model, the C_s outcomes predicted by the GSC-GBRT model were more in line with their observed values. In addition, the prediction accuracy of the GSC-GBRT test set after the hyperparameters' modification was higher than that of the default GBRT model, according to the optimized hyperparameter findings (see Table 4). For example, the R^2 and RMSE values for the GSC-GBRT model were 0.9612 and 2.3214, respectively, whereas they were 0.9216 and 3.4390 for the GBRT model. This demonstrates that, for the eco-friendly concrete dataset used for the prediction procedure of the C_s , GSC-GBRT could better match the complicated connection between the component factors influencing the C_s and had superior generalization capacity.

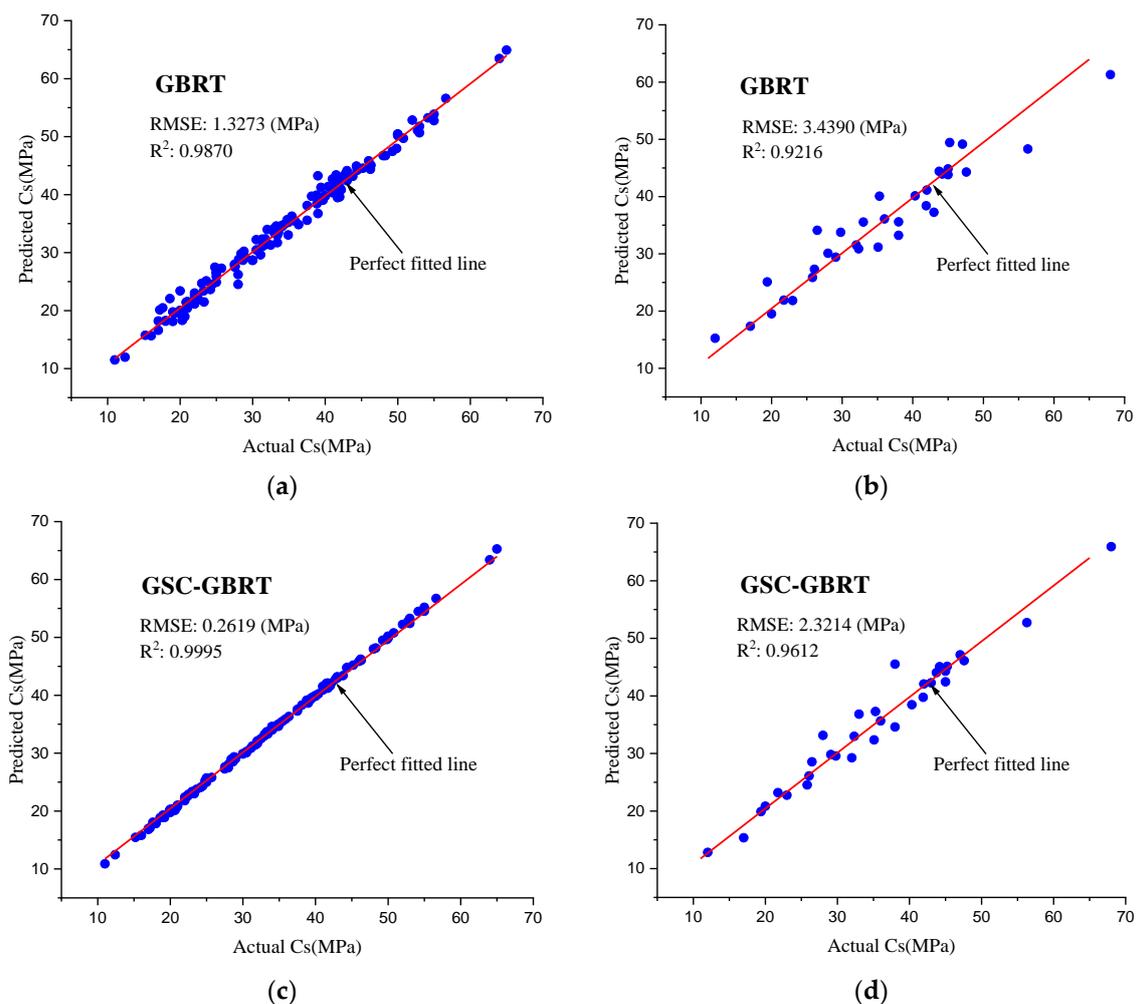
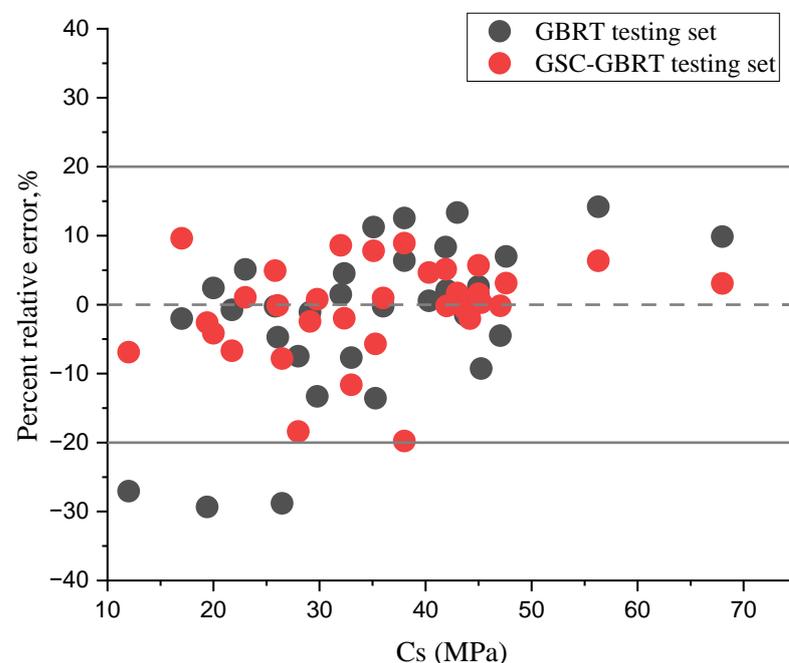


Figure 5. Results of the GSC-GBRT and GBRT models for the testing and training datasets. (a) GBRT training set. (b) GBRT testing set. (c) GSC-GBRT training set. (d) GSC-GBRT testing set.

Table 4. GSC-GBRT and GBRT model performance in predicting the C_s for training and test datasets.

Sets	Model	RMSE	R^2	VAF(%)
Training	GSC-GBRT	0.2619	0.9995	99.95
	GBRT	1.3273	0.9870	98.66
Testing	GSC-GBRT	2.3214	0.9612	96.07
	GBRT	3.4390	0.9216	91.40

Figure 6 displays the distribution of the relative percentage error between the values of the GBRT and GSC-GBRT models and the C_s of environmentally friendly concrete. Based on this graph, it can be stated that the relative error distributions around the zero-error line for both models during testing are suitable. However, as shown by the reported findings in Figure 6, the GSC-GBRT model appears more capable of making predictions than the GBRT model.

**Figure 6.** Error distribution for the developed models.

The results of residuals produced by GSC-GBRT and default GBRT prediction of the C_s of environmentally friendly concrete mixtures at various ages are shown in Figure 7. Each figure shows red dots for the predicted values and black square for measured values. The predicted and measured values of the training set was relatively close for both methods. However, some of the GBRT calculations included inaccuracies exceeding 5 MPa. As it can be noticed, the testing set analysis showed that while GSC-GBRT maintained relatively constant prediction findings, several GBRT predicted points significantly differed from the actual C_s values.

The multiple linear regression and M5P models [59,60] are compared with the GSC-GBRT model in order to more accurately depict the accuracy of the last one. Multiple linear regression and M5P were used to obtain the predictions, and an 8:2 dataset division ratio was used. Figure 8 depicts the correlation between the estimated and actual C_s .

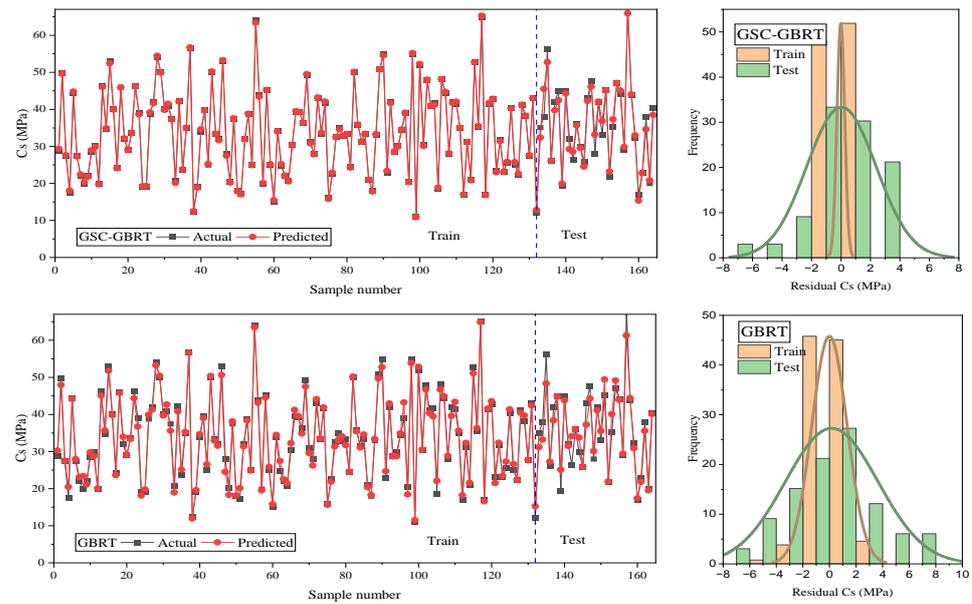


Figure 7. Model errors between targets and predictions for testing set (a) GSC-GBRT, (b) GBRT.

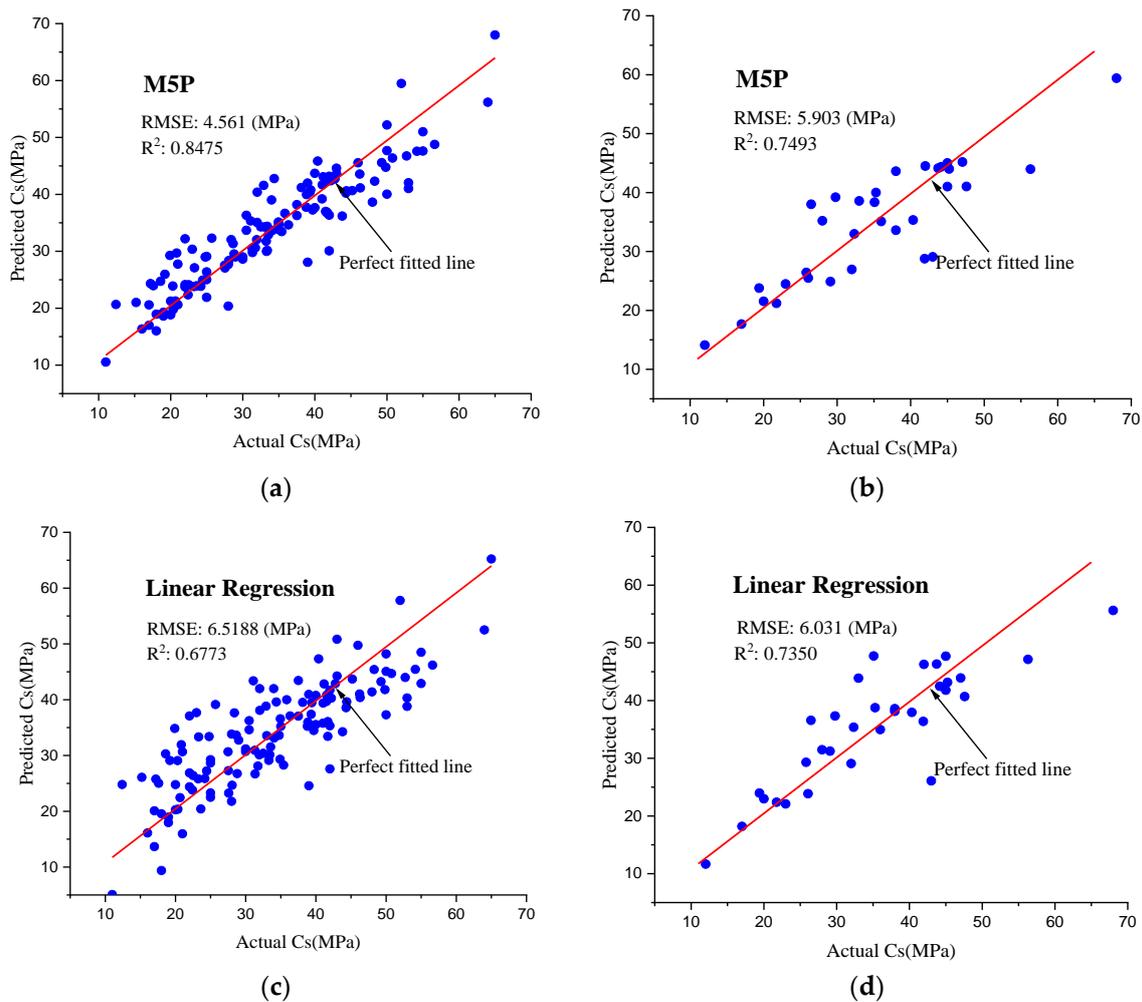


Figure 8. Results of the M5P and linear regression models for the testing and training datasets. (a) M5P training set. (b) M5P testing set. (c) Linear regression training set. (d) Linear regression testing set.

Table 5 displays the evaluation results for the test set using multiple linear regression and M5P. It can be shown that the GSC-GBRT ensemble learning model outperforms both of the other models. Multiple weak learners produced by different learning algorithms are combined in ensemble learning. To encourage more accurate predictions, weak learners who perform well are given larger weights, while weak learners who perform poorly are given lower weights. On the test set, the GSC-GBRT ensemble learning model performed the best. When compared to the linear regression, the R^2 was up 31%, and the RMSE was down 62%. R^2 increased by 28%, and RMSE dropped by 61% when compared to the M5P model. Overall, the ensemble learning model appears to perform significantly better for the eco-friendly concrete Cs prediction than the traditional machine learning methods.

Table 5. GSC-GBRT, linear regression, and M5P models' performance in predicting Cs for training and test datasets.

Model	RMSE	R^2	VAF(%)
GSC-GBRT	2.3214	0.9612	96.07
Linear regression	6.031	0.735	73.49
M5P	5.903	0.7493	74.87

7. Interpretation of the GBRT Model

In the absence of supporting theories, mathematical computations, or operational processes, the output findings or predictions of ML modeling can be challenging to explain [61]. However, the contributions of input variables can be analyzed using feature importance, sensitivity analysis, or partial dependency analysis in order to evaluate those outputs and comprehend the trained models. Here, the multicollinearity problem and potential synergistic effects of the variables were effectively treated using the SHAP approach. In this section, the “SHapley Additive Explanations” approach [62], combined with the GSC-GBRT model, explains and clarifies the contribution of each input variable to the concrete Cs prediction. According to Figure 9a, which displays the mean absolute SHAP values for each feature in the Cs modeling, the age in the concrete mixture has the highest mean SHAP value among the five input features. The descending order of input variable impact on GSC-GBRT model prediction accuracy is: AGE > W/B > GGBFS% > SP > RCA%.

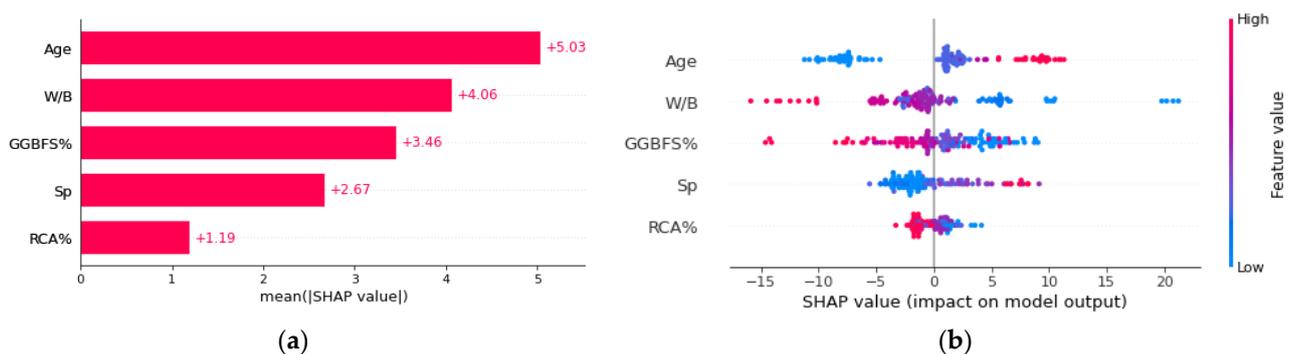


Figure 9. The relative importance of each feature and SHAP summary plot: (a) relative importance; (b) SHAP summary plot.

The input features SHAP values are shown in Figure 9b, with red denoting a high feature value and blue denoting a low feature value. The corresponding feature will benefit the output goal if the feature SHAP value is positive. When the SHAP value is more significant, the influence has a greater impact. For example, the age (red points) in Figure 9b exhibited noticeably high SHAP values, which showed that the higher curing age had a detrimental effect on the prediction of the Cs. It is widely established from Figure 9b that the W/B content, RCA%, and GGBFS% have a negative association with Cs. This is in line with earlier research [8,63].

A web application was developed using the suggested GSC-GBRT model to forecast the C_s of eco-friendly concrete at the Streamlit library. The user can theoretically forecast the C_s of eco-friendly concrete using the trained ML model based on the compiled dataset. W/B, RA%, GGBFS%, superplasticizer (kg), and age (days) are the first parameters the user enters using the online application sliders and radio buttons. In the next step, the concrete C_s in MPa is computed. The parameter ranges provided in the application match the feature ranges in the datasets used for ML training. The Streamlit web application can be accessed at the link given in [64].

8. Conclusions

This research proposed a hybrid GSC-GBRT model for the prediction of the C_s of sustainable concrete. The GridSearchCV approach was firstly used to find the optimum parameters, and the optimized model was then used to forecast the C_s . The hybrid GSC-GBRT model obtained higher prediction accuracy and reduced error with $R^2 = 0.9612$ and RMSE = 2.3214 when compared to the evaluation metrics of the original GBRT model with $R^2 = 0.9216$ and RMSE = 3.4390 for the test set. The suggested GSC-GBRT model surpasses the initial GBRT model in assessment metrics, and it is suggested to be used as a tool for pre-estimating the C_s of concrete using the mix ratio prior to design and mixing.

According to the SHAP-based research, W/B and age are the two input factors that most significantly impact the concrete C_s among the five considered. Age and superplasticizer positively impact the output, and the C_s rises as a result. On the other hand, a rise in W/B, GGBFS%, and RA% causes the C_s to fall. Therefore, designers and engineers can use the significance and contribution of these factors to the output outcomes as a guide. Finally, in Streamlit, a web application for predicting C_s of eco-friendly concrete, was developed. The cloud has been used to deliver the application light version. Any web browser, including mobile ones, can be used to access and use it.

The GSC-GBRT model has several limitations while being competent and acceptable for estimating the C_s of eco-friendly concrete. Firstly, 164 experiments from the literature were used to construct the eco-friendly concrete dataset. The accuracy of the prediction models is significantly influenced by the completeness of the data, quantity, quality, and distribution of the input parameters. As new experimental data become available, the dataset may benefit from being updated. Secondly, like any ML method, the SHAP explanations and GSC-GBRT findings may only apply to the tested input data ranges.

Author Contributions: Z.M.A.: conceptualization, modeling, and write up. Y.M.J.: visualization and review. S.N.H.: writing, validation and supervision. H.I.: review, writing original draft and funding. L.F.A.B.: review, writing and funding. H.M.H.: review, editing and visualization. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Sivakrishna, A.; Adesina, A.; Awoyera, P.; Kumar, K.R. Green concrete: A review of recent developments. *Mater. Today Proc.* **2020**, *27*, 54–58. [[CrossRef](#)]
2. Abd Elhakam, A.; Awad, E. Influence of self-healing, mixing method and adding silica fume on mechanical properties of recycled aggregates concrete. *Constr. Build. Mater.* **2012**, *35*, 421–427. [[CrossRef](#)]
3. Bairagi, N.; Ravande, K.; Pareek, V. Behaviour of concrete with different proportions of natural and recycled aggregates. *Resour. Conserv. Recycl.* **1993**, *9*, 109–126. [[CrossRef](#)]
4. Debieb, F.; Courard, L.; Kenai, S.; Degeimbre, R. Mechanical and durability properties of concrete using contaminated recycled aggregates. *Cem. Concr. Compos.* **2010**, *32*, 421–426. [[CrossRef](#)]

5. Kisku, N.; Joshi, H.; Ansari, M.; Panda, S.; Nayak, S.; Dutta, S.C. A critical review and assessment for usage of recycled aggregate as sustainable construction material. *Constr. Build. Mater.* **2017**, *131*, 721–740. [[CrossRef](#)]
6. Kou, S.; Poon, C.S. Enhancing the durability properties of concrete prepared with coarse recycled aggregate. *Constr. Build. Mater.* **2012**, *35*, 69–76. [[CrossRef](#)]
7. Sri Ravindrarajah, R.; Tam, C. Properties of concrete made with crushed concrete as coarse aggregate. *Mag. Concr. Res.* **1985**, *37*, 29–38. [[CrossRef](#)]
8. Moghadam, A.S.; Omidinasab, F.; Goodarzi, S.M. Characterization of concrete containing RCA and GGBFS: Mechanical, microstructural and environmental properties. *Constr. Build. Mater.* **2021**, *289*, 123134. [[CrossRef](#)]
9. Çakır, Ö. Experimental analysis of properties of recycled coarse aggregate (RCA) concrete with mineral additives. *Constr. Build. Mater.* **2014**, *68*, 17–25. [[CrossRef](#)]
10. Majhi, R.; Nayak, A. Bond, durability and microstructural characteristics of ground granulated blast furnace slag based recycled aggregate concrete. *Constr. Build. Mater.* **2019**, *212*, 578–595. [[CrossRef](#)]
11. Rashad, A.M. A Brief Review on Blast-Furnace Slag and Copper Slag as Fine Aggregate in Mortar and Concrete Based on Portland Cement. *Rev. Adv. Mater. Sci.* **2016**, *44*, 221–237.
12. Ann, K.Y.; Moon, H.; Kim, Y.; Ryou, J. Durability of recycled aggregate concrete using pozzolanic materials. *Waste Manag.* **2008**, *28*, 993–999. [[CrossRef](#)] [[PubMed](#)]
13. Afroughsabet, V.; Biolzi, L.; Ozbakkaloglu, T. Influence of double hooked-end steel fibers and slag on mechanical and durability properties of high performance recycled aggregate concrete. *Compos. Struct.* **2017**, *181*, 273–284. [[CrossRef](#)]
14. De Domenico, D.; Faleschini, F.; Pellegrino, C.; Ricciardi, G. Structural behavior of RC beams containing EAF slag as recycled aggregate: Numerical versus experimental results. *Constr. Build. Mater.* **2018**, *171*, 321–337. [[CrossRef](#)]
15. Menéndez, G.; Bonavetti, V.; Irassar, E. Strength development of ternary blended cement with limestone filler and blast-furnace slag. *Cem. Concr. Compos.* **2003**, *25*, 61–67. [[CrossRef](#)]
16. Han, T.; Siddique, A.; Khayat, K.; Huang, J.; Kumar, A. An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete. *Constr. Build. Mater.* **2020**, *244*, 118271. [[CrossRef](#)]
17. Dabiri, H.; Kioumars, M.; Kheyroddin, A.; Kandiri, A.; Sartipi, F. Compressive strength of concrete with recycled aggregate; a machine learning-based evaluation. *Clean. Mater.* **2022**, *3*, 100044. [[CrossRef](#)]
18. Liu, K.; Alam, M.S.; Zhu, J.; Zheng, J.; Chi, L. Prediction of carbonation depth for recycled aggregate concrete using ANN hybridized with swarm intelligence algorithms. *Constr. Build. Mater.* **2021**, *301*, 124382. [[CrossRef](#)]
19. Munir, M.J.; Kazmi, S.M.S.; Wu, Y.-F.; Lin, X.; Ahmad, M.R. Development of a novel compressive strength design equation for natural and recycled aggregate concrete through advanced computational modeling. *J. Build. Eng.* **2022**, *55*, 104690. [[CrossRef](#)]
20. Boğa, A.R.; Öztürk, M.; Topcu, I.B. Using ANN and ANFIS to predict the mechanical and chloride permeability properties of concrete containing GGBFS and CNI. *Compos. Part B Eng.* **2013**, *45*, 688–696. [[CrossRef](#)]
21. Mai, H.-V.T.; Nguyen, T.-A.; Ly, H.-B.; Tran, V.Q. Prediction Compressive Strength of Concrete Containing GGBFS using Random Forest Model. *Adv. Civ. Eng.* **2021**, *2021*, 6671448. [[CrossRef](#)]
22. Han, I.-J.; Yuan, T.-F.; Lee, J.-Y.; Yoon, Y.-S.; Kim, J.-H. Learned prediction of compressive strength of GGBFS concrete using hybrid artificial neural network models. *Materials* **2019**, *12*, 3708. [[CrossRef](#)] [[PubMed](#)]
23. Gholampour, A.; Gandomi, A.H.; Ozbakkaloglu, T. New formulations for mechanical properties of recycled aggregate concrete using gene expression programming. *Constr. Build. Mater.* **2017**, *130*, 122–145. [[CrossRef](#)]
24. Gholampour, A.; Mansouri, I.; Kisi, O.; Ozbakkaloglu, T. Evaluation of mechanical properties of concretes containing coarse recycled aggregate using multivariate adaptive regression splines (MARS), M5 model tree (M5Tree), and least squares support vector regression (LSSVR) models. *Neural Comput. Appl.* **2020**, *32*, 295–308. [[CrossRef](#)]
25. Deng, F.; He, Y.; Zhou, S.; Yu, Y.; Cheng, H.; Wu, X. Compressive strength prediction of recycled concrete based on deep learning. *Constr. Build. Mater.* **2018**, *175*, 562–569. [[CrossRef](#)]
26. Xu, J.; Zhao, X.; Yu, Y.; Xie, T.; Yang, G.; Xue, J. Parametric sensitivity analysis and modelling of mechanical properties of normal-and high-strength recycled aggregate concrete using grey theory, multiple nonlinear regression and artificial neural networks. *Constr. Build. Mater.* **2019**, *211*, 479–491. [[CrossRef](#)]
27. Duan, J.; Asteris, P.G.; Nguyen, H.; Bui, X.-N.; Moayedi, H. A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model. *Eng. Comput.* **2021**, *37*, 3329–3346. [[CrossRef](#)]
28. 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](#)]
29. Saridemir, M.; Topcu, İ.B.; Özcan, F.; Severcan, M.H. Prediction of long-term effects of GGBFS on compressive strength of concrete by artificial neural networks and fuzzy logic. *Constr. Build. Mater.* **2009**, *23*, 1279–1286. [[CrossRef](#)]
30. Kandiri, A.; Golafshani, E.M.; Behnood, A. Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm. *Constr. Build. Mater.* **2020**, *248*, 118676. [[CrossRef](#)]
31. Hao, X.; Hu, X.; Liu, T.; Wang, C.; Wang, L. Estimating urban PM_{2.5} concentration: An analysis on the nonlinear effects of explanatory variables based on gradient boosted regression tree. *Urban Clim.* **2022**, *44*, 101172. [[CrossRef](#)]
32. Wei, L.; Yuan, Z.; Zhong, Y.; Yang, L.; Hu, X.; Zhang, Y. An improved gradient boosting regression tree estimation model for soil heavy metal (Arsenic) pollution monitoring using hyperspectral remote sensing. *Appl. Sci.* **2019**, *9*, 1943. [[CrossRef](#)]

33. Shatnawi, A.; Alkassar, H.M.; Al-Abdaly, N.M.; Al-Hamdany, E.A.; Bernardo, L.F.A.; Imran, H. Shear Strength Prediction of Slender Steel Fiber Reinforced Concrete Beams Using a Gradient Boosting Regression Tree Method. *Buildings* **2022**, *12*, 550. [[CrossRef](#)]
34. Nyakilla, E.E.; Jun, G.; Kasimu, N.A.; Robert, E.F.; Innocent, N.; Mohamedy, T.; Shaame, M.; Ngata, M.R.; Mabeyo, P.E. Application of machine learning in the prediction of compressive, and shear bond strengths from the experimental data in oil well cement at 80 °C. Ensemble trees boosting approach. *Constr. Build. Mater.* **2022**, *317*, 125778. [[CrossRef](#)]
35. Nunez, I.; Marani, A.; Nehdi, M.L. Mixture optimization of recycled aggregate concrete using hybrid machine learning model. *Materials* **2020**, *13*, 4331. [[CrossRef](#)]
36. He, X.; Luo, J.; Li, P.; Zuo, G.; Xie, J. A hybrid model based on variational mode decomposition and gradient boosting regression tree for monthly runoff forecasting. *Water Resour. Manag.* **2020**, *34*, 865–884. [[CrossRef](#)]
37. Guo, R.; Fu, D.; Sollazzo, G. An ensemble learning model for asphalt pavement performance prediction based on gradient boosting decision tree. *Int. J. Pavement Eng.* **2021**, 1–14. [[CrossRef](#)]
38. Danandeh Mehr, A.; Fathollahzadeh Attar, N. A gradient boosting tree approach for SPEI classification and prediction in Turkey. *Hydrol. Sci. J.* **2021**, *66*, 1653–1663. [[CrossRef](#)]
39. Cheng, Z.; Yang, Y.; Zhang, H. Interpretable ensemble machine-learning models for strength activity index prediction of iron ore tailings. *Case Stud. Constr. Mater.* **2022**, *17*, e01239. [[CrossRef](#)]
40. Chen, S.-Z.; Zhang, S.-Y.; Han, W.-S.; Wu, G. Ensemble learning based approach for FRP-concrete bond strength prediction. *Constr. Build. Mater.* **2021**, *302*, 124230. [[CrossRef](#)]
41. Wan, Z.; Xu, Y.; Šavija, B. On the use of machine learning models for prediction of compressive strength of concrete: Influence of dimensionality reduction on the model performance. *Materials* **2021**, *14*, 713. [[CrossRef](#)] [[PubMed](#)]
42. Sarkhani Benemaran, R.; Esmaili-Falak, M.; Javadi, A. Predicting resilient modulus of flexible pavement foundation using extreme gradient boosting based optimised models. *Int. J. Pavement Eng.* **2022**, 1–20. [[CrossRef](#)]
43. Nguyen-Sy, T.; Wakim, J.; To, Q.-D.; Vu, M.-N.; Nguyen, T.-D.; Nguyen, T.-T. Predicting the compressive strength of concrete from its compositions and age using the extreme gradient boosting method. *Constr. Build. Mater.* **2020**, *260*, 119757. [[CrossRef](#)]
44. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *the J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
45. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
46. Li, B.; Friedman, J.; Olshen, R.; Stone, C. Classification and regression trees (CART). *Biometrics* **1984**, *40*, 358–361.
47. He, Q.; Kamarianakis, Y.; Jintanakul, K.; Wynter, L. Incident duration prediction with hybrid tree-based quantile regression. In *Advances In Dynamic Network Modeling in Complex Transportation Systems*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 287–305.
48. Liu, X.; Liu, T.; Feng, P. Long-term performance prediction framework based on XGBoost decision tree for pultruded FRP composites exposed to water, humidity and alkaline solution. *Compos. Struct.* **2022**, *284*, 115184. [[CrossRef](#)]
49. Mokhtari, K.E.; Higdon, B.P.; Başar, A. Interpreting financial time series with SHAP values. In Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering, Markham, ON, Canada, 4–6 November 2019; pp. 166–172.
50. Naser, A.H.; Badr, A.H.; Henedy, S.N.; Ostrowski, K.A.; Imran, H. Application of Multivariate Adaptive Regression Splines (MARS) approach in prediction of compressive strength of eco-friendly concrete. *Case Stud. Constr. Mater.* **2022**, *17*, e01262. [[CrossRef](#)]
51. Gesoglu, M.; Güneysi, E.; Öz, H.Ö.; Taha, I.; Yasemin, M.T. Failure characteristics of self-compacting concretes made with recycled aggregates. *Constr. Build. Mater.* **2015**, *98*, 334–344. [[CrossRef](#)]
52. Poon, C.S.; Lam, L.; Wong, Y. A study on high strength concrete prepared with large volumes of low calcium fly ash. *Cem. Concr. Res.* **2000**, *30*, 447–455. [[CrossRef](#)]
53. Munir, M.J.; Kazmi, S.M.S.; Wu, Y.-F.; Lin, X.; Ahmad, M.R. Axial stress-strain performance of recycled aggregate concrete reinforced with macro-polypropylene fibres. *Sustainability* **2021**, *13*, 5741. [[CrossRef](#)]
54. Kazmi, S.M.S.; Munir, M.J.; Wu, Y.-F.; Patnaikuni, I.; Zhou, Y.; Xing, F. Influence of different treatment methods on the mechanical behavior of recycled aggregate concrete: A comparative study. *Cem. Concr. Compos.* **2019**, *104*, 103398. [[CrossRef](#)]
55. Aliabdo, A.A.; Abd-Elmoaty, A.-E.M.; Hassan, H.H. Utilization of crushed clay brick in concrete industry. *Alex. Eng. J.* **2014**, *53*, 151–168. [[CrossRef](#)]
56. Dash, A.P.; Sahoo, K.; Panda, H.S.; Pradhan, A.; Jena, B. Experimental study on the effect of superplasticizer on workability and strength characteristics of recycled coarse aggregate concrete. *Mater. Today: Proc.* **2022**, *60*, 488–493.
57. Djelloul, O.K.; Menadi, B.; Wardeh, G.; Kenai, S. Performance of self-compacting concrete made with coarse and fine recycled concrete aggregates and ground granulated blast-furnace slag. *Adv. Concr. Constr.* **2018**, *6*, 103.
58. Lu, J.; Zhang, Y.; Chen, M.; Wang, L.; Zhao, S.; Pu, X.; Chen, X. Estimation of monthly 1 km resolution PM_{2.5} concentrations using a random forest model over “2+ 26” cities, China. *Urban Clim.* **2021**, *35*, 100734. [[CrossRef](#)]
59. Patil, S.V.; Rao, K.B.; Nayak, G. Prediction of recycled coarse aggregate concrete mechanical properties using multiple linear regression and artificial neural network. *J. Eng. Des. Technol.* **2021**, ahead-of-print. [[CrossRef](#)]
60. Behnood, A.; Behnood, V.; Gharehveran, M.M.; Alyamac, K.E.J.C.; Materials, B. Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm. *Constr. Build. Mater.* **2017**, *142*, 199–207. [[CrossRef](#)]

61. Lipton, Z.C. The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue* **2018**, *16*, 31–57. [[CrossRef](#)]
62. 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](#)]
63. Tran, V.Q.; Dang, V.Q.; Ho, L.S. Evaluating compressive strength of concrete made with recycled concrete aggregates using machine learning approach. *Constr. Build. Mater.* **2022**, *323*, 126578. [[CrossRef](#)]
64. Concrete Compressive Strength Prediction. Available online: <https://hamza19901990-eco-frendly-concrete-comp-comp-concrete-ml-4sasy1.streamlitapp.com> (accessed on 4 October 2022).