Article

Evaluating the Strength and Impact of Raw Ingredients of Cement Mortar Incorporating Waste Glass Powder Using Machine Learning and SHapley Additive ExPlanations (SHAP) Methods

Hassan Ali Alkadhim 1, Muhammad Nasir Amin 1,*, Waqas Ahmad 2, Kaffayatullah Khan 1, Sohaib Nazar 2, Muhammad Iftikhar Faraz 3 and Muhammad Imran 4

1 Department of Civil and Environmental Engineering, College of Engineering, King Faisal University, Al-Ahsa 31982, Saudi Arabia
2 Department of Civil Engineering, COMSATS University Islamabad, Abbottabad 22060, Pakistan
3 Department of Mechanical Engineering, College of Engineering, King Faisal University, Al-Ahsa 31982, Saudi Arabia
4 School of Civil and Environmental Engineering (SCEE), National University of Sciences & Technology (NUST), Islamabad 44000, Pakistan
* Correspondence: mgadir@kfu.edu.sa; Tel.: +966-13-589-5431; Fax: +966-13-581-7068

Abstract: This research employed machine learning (ML) and SHapley Additive ExPlanations (SHAP) methods to assess the strength and impact of raw ingredients of cement mortar (CM) incorporated with waste glass powder (WGP). The data required for this study were generated using an experimental approach. Two ML methods were employed, i.e., gradient boosting and random forest, for compressive strength (CS) and flexural strength (FS) estimation. The performance of ML approaches was evaluated by comparing the coefficient of determination ($R^2$), statistical checks, k-fold assessment, and analyzing the variation between experimental and estimated strength. The results of the ML-based modeling approaches revealed that the gradient boosting model had a good degree of precision, but the random forest model predicted the strength of the WGP-based CM with a greater degree of precision for CS and FS prediction. The SHAP analysis revealed that fine aggregate was a critical raw material, with a stronger negative link to the strength of the material, whereas WGP and cement had a greater positive effect on the strength of CM. Utilizing such approaches will benefit the building sector by supporting the progress of rapid and inexpensive approaches for identifying material attributes and the impact of raw ingredients.

Keywords: cement mortar; waste glass powder; building material; compressive strength; flexural strength

1. Introduction

Several practices, including manufacturing, mining, electricity generation, steel and iron metallurgy, production of electronic devices, etc., result in large volumes of solid waste [1]. Numerous harmful wastes are combustible, caustic, ignitable, virulent, and chemically reactive, and their dumping in landfills has caused substantial economic losses [2,3]. Therefore, it is desirable to recycle or reuse the solid waste in building materials [4–6]. In the building sector, cement mortar (CM) is commonly used [7–9]. Various strategies have been adopted by researchers to enhance the performance of CM [10,11]. For example, to improve the performance of CM, waste materials may be utilized as an alternative to aggregate [12,13], reinforcing fibers [14,15], and cement substitutes [16,17]. Due to the partial replacement of aggregates and cement, natural resources may be conserved, and CO$_2$ emissions might decrease [18–20]. It has also been noticed that using some waste materials enhances the performance of CM [21,22]. Globally, a substantial amount of glass waste (GW) is generated, with a substantial volume of GW discarded in landfills [23].
Although a growing number of towns are producing more municipal solid waste, landfill space is becoming gradually insufficient, especially in large cities [24]. Compared to other wastes (plastic and wood), GW is more chemically stable. Long-term burial of GW does not render it biodegradable [25]. Moreover, certain types of glass, such as cathode ray tubes (CRT), contain toxic elements, such as lead, cadmium, mercury, and beryllium, polluting underground soil and water [26]. China generates over 43 million tons of CRT glass annually [27], posing major environmental risks and endangering human health. GW may be recycled, crushed and utilized to partly substitute aggregates and cement in CM [28–30]. Prior research has demonstrated that GW recycling in CM is the most effective method [28,29]. Thus, the use of GW as a substitute for aggregate in construction materials will preserve natural resources and facilitate waste management. In the same way, incorporating GW as a cement substitute will help decrease cement requirements and, consequently, CO$_2$ emissions [31,32].

Numerous experiments have been conducted to assess the characteristics of CM and compressive strength (CS) is interpreted as significant [18]. The CS of the CM provides crucial information on its various properties. The CS of concrete is connected, either directly or indirectly, to a number of mechanical and durability properties [33]. Similarly, flexural strength (FS) is an essential component that must be considered in the design of CM structures, as it affects deflection characteristics, flexural cracking, brittleness ratio, and shear strength [34]. Currently, analytical models for the strength of materials are being developed in an effort to eliminate tests and resources that are unneeded. Various traditional models, including best-fit curves, are used to simulate the material characteristics (developed using regression analysis). Due to the nonlinear nature of CM [35,36], conventional regression approaches may not effectively represent the material’s inherent behavior. In addition, regression methods may exaggerate the significance of some variables [37]. Methods based on artificial intelligence, such as supervised machine learning (ML), are amongst the most highly advanced modeling approaches employed in the current topic area [38–42]. These approaches simulate reactions based on input characteristics, and testing validates the yield models. The use of ML approaches to predict the properties of CM and bituminous blends is gaining popularity [43–45]. The bulk of prior ML-based investigations focused on forecasting the strength of conventional CM [46–48], whereas just a handful focused on predicting the properties of CM that integrated GW. Therefore, further ML-based studies are required to examine the validity of ML techniques on CM that contains WG. Moreover, it is challenging to investigate the impact of several raw ingredients on the material strength in a single study using experimental methods. Thus, it is required to apply advanced methods to examine the impact of various raw ingredients on the material strength in a fast and efficient manner.

This work utilized experimental data to assess the CS and FS of CM with waste glass powder (WGP), used to partially replace cement and fine aggregate, by employing ML methods. Samples of CM were cast using varied quantities of WGP as cement and fine aggregate replacements from 0 to 15%, with an increment of 2.5%. The CS and FS of the samples that contained WGP were recorded at 28 days of age. Two ML methods were employed, i.e., random forest regressor (RFR) and gradient boosting regressor (GBR), to achieve the study’s aims. Both GBR and RFR are ensemble ML methods [49]. Because it is obvious from the literature that ensemble ML approaches outperform individual ML methods [45,50], only ensemble ML methods are selected in this research to determine the best predictor. The performance evaluation of each model was carried out and compared by the coefficient of determination ($R^2$), statistical checks, $k$-fold method, and variation in the predicted results (errors). Experimental-based studies entail extensive efforts, expenditures, and time for material collection, sample casting, curing, and performing tests. By decreasing these efforts using advanced techniques such as ML, the construction sector will benefit as a result. Due to the fact that several raw ingredients, such as cement, fine aggregate, water, silica fume, superplasticizer, and WGP, influence the strength of CM, it is difficult to measure their combined influence using experimental methods. In this instance,
a SHapley Additive ExPlanations (SHAP) analysis was performed to explore the interaction and impact of raw ingredients on the CS and FS of CM. A data sample is needed for ML techniques, which can be generated from the experimental approach. The data generated may consequently be employed to train ML techniques and approximate material properties. The present study used 6 input parameters to foretell the CS and FS of WGP-based CM and assess the efficacy of each ML approach and the impact of raw ingredients on its strength.

2. Research Methods

2.1. Dataset Used for Modeling

In order to obtain the desired results, ML approaches require a broad collection of input parameters [51]. In this regard, an experimental study was carried out using six raw ingredients, including cement, fine aggregate (FA), water, silica fume (SF), superplasticizer (SP), and WGP. Samples of CM were cast with varying amounts of WGP as a cement and fine aggregate replacement, ranging from 0% to 15% with a 2.5% increment. Compressive and flexural strength tests were performed on the samples of 50 mm cubes and 40 mm \( \times \) 40 mm \( \times \) 160 mm prisms, respectively, after 28 days of water curing to assess the CS and FS. In this study, data sample was generated from the experimental work and used to train ML models. All six raw ingredients were taken as inputs, and CS and FS as the outputs for ML-based modeling. Table 1 provides the descriptive statistics for all inputs and outputs utilized for modeling. Mode, median, and mean illustrate the fundamental tendencies, whereas standard deviation, minimum, and maximum highlight variation. The frequency dispersion of each input and output component is provided in Figure 1.

| Parameter       | Cement (kg/m\(^3\)) | FA (kg/m\(^3\)) | Water (kg/m\(^3\)) | SF (kg/m\(^3\)) | SP (kg/m\(^3\)) | WGP (kg/m\(^3\)) | CS (MPa) | FS (MPa) |
|-----------------|---------------------|-----------------|---------------------|-----------------|-----------------|------------------|----------|----------|
| Mean            | 732.51              | 735.62          | 191.33              | 151.67          | 38.17           | 61.65            | 43.48    | 5.00     |
| Standard error  | 4.96                | 5.04            | 0.87                | 2.20            | 0.17            | 3.36             | 0.56     | 0.05     |
| Median          | 722.00              | 729.00          | 191.00              | 153.00          | 38.00           | 60.75            | 41.40    | 4.98     |
| Mode            | 810.00              | 810.00          | 203.00              | 122.00          | 40.50           | 40.00            | 40.20    | 4.90     |
| Standard deviation | 53.62            | 54.51            | 9.43                | 23.80           | 1.85            | 36.34           | 6.09     | 0.56     |
| Minimum         | 612                 | 612              | 180                 | 122             | 36              | 0                | 34.24    | 3.55     |
| Maximum         | 810                 | 810              | 203                 | 180             | 40.5            | 121.5           | 62.46    | 6.21     |

2.2. Machine Learning-Based Modeling

Using the experimental data, the CS of CM that contained WGP was evaluated. The procedures utilized cement, fine aggregate, water, silica fume, superplasticizer, and WGP as inputs, with CS and FS serving as the outputs. Ensemble ML approaches with Python code, and Anaconda Navigator software were utilized to achieve the aims of the research. The ML models were operated using Spyder (version 5.1.5). GBR and RFR ML methods were used to assess the CS and FS of WGP-based CM. These ML algorithms are generally used to estimate required results using input variables. These approaches may be used to forecast a material’s strength, temperature resistance, and durability [52,53]. During the modeling phase, six input characteristics and two outputs (CS and FS) were used. The proportion of experimental data used for modeling was 30% for testing and 70% for validation. The \( R^2 \) value of the expected outcome indicates the exactness of a model. The \( R^2 \) value reveals the amount of deviation; a number close to zero indicates greater variation, whereas a number close to one suggests that the prediction model and experimental results are almost entirely matched [54]. On both models, k-fold, statistical, and error assessments, including mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE), were performed. Figure 2 demonstrates the sequence of the modeling techniques. The succeeding subsections describe the ML methods and validation approaches utilized in this work.
2.2. Machine Learning-Based Modeling

Using the experimental data, the CS of CM that contained WGP was evaluated. The procedures utilized cement, fine aggregate, water, silica fume, superplasticizer, and WGP as inputs, with CS and FS serving as the outputs. Ensemble ML approaches with Python code, and Anaconda Navigator software were utilized to achieve the aims of the research. The ML models were operated using Spyder (version 5.1.5). GBR and RFR ML methods were used to assess the CS and FS of WGP-based CM. These ML algorithms are generally used to estimate required results using input variables. These approaches may be used to forecast a material’s strength, temperature resistance, and durability [52,53]. During the modeling phase, six input characteristics and two outputs (CS and FS) were used. The proportion of experimental data used for modeling was 30% for testing and 70% for validation. The R² value of the expected outcome indicates the exactness of a model. The R² value reveals the amount of deviation; a number close to zero indicates greater variation, whereas a number close to one suggests that the prediction model and experimental results are almost entirely matched [54]. On both models, k-fold, statistical, and error assessments, including mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE), were performed. Figure 2 demonstrates the sequence of the modeling techniques. The succeeding subsections describe the ML methods and validation approaches utilized in this work.

Figure 1. Relative frequency dispersal of input and output parameters: (a) Cement; (b) Fine aggregate; (c) Water; (d) Silica fume; (e) Superplasticizer; (f) Waste glass powder; (g) Compressive strength; (h) Flexural strength.
Spyder (version: 5.1.5) was used from Anaconda navigator.

Dataset was developed using experimental methods. Cube and prism samples were casted to for testing after 28-days of water curing. Inputs: cement, fine aggregate, water, silica fume, superplasticizer, and waste glass powder. Outputs: Compressive and flexural strengths.

Compressive and flexural strength prediction using gradient boosting and random forest techniques. Spyder (version: 5.1.5) was used from Anaconda navigator.

Coefficient of determination ($R^2$) and variation between experimental and predicted results (errors). Statistical checks (MAE, MAPE, and RMSE) and k-fold assessment.

Figure 2. Sequence of dataset development, modeling, and validation methods.

2.2.1. Gradient Boosting Regressor

Friedman [55] suggested this ensemble approach for classification and regression. GBR is similar to other boosting approaches, but is restricted to regression alone. As observed in Figure 3, each training set repetition is picked at random and verified by the base model in this approach. GBR's accuracy and speed may be improved by randomly subsampling the training data, which eventually prevents overfitting. The lower the amount of training data samples, the greater the rate of regression to fit. GBR requires the shrinkage rate and n-trees tuning variables, where n-trees represents the figure of trees generated. Here, the number of n trees should not be too low, and the shrinkage factor, also known as the learning rate, applies to each expansion tree.

Figure 3. Sequence of gradient boosting regressor training procedure [56].

2.2.2. Random Forest Regressor

RFR is accomplished by means of random split selection on bagged decision trees [57]. The construction and procedure of the RFR model are depicted schematically in Figure 4. Each tree in the forest is generated by utilizing an arbitrarily selected training set, and each split within a tree is constructed by utilizing an arbitrarily chosen subgroup of input
parameters, developing a forest of trees [38]. This element of uncertainty increases the tree’s variety. The forest is comprised solely of mature binary trees. The RFR approach has shown to be a highly successful regression tool for general applications. The approach, which accumulates the calculations of many randomized decision trees, has demonstrated greater precision when the number of variables surpasses the number of interpretations. In addition, it is adaptable to both large-scale and ad hoc learning tasks, delivering metrics of varying significance [57].

Figure 4. Structure of random forest-based modeling [59].

2.2.3. Validation of Machine Learning-Based Models

Statistical checks and k-fold techniques were adopted to confirm the deployed ML algorithms. Typically, the k-fold method is employed to evaluate the effectiveness of a model by randomly splitting relevant data samples into 10 classes [60]. As shown in Figure 5, nine classes are used to train ML models, whereas only one is used for validation. The ML approach is more accurate when errors are smaller and $R^2$ is larger. In addition, the desired result requires 10 repeats of the technique. This effort greatly adds to the model’s outstanding accuracy. In addition, each ML technique’s precision was statistically evaluated using error evaluation (MAE, MEPE, and RMSE). The accuracy of the ML methods’ projections was statistically evaluated using Equations (1)–(3) derived from past research [61,62].

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |P_i - T_i|, \quad (1)
\]

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - T_i)^2}, \quad (2)
\]

\[
\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|P_i - T_i|}{T_i}, \quad (3)
\]

where $n$ = size of the dataset, $P_i$ = estimated results, and $T_i$ = experimental results.
2.3. SHAP Analysis

This research also determined global feature impacts and analyzed feature relations with CM, using a game theory approach known as SHAP [64]. SHAP analysis enhances the expandability of the suggested model. In this method, each case prediction is proved by computing all impact-considered characteristics, using SHapley values derived from the coalition game theory. Each feature’s impact on the SHapley value is somewhat averaged over all feasible combinations. The values of SHAP are directly proportional to the impact of characteristics. The mean of each input’s SHAP value is used to determine the global effect of each feature. These values are then arranged in order of descending significance, followed by the charting of SHAP values. On the SHAP plot, the SHAP value for each raw ingredient is represented by a single point. The X and Y axes show SHAP values and the significance of a feature, respectively. On the Y axis, its higher placement indicates the greater effect of the characteristic on the output, and a color gradient from light to dark is utilized to illustrate its significance. The interaction between characteristics and their effect on the outcomes is illustrated by SHAP plots with a color scheme that indicates feature interaction. This strategy provides greater information than conventional partial dependency graphs [65]. The $\phi^i(f)$ is the allotted impact for an ingredient impact, summed for the model’s outcome $f(x_i)$ to obtain the likely feature patterns [66]. The $\phi^i(f)$ is determined by Equation (4) as follows:

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

(4)

where

- $S$ = the ingredient subset,
- $x_i$ = ingredient $j$,
- $p$ = the ingredient number in the model.

In this technique, the significance of a feature is determined by measuring and estimating errors, while using a fixed feature value. Consideration is given to the estimated...
error sensitivity to allocate weight to the ingredient importance, while affecting its value. Moreover, SHAP illustrates the performance of the trained ML model. SHAP considers a different feature designation strategy, namely the linear addition of inputs, to provide an explainable model based on the model’s conclusion. For instance, a model with input parameters \( x_i \), where \( i \) is in the range from 1 to \( k \), and \( k \) represents the quantity of input parameters, and where \( h(x_k) \) represents a description model with \( x_k \) as a simple input, Equation (5) is used to illustrate an original model \( f(x) \).

\[
f(x) = h(x) = \varnothing_0 + \sum_{i=1}^{p} \varnothing_i x_i
\]  

(5)

where

\( p \) = the input feature number

\( \varnothing_0 \) = the constant.

In addition, \( x = m_i(x_i) \), i.e., the mapping function is interlinked with both \( x \) and \( x_i \) input factors. Lundberg and Lee [67] provided Equation (5), where \( h(x) \), i.e., the estimation value, was increased by \( \varnothing_0 \), \( \varnothing_1 \), and \( \varnothing_3 \) terms and a drop of \( \varnothing_4 \) in \( h(x_i) \) was also identified (see Figure 6). A solution with a single value to Equation (5) incorporates three advantageous features, namely local accuracy, reliability, and missingness. Reliability verifies that no attribute decrease was assigned to the appropriate feature in a modification to a more influential feature. For missingness, it is established that missing features have no significant value; therefore, \( \varnothing_i = 0 \) is applied by \( x_i^s = 0 \). For local accuracy, it is established that the sum-up for attributing features as an output function includes a model that needs to match output \( f \) to \( x_k \) as the simplified input. \( x = m_i x_i \) signifies the achievement of local precision.

![Simple Case of SHAP Attributes](image)

**Figure 6.** Attributes of SHAP [68].

3. Results and Analysis

3.1. Compressive Strength Models

3.1.1. Gradient Boosting Regressor Model

Figure 7 illustrates the results of the GBR model for estimating the WGP-based CM’s CS. Figure 7a depicts the association among experimental and projected CS. The GBR model forecasted CS with a high level of precision and little variance between the experimental and predicted results. The \( R^2 \) value of 0.93 indicates that the GBR strategy for estimating the CS of WGP-based CM is exact, and the experimental and predicted results are in excellent agreement. Figure 7b depicts the experimental, predicted, and divergent value (errors) dispersion for the GBR method. The error value distribution ranged from 0.01 MPa to 5.0 MPa, with a mean of 1.25 MPa. In addition, the proportionate dispersion of errors was examined, and it was determined that 47.2% of the values were lower than 1 MPa, 30.6% fell among 1–2 MPa, and 22.2% were higher than 2 MPa. The division of divergent data (errors) indicates that the GBR approach predicted the CS of WGP-based CM accurately.
3.1.2. Random Forest Regressor Model

Figure 8 displays the results of the RFR approach used to forecast the CS of the WGP-based CM. The correlation among experimental and estimated CS can be observed in Figure 8a. Compared to the GBR model utilized in the current research, the RFR method produced more accurate results and the smallest discrepancy between the experimental and anticipated findings. The $R^2$ value of 0.94 for the RFR model is indicative of its greater accuracy. Figure 8b illustrates the distribution of experimental, estimated, and divergent values (error) using the RFR method. The least, average, and maximum errors...
were 0.16 MPa, 1.10 MPa, and 2.81 MPa, respectively. Analyzing the error value distribution revealed that 55.6% fell below 1 MPa, 30.6% fell between 1 and 2 MPa, and 13.9% exceeded 2 MPa. Consequently, the error distribution suggested that the RFR model is more exact than the GBR model. The RFR model is more accurate because, in the RFR training process, each tree produces regression, and the forest with the highest number of votes is chosen as the model.

Figure 8 displays the results of the RFR approach used to forecast the CS of the WGP-based CM. The correlation among experimental and estimated CS can be observed in Figure 8a. Compared to the GBR model utilized in the current research, the RFR method produced more accurate results and the smallest discrepancy between the experimental and anticipated findings. The R² value of 0.94 for the RFR model is indicative of its greater accuracy. Figure 8b illustrates the distribution of experimental, estimated, and divergent values (error) using the RFR method. The least, average, and maximum errors were 0.16 MPa, 1.10 MPa, and 2.81 MPa, respectively. Analyzing the error value distribution revealed that 55.6% fell below 1 MPa, 30.6% fell between 1 and 2 MPa, and 13.9% exceeded 2 MPa. Consequently, the error distribution suggested that the RFR model is more exact than the GBR model. The RFR model is more accurate because, in the RFR training process, each tree produces regression, and the forest with the highest number of votes is chosen as the model.
3.2. Flexural Strength Models

3.2.1. Gradient Boosting Regressor Model

Figure 9 depicts the findings of the GBR approach for determining the FS of the WGP-based CM. The link between experimental data and expected results is depicted in Figure 9a. The GBR approach predicted FS with a satisfactory degree of precision and less variation among experimental and estimated FS. The $R^2$ value of 0.90 suggests that the GBR technique for predicting the FS of WGP-based CM is satisfactory, with good agreement between the experimental and anticipated findings. Figure 9b displays the dispersion of experimental, anticipated, and divergent values (errors) using the GBR approach. The distribution of error values varied from 0.01 MPa to 0.38 MPa, with a mean of 0.12 MPa. Furthermore, the proportional dispersion of error values was investigated, and it was found that 50% of the errors were lower than 0.1 MPa, 30.6% were among 0.1–0.2 MPa, and 19.4% were larger than 0.2 MPa. The distribution of divergent data (errors) suggests that the GBR technique effectively anticipated the FS of WGP-based CM.

![Figure 9. GBR-FS model: (a) correlation amongst experimental and predicted FS; (b) distribution of experimental and predicted FS and error values.](image-url)
3.2.2. Random Forest Regressor Model

The findings of the RFR technique used to anticipate the FS of the WGP-based CM are shown in Figure 10. Figure 10a represents the link among experimental and predicted FS. The RFR technique yielded more accurate results and the least disparity between experimental and predicted outcomes. The RFR model’s $R^2$ of 0.91 indicates its superior accuracy. Figure 10b depicts the RFR method’s distribution of experimental, estimated, and divergent values (errors). The average and maximum errors were determined to be 0.10 MPa and 0.37 MPa, respectively. The error value distribution indicated that 61.1% were less than 0.1 MPa, 25.0% were between 0.1 and 0.2 MPa, and 13.9% were greater than 0.2 MPa. As a result of the error distribution, the RFR model was shown to be more exact than the GBR model. It can be concluded that the RFR method is more accurate in predicting the CS and FS of WGP-based CM. However, the accuracy of the GBR model is also at an acceptable level. Hence, both models can be employed to assess the strength of CM incorporated with WGP.

Figure 10. RFR-FS model: (a) correlation amongst experimental and predicted FS; (b) distribution of experimental and predicted FS and error values.
3.3. Validation of Machine Learning Models

The results of the error evaluations (MAE, MAPE, and RMSE) using Equations (1)–(3) above for both CS and FS estimation models are shown in Table 2. For the CS prediction, it was found that the MAE values for GBR and RFR were 1.254 MPa and 1.095 MPa, respectively. MAPE values for GBR and RFR were determined to be 2.90% and 2.60%, respectively. In addition, RMSE values for GBR and RFR were calculated to be 1.597 MPa and 1.331 MPa, respectively. These assessments also indicated that the RFR approach is more precise than the GBR due to its lower error rate. Similarly, for the FS prediction, the MAE, MAPE, and RMSE values for the GBR model were 0.124 MPa, 2.50%, and 0.152 MPa, respectively, whereas MAE, MAPE, and RMSE values for the RFR model were 0.104 MPa, 2.10%, and 0.137 MPa, respectively. These errors also validated the higher precision of the RFR model in estimating the FS of CM incorporated with WGP. The results of computing \( R^2 \), RMSE, and MAE to validate the k-fold method are provided in Table 3. Figures 11 and 12 were created for the comparison of both ML methods’ k-fold analyses for the CS and FS prediction, respectively. For the CS estimation, the range of MAE values for the GBR technique was 1.25 MPa to 5.09 MPa, with a mean of 2.42 MPa. The RFR model’s MAE ranged from 1.10 MPa to 4.22 MPa, with a mean of 2.23 MPa. In the same way, the average RMSE values for the GBR and RFR techniques were 3.01 MPa and 2.46 MPa, respectively. However, the average \( R^2 \) values for GBR and RFR were 0.72 and 0.75, respectively. Comparatively, the RFR model with the lowest errors and the greater \( R^2 \) is the more accurate in predicting the CS of WGP-based CM. Similar findings were noticed for the FS prediction models. The average MAE, RMSE, and \( R^2 \) values for the GBR model were 0.30 MPa, 0.39 MPa, and 0.68, respectively, whereas for the RFR model, the average MAE, RMSE, and \( R^2 \) values were 0.26 MPa, 0.31 MPa, and 0.70, respectively. The assessment of these errors and \( R^2 \) values from the k-fold approach also confirmed the higher accuracy of the RFR model. However, the precision of the GBR model is also satisfactory. Hence, both GBR and RFR models might be employed to assess the CS and FS of CM with higher accuracy.

Table 2. Error evaluations of the developed models using statistical checks.

| ML Method | CS Model | FS Model |
|-----------|----------|----------|
|           | MAE (MPa) | MAPE (%) | RMSE (MPa) | MAE (MPa) | MAPE (%) | RMSE (MPa) |
| GBR       | 1.254     | 2.90     | 1.597      | 0.124     | 2.50     | 0.152      |
| RFR       | 1.095     | 2.60     | 1.331      | 0.104     | 2.10     | 0.137      |

Table 3. Results of k-fold assessment.

| K-Fold | CS         | FS         |
|--------|------------|------------|
|        | GBR        | RFR        | GBR        | RFR        |
|        | MAE | RMSE | \( R^2 \) | MAE | RMSE | \( R^2 \) | MAE | RMSE | \( R^2 \) |
| 1      | 1.60 | 2.48 | 0.53 | 2.04 | 2.61 | 0.40 | 0.23 | 0.35 | 0.89 | 0.13 | 0.29 | 0.88 |
| 2      | 1.62 | 4.06 | 0.86 | 1.49 | 1.46 | 0.94 | 0.32 | 0.30 | 0.73 | 0.26 | 0.26 | 0.91 |
| 3      | 1.52 | 1.60 | 0.65 | 1.90 | 1.87 | 0.70 | 0.27 | 0.32 | 0.83 | 0.29 | 0.31 | 0.61 |
| 4      | 2.99 | 4.01 | 0.82 | 3.51 | 4.63 | 0.77 | 0.12 | 0.38 | 0.85 | 0.10 | 0.27 | 0.63 |
| 5      | 4.41 | 5.48 | 0.93 | 3.04 | 1.33 | 0.53 | 0.42 | 0.64 | 0.66 | 0.39 | 0.43 | 0.75 |
| 6      | 1.61 | 1.77 | 0.85 | 1.23 | 2.08 | 0.89 | 0.26 | 0.34 | 0.45 | 0.21 | 0.26 | 0.68 |
| 7      | 5.09 | 2.70 | 0.57 | 4.22 | 2.31 | 0.89 | 0.43 | 0.48 | 0.57 | 0.35 | 0.24 | 0.81 |
| 8      | 1.25 | 2.17 | 0.33 | 1.13 | 1.82 | 0.78 | 0.23 | 0.39 | 0.31 | 0.25 | 0.31 | 0.59 |
| 9      | 2.50 | 3.97 | 0.81 | 2.64 | 3.03 | 0.91 | 0.48 | 0.54 | 0.64 | 0.40 | 0.44 | 0.62 |
| 10     | 1.63 | 1.82 | 0.83 | 1.10 | 3.44 | 0.71 | 0.24 | 0.15 | 0.86 | 0.24 | 0.29 | 0.54 |
GBR and RFR models might be employed to assess the CS and FS of CM with higher accuracy.

Table 2. Error evaluations of the developed models using statistical checks.

| ML Method | CS Model | FS Model |
|-----------|----------|----------|
|            | MAE (MPa) | MAPE (%) | RMSE (MPa) | MAE (MPa) | MAPE (%) | RMSE (MPa) |
| GBR       | 1.254     | 2.90     | 1.597      | 0.124     | 2.50      | 0.152      |
| RFR       | 1.095     | 2.60     | 1.331      | 0.104     | 2.10      | 0.137      |

Figure 11. Results of k-fold assessment, indicating MAE, RMSE, and $R^2$ for GBR and RFR models for the CS estimation.

Table 3. Results of k-fold assessment.

| K-fold | CS | FS |
|--------|----|----|
|        | GBR MAE | RMSE | $R^2$ | RFR MAE | RMSE | $R^2$ |
| 1      | 1.60 | 2.48 | 0.53 | 2.04 | 2.61 | 0.40 |
| 2      | 1.62 | 4.06 | 0.86 | 1.49 | 1.46 | 0.94 |
| 3      | 1.52 | 1.60 | 0.65 | 1.90 | 1.87 | 0.70 |
| 4      | 2.99 | 4.01 | 0.82 | 3.51 | 4.63 | 0.77 |
| 5      | 4.41 | 5.48 | 0.93 | 3.04 | 1.33 | 0.53 |
| 6      | 1.61 | 1.77 | 0.85 | 1.23 | 2.08 | 0.89 |
| 7      | 5.09 | 2.70 | 0.57 | 4.22 | 2.31 | 0.89 |
| 8      | 1.25 | 2.17 | 0.33 | 1.13 | 1.82 | 0.78 |
| 9      | 2.50 | 3.97 | 0.81 | 2.64 | 3.03 | 0.91 |
| 10     | 1.63 | 1.82 | 0.83 | 1.10 | 3.44 | 0.71 |

Figure 12. Results of k-fold assessment, indicating MAE, RMSE, and $R^2$ for GBR and RFR models for the FS estimation.

3.4. Impact of Raw Ingredients

In this research, the effect of raw materials on the CS and FS performance of CM was investigated. The SHAP tree explainer is employed in the entire dataset to provide a more detailed explanation of the global feature effects by including local SHAP explanations.

Figure 13 illustrates the results of the violin SHAP plot for all raw materials regarding the CS and FS of WGP-based CM. In this graph, each parameter value is characterized by a separate hue, and the resultant SHAP value on the x axis shows a raw ingredient’s contribution. FA is a raw material that has a greater impact on strength, as observed by the larger negative association between this feature and the CS of WGP-based CM (more red dots on the negative side). This implies that a rise in FA amount would likely lead to a decline in strength. It was established that the influence of WGP is more favorable (more red dots on the positive side), indicating that as WGP content increases, the material strength improves. However, a negative effect has also been found, suggesting that utilizing WGP in excess of the optimal quantity may reduce the strength. The effect of cement on the CS and FS was likewise shown to be more favorable, indicating that a larger cement fraction raises the strength. On the other hand, the influence of water on the CS and FS was shown to be more negative, suggesting that the water amount must be maintained low in order to attain a greater material strength. Due to the lack of input value variation in the employed dataset, it was concluded that the effect of SF and SP was unclear. Using a larger dataset with a broader variety of input characteristics may result in improved associations.
3.4. Impact of Raw Ingredients

In this research, the effect of raw materials on the CS and FS performance of CM was investigated. The SHAP tree explainer is employed in the entire dataset to provide a more detailed explanation of the global feature effects by including local SHAP explanations. Figure 13 illustrates the results of the violin SHAP plot for all raw materials regarding the CS and FS of WGP-based CM. In this graph, each parameter value is characterized by a separate hue, and the resultant SHAP value on the x axis shows a raw ingredient’s contribution. FA is a raw material that has a greater impact on strength, as observed by the larger negative association between this feature and the CS of WGP-based CM (more red dots on the negative side). This implies that a rise in FA amount would likely lead to a decline in strength. It was established that the influence of WGP is more favorable (more red dots on the positive side), indicating that as WGP content increases, the material strength improves. However, a negative effect has also been found, suggesting that utilizing WGP in excess of the optimal quantity may reduce the strength. The effect of cement on the CS and FS was likewise shown to be more favorable, indicating that a larger cement fraction raises the strength. On the other hand, the influence of water on the CS and FS was shown to be more negative, suggesting that the water amount must be maintained low in order to attain a greater material strength. Due to the lack of input value variation in the employed dataset, it was concluded that the effect of SF and SP was unclear. Using a larger dataset with a broader variety of input characteristics may result in improved associations.

![Figure 13. SHAP plot, indicating the impact of raw ingredients on CS and FS.](image-url)

Figure 14 depicts the correlations between raw ingredients and their influence on the CS and FS of WGP-based CM. Figure 14a depicts the interaction of cement. The graph illustrates that as the cement amount grows, the material strength increases and mostly interacts with the FA. In contrast, increased FA levels have a detrimental effect on CS and FS (Figure 14b) and interact mostly with WGP. In addition, as shown in Figure 14c, water mostly interacts with WGP, and raising its value has a detrimental effect on its strength. Therefore, the water content should be reduced to increase its strength. Figure 14d,e demonstrate that the influence of SF and SP was uncertain based on the employed dataset, due to the lower variation in quantities for SF and SP. The incorporation of WGP into CM was shown to be advantageous (see Figure 14f). Among the possible causes are the filler effect and the pozzolanic property of WGP. The filler effect decreases porosity, resulting in...
a compact and dense matrix. The larger concentration of SiO$_2$ in the chemical composition of glass reacts with Ca(OH)$_2$ in the matrix to create a thick C-S-H gel, hence improving the performance of the material [69,70]. Utilizing WGP up to the optimum amount will assist in enhancing the strength of CM. Therefore, WGP may be utilized in the range of 80 to 120 kg/m$^3$ to increase material strength. Moreover, WGP interacts primarily with the FA, among other input characteristics. This indicates that the use of WGP as a substitute for FA may result in greater strength compared to its usage as a cement replacement. It is crucial to note that these results are dependent on the kinds of raw materials and the size of the dataset studied in this research. Utilizing varied input settings and data samples could result in distinct outputs.

4. Discussions
Globally, a substantial quantity of GW is produced, most of which is discarded in landfills, posing health and environmental risks [23]. Furthermore, CM is the most common construction material, but its overuse depletes natural resources and generates CO$_2$. GW has the possibility to be used as a partial substitute for fine aggregate and cement in CM, which is an environmentally favorable technique. Therefore, the use of GW in CM will decrease environmental impacts through the elimination of waste, preservation of natural resources, and reduction in CO$_2$ emissions. Using ML-based modeling and SHAP analysis techniques, this research aimed to expand knowledge about the application of WGP in
CM. This work estimated the CS and FS of WGP-based CM using GBR and RFR ML techniques. The accuracy of each technique was evaluated to ascertain which is the most exact predictor. Compared to the GBR method, with an $R^2$ of 0.93 and 0.89 for CS and FS estimation, respectively, the RFR method produced more accurate results, with an $R^2$ of 0.94 and 0.91 for CS and FS prediction, respectively. The disparity between the experimental and anticipated outcomes (errors) further substantiated the superior accuracy of the RFR approach. Compared to the GBR models, the experimental and estimated results for the RFR models showed good agreement, as demonstrated by the error analysis. In predicting the strength of CM, previous research has similarly demonstrated that the RFR method is more exact than the GBR method [61,71,72].

Moreover, the accuracy of both models was examined using statistical and k-fold approaches. When the degree of divergence (MAE, MAPE, and RMSE) is low and $R^2$ is high, a model is more exact. Nevertheless, defining and recommending the optimal ML method for predicting attributes in various study areas is difficult, since the accuracy of an ML methodology is primarily reliant on the size of inputs and data samples used to execute algorithms [61]. Ensemble ML methods repeatedly employ the weak learner by developing sub-models that are trained on the data sample and tweaked to enhance the $R^2$ value, thus resulting in more accurate outputs than the individual ML models. Figure 15 shows the dispersal of $R^2$ for the GBR and RFR sub-models. The $R^2$ for GBR-CS sub-models ranged from 0.876 to 0.932, with a mean of 0.903. In addition, the $R^2$ for the RFR-CS sub-models varied between 0.927 and 0.944, with an average of 0.938. Likewise, the average $R^2$ values for the GBR-FS and RFR-FS sub-models were 0.866 and 0.905, respectively. These findings reveal that the RFR sub-models are more precise than the GBR sub-models. Additionally, a SHAP analysis is conducted to investigate the interaction and influence of raw materials on the CS and FS of WGP-based CM. FA was demonstrated to be a very effective raw material, exhibiting a stronger negative association with the material’s strength. However, WGP was found to have a greater positive impact on the strength of CM. While WGP has been demonstrated to have beneficial effects, there is evidence that exceeding the optimal dosage may have negative consequences on performance. In addition, the impact of cement on CS and FS was shown to be more beneficial, indicating that strength increases with increasing cement content. However, due to the lack of variation in SF and SP in the data sample, their effect was ambiguous, and larger datasets with more input features may result in better associations.

![Figure 15. Distribution of $R^2$ for the utilized ML models.](image-url)
5. Conclusions

This research aimed to employ experimental data to develop machine learning (ML)-based models to evaluate the compressive strength (CS) and flexural strength (FS) of cement mortar (CM) that contained waste glass powder (WGP). Two types of ensemble ML approaches, including gradient boosting regressor (GBR) and random forest regressor (RFR), were used to forecast the CS and FS. Moreover, SHapley Additive ExPlanations (SHAP) analysis was carried out to study the impact of raw ingredients on the strength of CM. This research reached the following conclusions:

- It was determined from the modeling methods that the GBR models had a satisfactory degree of precision, with an $R^2$ of 0.93 and 0.89 for CS and FS prediction, respectively, while the RFR models had a higher degree of precision, with an $R^2$ of 0.94 and 0.91 for CS and FS prediction, respectively.
- The average variation between predicted and experimental CS (error) in GBR and RFR models was determined to be 1.25 MPa and 1.10 MPa, respectively. Similarly, the average error values in predicting the FS of CM in the GBR and RFR models were 0.12 MPa and 0.10 MPa, respectively. These errors also confirmed the acceptable precision of the RFR models and higher accuracy of the RFR models in forecasting the strength of WGP-based CM.
- The SHAP study revealed that fine aggregate (FA) was a crucial raw material, with a higher negative correlation to the material’s strength. WGP and cement had a stronger favorable impact on CM’s strength. Due to the deficiency of variance in silica fume (SF) and superplasticizer (SP) in the data set, the effect of SF and SP was unclear.
- New techniques, such as ML-based modeling and SHAP analysis, will aid the building industry by fostering the advancement of fast and economical ways of determining material properties and the impact of raw ingredients.

This study employed data for which experimental work was performed in a controlled environment (laboratory). It is suggested that in future studies, actual on-site conditions, such as humidity, temperature, curing, etc., need to be incorporated during the modeling phase to examine their impact on the material’s strength.

**Author Contributions:** H.A.A.: data curation, methodology, resources, funding acquisition and visualization. M.N.A.: conceptualization, resources, funding acquisition, supervision, project administration, writing, reviewing, and editing. W.A.: conceptualization, software, methodology, validation, investigation, supervision, visualization, writing original draft, reviewing, and editing. K.K.: methodology, resources, writing, reviewing, and editing. S.N.: data curation, visualization, writing, reviewing, and editing. M.I.F.: visualization and validation. M.I.: visualization and validation. 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. GRANT1653), through its KFU Research Summer Initiative. The APC was also funded by Project No. GRANT1653.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

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

**Acknowledgments:** The authors acknowledge the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Project No. GRANT1653), through its KFU Research Summer Initiative. The authors also acknowledge the financial support that made this study possible.

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

1. Vigil de la Villa Mencía, R.; Frias, M.; Ramírez, S.M.; Carrasco, L.F.; Giménez, R.G. Concrete/Glass Construction and Demolition Waste (CDW) Synergies in Ternary Eco-Cement-Paste Mineralogy. *Materials* 2022, 15, 4661. [CrossRef] [PubMed]

2. Kang, S.; Zhao, Y.; Wang, W.; Zhang, T.; Chen, T.; Yi, H.; Rao, F.; Song, S. Removal of methylene blue from water with montmorillonite nanosheets/chitosan hydrogels as adsorbent. *Appl. Surf. Sci.* 2018, 448, 203–211. [CrossRef]

3. Wang, W.; Zhao, Y.; Bai, H.; Zhang, T.; Ibarra-Galvan, V.; Song, S. Methylene blue removal from water using the hydrogel beads of poly (vinyl alcohol)-sodium alginate-chitosan-montmorillonite. *Carbohydr. Polym.* 2018, 198, 518–528. [CrossRef] [PubMed]

4. Ahmad, W.; Ahmad, A.; Ostrowski, K.A.; Aslam, F.; Joyklad, P. A scientometric review of waste material utilization in concrete for sustainable construction. *Case Stud. Constr. Mater.* 2021, 15, e00683. [CrossRef]

5. Li, X.; Qin, D.; Hu, Y.; Ahmad, W.; Ahmad, A.; Aslam, F.; Joyklad, P. A systematic review of waste materials in cement-based composites for construction applications. *J. Build. Eng.* 2021, 45, 103447. [CrossRef]

6. Alyousef, R.; Ahmad, W.; Ahmad, A.; Aslam, F.; Joyklad, P.; Alabduljabbar, H. Potential use of recycled plastic and rubber aggregate in cementitious materials for sustainable construction: A review. *J. Clean. Prod.* 2021, 329, 129736. [CrossRef]

7. Khan, M.; Cao, M.; Xie, C.; Ali, M. Effectiveness of hybrid steel-basalt fiber reinforced concrete under compression. *Eng. Sci. Technol. Int. J.* 2017, 20, 151–159. [CrossRef]

8. Smarzewski, P.; Barnat-Hunek, D. Mechanical and durability related properties of high performance concrete made with coal cinder and waste foundry sand. *Constr. Build. Mater.* 2016, 121, 9–17. [CrossRef]

9. Khan, M.; Cao, M.; Hussain, A.; Chu, S.H. Effect of silica-fume content on performance of CaCO3 whisker and basalt fiber at matrix interface in cement-based composites. *Case Stud. Constr. Mater.* 2021, 300, 124046. [CrossRef]

10. Xie, C.; Cao, M.; Khan, M.; Yin, H.; Guan, J. Review on different testing methods and factors affecting fracture properties of fiber reinforced cementitious composites. *Constr. Build. Mater.* 2020, 273, 121766. [CrossRef]

11. Asutkar, P.; Shinde, S.; Patel, R. Study on the behaviour of rubber aggregates concrete beams using analytical approach. *Eng. Sci. Technol. Int. J.* 2017, 20, 151–159. [CrossRef]

12. Martínez-Barrera, G.; del Coz-Díaz, J.J.; Álvarez-Rabanal, F.P.; Gayarre, F.L.; Martínez-López, M.; Cruz-Olivares, J. Waste tire rubber particles modified by gamma radiation and their use as modifiers of concrete. *Case Stud. Constr. Mater.* 2020, 12, e00321. [CrossRef]

13. Amin, M.N.; Ahmad, W.; Khan, K.; Sayed, M.M. Mapping Research Knowledge on Rice Husk Ash Application in Concrete: A Scientometric Review. *Materials* 2022, 15, 4343. [CrossRef] [PubMed]

14. He, Z.-H.; Han, X.-D.; Zhang, M.-Y.; Yuan, Q.; Shi, J.-Y.; Zhan, P.-M. A novel development of green UHPC containing waste concrete powder derived from construction and demolition waste. *Powder Technol.* 2022, 398, 117075. [CrossRef]

15. Jain, K.L.; Sancheti, G.; Gupta, L.K. Durability performance of waste granite and glass powder added concrete. *Constr. Build. Mater.* 2020, 252, 119075. [CrossRef]

16. Martínez-Barrera, G.; del Coz-Díaz, J.J.; Álvarez-Rabanal, F.P.; Gayarre, F.L.; Martínez-López, M.; Cruz-Olivares, J. Waste tire rubber particles modified by gamma radiation and their use as modifiers of concrete. *Case Stud. Constr. Mater.* 2020, 12, e00321. [CrossRef]

17. Qin, D.; Hu, Y.; Li, X. Waste Glass Utilization in Cement-Based Materials for Sustainable Construction: A Review. *Crystals* 2021, 11, 710. [CrossRef]

18. Ahmad, W.; Ahmad, A.; Ostrowski, K.A.; Aslam, F.; Joyklad, P.; Zajdel, P. Sustainable approach of using sugarcane bagasse as cement-based composites: A systematic review. *Case Stud. Constr. Mater.* 2021, 15, e00698. [CrossRef]

19. Meena, A.; Singh, R. *Comparative Study of Waste Glass Powder as Pozzolanic Material in Concrete*; National Institute of Technology: Rourkela, India, 2012.

20. Pant, D.; Singh, P. Pollution due to hazardous glass waste. *Environ. Sci. Pollut. Res.* 2014, 21, 2414–2436. [CrossRef] [PubMed]

21. Singh, N.; Li, J.; Zeng, X. Solutions and challenges in recycling waste cathode-ray tubes. *J. Clean. Prod.* 2016, 133, 188–200. [CrossRef]

22. Federico, L.; Chidiae, S. Waste glass as a supplementary cementitious material in concrete-critical review of treatment methods. *Cem. Concr. Compos.* 2009, 31, 606–610. [CrossRef]

23. Mohajerani, A.; Vajna, J.; Cheung, T.H.H.; Kurmus, H.; Arulrajah, A.; Horpibulsuk, S. Practical recycling applications of crushed waste glass in construction materials: A review. *Constr. Build. Mater.* 2017, 156, 443–467. [CrossRef]
30. Paul, S.C.; Šavija, B.; Babafemi, A.J. A comprehensive review on mechanical and durability properties of cement-based materials containing waste recycled glass. J. Clean. Prod. 2018, 198, 891–906. [CrossRef]
31. Sudharsan, N.; Palanisamy, T.; Yaragal, S.C. Environmental sustainability of waste glass as a valuable construction material—A critical review. Ecol. Environ. Conserv. 2018, 24, 5331–5338.
32. Jiang, Y.; Ling, T.-C.; Mo, K.H.; Shi, C. A critical review of waste glass powder—Multiple roles of utilization in cement-based materials and construction products. J. Environ. Manag. 2019, 242, 440–449. [CrossRef] [PubMed]
33. Singh, N.; Kumar, P.; Goyal, P. Reviewing the behaviour of high volume fly ash based self compacting concrete. J. Build. Eng. 2019, 26, 100882. [CrossRef]
34. Meisuh, B.K.; Kankam, C.K.; Buabin, T.K. Effect of quarry rock dust on the flexural strength of concrete. Constr. Build. Mater. 2018, 8, 16–22. [CrossRef]
35. Awoyera, P.O. Nonlinear finite element analysis of steel fibre-reinforced concrete beam under static loading. J. Eng. Sci. Technol. 2016, 11, 1669–1677.
36. 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]
37. Sadrmomtazi, A.; Sobhani, J.; Mirgozar, M.A. Modeling compressive strength of EPS lightweight concrete using regression, neural network and ANFIS. Constr. Build. Mater. 2013, 42, 205–216. [CrossRef]
38. Ilyas, I.; Zafar, A.; Afzal, M.T.; Javed, M.F.; Alrowais, R.; Althoey, F.; Mohamed, A.M.; Mohamed, A.; Vatin, N.I. Advanced Machine Learning Modeling Approach for Prediction of Compressive Strength of FRP Confined Concrete Using Multiphysics Genetic Expression Programming. Polymers 2022, 14, 1789. [CrossRef]
39. Nafees, A.; Khan, S.; Javed, M.F.; Alrowais, R.; Mohamed, A.M.; Mohamed, A.; Vatin, N.I. Forecasting the Mechanical Properties of Plastic Concrete Employing Experimental Data Using Machine Learning Algorithms: DT, MLPNN, SVM, and RF. Polymers 2022, 14, 1583. [CrossRef]
40. Nafees, A.; Amin, M.N.; Khan, K.; Nazir, K.; Ali, M.; Javed, M.F.; Aslam, F.; Musarat, M.A.; Vatin, N.I. Modeling of Mechanical Properties of Silica Fume-Based Green Concrete Using Machine Learning Techniques. Polymers 2022, 14, 30. [CrossRef]
41. Khan, K.; Ahmad, A.; Amin, M.N.; Ahmad, W.; Nazar, S.; Arab, A.M.A. Comparative Study of Experimental and Modeling of Fly Ash-Based Concrete. Polymers 2022, 15, 3762. [CrossRef]
42. Nafees, A.; Javed, M.F.; Khan, S.; Nazir, K.; Farooq, F.; Aslam, F.; Musarat, M.A.; Vatin, N.I. Predictive Modeling of Mechanical Properties of Silica Fume-Based Green Concrete Using Artificial Intelligence Approaches: MLPNN, ANFIS, and GEP. Materials 2021, 14, 7531. [CrossRef] [PubMed]
43. Shah, 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] [PubMed]
44. Shafabakhsh, G.H.; Anj, O.J.; Talebsafa, M. Artificial neural network modeling (ANN) for predicting rutting performance of nano-modified hot-mix asphalt mixtures containing steel slag aggregates. Constr. Build. Mater. 2015, 85, 136–143. [CrossRef]
45. Khan, K.; Ahmad, W.; Amin, M.N.; Aslam, F.; Ahmad, A.; Al-Faidi, M.A. Comparison of Prediction Models Based on Machine Learning for the Compressive Strength Estimation of Recycled Aggregate Concrete. Materials 2022, 15, 3430. [CrossRef]
46. Sharma, N.; Thakur, M.S.; Sihag, P.; Malik, M.A.; Kumar, R.; Abbas, M.; Saleel, C.A. Machine Learning Techniques for Evaluating Concrete Strength with Waste Marble Powder. Materials 2022, 15, 5811. [CrossRef]
47. Shah, H.A.; Yuan, Q.; Akmal, U.; Shah, S.A.; Salmi, A.; Awad, Y.A.; Shah, L.A.; Ifikhar, Y.; Javed, M.H.; Khan, M.I. Application of Machine Learning Techniques for Predicting Compressive, Splitting Tensile, and Flexural Strengths of Concrete with Metakaolin. Materials 2022, 15, 5435. [CrossRef]
48. Qi, C.; Huang, B.; Wu, M.; Wang, K.; Yang, S.; Li, G. Concrete Strength Prediction Using Different Machine Learning Processes: Effect of Slag, Fly Ash and Superplasticizer. Materials 2022, 15, 5369. [CrossRef]
49. Khan, K.; Ahmad, W.; Amin, M.N.; Ahmad, A.; Alabdullah, A.A. Compressive Strength Evaluation of Steel-Fiber-Reinforced Concrete and Raw Material Interactions Using Advanced Algorithms. Polymers 2022, 14, 3065. [CrossRef]
50. Song, H.; Ahmad, A.; Farooq, F.; Ostrowski, K.A.; Mašlak, M.; Czarnecki, S.; Aslam, F. Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. Constr. Build. Mater. 2021, 308, 125021. [CrossRef]
51. Sufian, M.; Ullah, S.; Ostrowski, K.A.; Ahmad, A.; Zia, A.; Śliwa-Wieczorek, K.; Siddiq, M.; Awan, A.A. An Experimental and Empirical Study on the Use of Waste Marble Powder in Construction Material. Materials 2021, 14, 3829. [CrossRef]
52. Song, Y.-Y.; Ying, L.U. Decision tree method: Applications for classification and prediction. Shanghai Arch. Psychiatry 2015, 27, 130. [PubMed]
53. Hillebrand, E.; Medeiros, M.C. The benefits of bagging for forecast models of realized volatility. Econom. Rev. 2010, 29, 571–593. [CrossRef]
54. Ahmad, A.; Ahmad, W.; Aslam, F.; Joyklad, P. Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques. Case Stud. Constr. Mater. 2022, 16, e00840. [CrossRef]
55. Friedman, J.H. Greedy function approximation: A gradient boosting machine. Ann. Stat. 2001, 29, 1189–1232. [CrossRef]
56. 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]
57. 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]
58. Grömping, U. Variable importance assessment in regression: Linear regression versus random forest. *Am. Stat.* 2009, 63, 308–319. [CrossRef]

59. 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]

60. Ahmad, A.; Chaiyasarn, K.; Farooq, F.; Ahmad, W.; Suparp, S.; Aslam, F. Compressive Strength Prediction via Gene Expression Programming (GEP) and Artificial Neural Network (ANN) for Concrete Containing RCA. *Buildings* 2021, 11, 324. [CrossRef]

61. Farooq, F.; Ahmed, W.; Akbar, A.; Aslam, F.; Alyousef, R. Predictive modeling for sustainable high-performance concrete from industrial wastes: A comparison and optimization of models using ensemble learners. *J. Clean. Prod.* 2021, 292, 126032. [CrossRef]

62. Aslam, F.; Farooq, F.; Amin, M.N.; Khan, K.; Waheed, A.; Akbar, A.; Javed, M.F.; Alyousef, R.; Alabdulijabbar, H. Applications of gene expression programming for estimating compressive strength of high-strength concrete. *Adv. Civ. Eng.* 2020, 2020, 885035. [CrossRef]

63. 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] [PubMed]

64. Lundberg, S. A Game Theoretic Approach to Explain the Output of Any Machine Learning Model. *Github*. 2021. Available online: https://github.com/slundberg/shap (accessed on 29 September 2022).

65. 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]

66. Molnar, C. Interpretable Machine Learning. Lulu.com. 2020. Available online: https://christophm.github.io/interpretable-ml-book/ (accessed on 29 September 2022).

67. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Processing Syst.* 2017, 30. [CrossRef]

68. 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]

69. Islam, G.S.; Rahman, M.; Kazì, N. Waste glass powder as partial replacement of cement for sustainable concrete practice. *Int. J. Sustain. Built Environ.* 2017, 6, 37–44. [CrossRef]

70. Matos, A.M.; Sousa-Coutinho, J. Durability of mortar using waste glass powder as cement replacement. *Constr. Build. Mater.* 2012, 36, 205–215. [CrossRef]

71. 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]

72. 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]