

Article

A Two-Level Machine Learning Prediction Approach for RAC Compressive Strength

Fei Qi ¹ and Hangyu Li ^{2,*}¹ School of Civil Engineering, Xinyang College, Xinyang 464000, China² School of Civil Engineering, Southeast University, Nanjing 211189, China

* Correspondence: civillyh@163.com

Abstract: Through the use of recycled aggregates, the construction industry can mitigate its environmental impact. A key consideration for concrete structural engineers when designing and constructing concrete structures is compressive strength. This study aims to accurately forecast the compressive strength of recycled aggregate concrete (RAC) using machine learning techniques. We propose a simplified approach that incorporates a two-layer stacked ensemble learning model to predict RAC compressive strength. In this framework, the first layer consists of ensemble models acting as base learners, while the second layer utilizes a random forest (RF) model as the meta-learner. A comparative analysis with four other ensemble learning models demonstrates the superior performance of the proposed stacked model in effectively integrating predictions from the base learners, resulting in enhanced model accuracy. The model achieves a low mean absolute error (MAE) of 2.599 MPa, a root mean squared error (RMSE) of 3.645 MPa, and a high R-squared (R^2) value of 0.964. Additionally, a Shapley (SHAP) additive explanation analysis reveals the influence and interrelationships of various input factors on the compressive strength of RAC, aiding design and construction professionals in optimizing raw material content during the RAC design and production process.

Keywords: machine learning; recycled aggregate concrete; regression prediction; Shapley additive explanations; compressive strength



Citation: Qi, F.; Li, H. A Two-Level Machine Learning Prediction Approach for RAC Compressive Strength. *Buildings* **2024**, *14*, 2885. <https://doi.org/10.3390/buildings14092885>

Academic Editors: Binsheng (Ben) Zhang and Alberto Taliercio

Received: 19 August 2024

Revised: 7 September 2024

Accepted: 11 September 2024

Published: 12 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Concrete stands as the predominant construction material on a global scale. A fraction amounting to 8% of global carbon dioxide released pollutants is linked to concrete production [1,2]. Furthermore, a large amount of natural aggregates (NAs), such as river sand and natural stone, are required in the manufacturing of concrete mixtures [3]. Consequently, ensuring a consistent and sufficient supply of concrete aggregates is crucial for the sustainability of the construction sector. Short-term deficiencies in aggregate availability have the potential to result in project delays and escalated costs.

In China and various other countries, the demand for NAs has experienced rapid growth, primarily driven by infrastructure construction projects. The depletion of existing NA resources necessitates the exploration and extraction of aggregates from new areas. However, these mining processes are not environmentally friendly and can lead to adverse impacts on the overall environment and ecosystems. Statistics show that the construction industry has a significant impact on the environment, accounting for over 50% of the exploitation of natural resources and 39% of carbon dioxide emissions [4]. These waste materials are often disposed of in landfills with the absence of recycling, causing detrimental effects on natural ecosystems. Therefore, employing recycled aggregates sourced from building and demolition debris can make a substantial contribution to environmental conservation and improve the sustainability of the architecture industry [5]. As a result, there has been a growing focus on investigating the repurposing of building and dismantling materials in recent years [6,7].

While extensive research has indeed been conducted on the physical and mechanical characteristics of recycled aggregate concrete (RAC), exploratory studies are still needed to address specific gaps and challenges. In particular, adopting efficient and intelligent methods to preliminarily assess and better understand the mechanical properties of RAC would greatly reduce testing costs and resource consumption. This approach will help substantiate RAC's practical application and enhance its implementation in various contexts. Extensive investigations have been carried out regarding RAC's compressive strength (CS), a critical mechanical attribute. Since recycled aggregates are a part of RAC, the relationship between the mechanical attribute and its constituent parts is typically complex [8]. The mechanical characteristics of concrete and the recycled components interact in a highly nonlinear way due to the diverse composition of recycled aggregates [9]. Currently, research on the compressive strength of RAC primarily relies on laboratory compression tests. However, these tests are often time-consuming and resource-intensive, particularly when adjusting the mix proportions to achieve the desired compressive strength. Additionally, traditional empirical formulas, such as polynomial and exponential functions incorporating parameters like rebound value and ultrasonic pulse velocity, can result in relatively large errors when estimating compressive strength [10–12]. Therefore, being able to provide an accurate estimate of compressive strength for a specific RAC mix before conducting laboratory compression tests would be highly valuable. In fact, numerous researchers have already conducted extensive laboratory compression tests on RAC, and the results from these tests represent a valuable resource. Fully utilizing and mining these data would have significant engineering value.

With the progress of artificial intelligence and computer technology, more intelligent methods are being developed for this goal. A data-driven method has been provided for learning and discerning relationships between the CS of RAC and its constituent elements. Specifically, the models markedly decrease the time and cost needed for sample preparation and testing. Previous studies have explored various intelligent models for predicting RAC compressive strength. For example, Aliakbar et al. [13] developed a prediction equation for the 28-day compressive strength of RAC based on 650 samples using gene expression programming (GEP), but the equation was complex and considered only a few influencing factors. Using three models, i.e., the artificial neural network (ANN), the adaptive neuro-fuzzy inference system (ANFIS), and multiple linear regression (MLR), with 14 input parameter variables, Faeze et al. [14] obtained promising results in predicting the 28-day compressive strength of RAC. Three models were used by Neela et al. [15] to forecast the 28-day compressive strength of RAC. Among them, ANN demonstrated better predictive performance compared to non-linear regression (NLR) and model tree (MT). Catherina et al. [16], Duan et al. [17], Mohamad et al. [18], and Jesús et al. [19] have also employed ANN for the same goal. According to the evaluation results, ANN demonstrated a better degree of prediction accuracy and can be a valuable tool for forecasting the compressive strength of RAC made from various types and sources of recycled aggregates.

In addition to simple ANN models, other models such as support vector machine (SVM), extreme gradient boosting (XGBoost), gradient boosting, Gaussian process, deep learning, and convolutional neural networks (CNNs) have also been employed to forecast the compressive strength of RAC [20–29]. Van et al. [30] utilized the particle swarm optimization (PSO) algorithm to form the corresponding hybrid models with gradient boosting (GB), XGBoost, and support vector regression (SVR). In these models, the GB_PSO model demonstrated the highest correlation coefficient ($R = 0.936$), accompanied by the lowest RMSE (5.560 MPa) and MAE (4.288 MPa). To explore the predictive performances of different models, Uma et al. [31] compared seven different machine learning models, namely XGBoost, K-nearest neighbors (KNNs), ANN, SVM, linear regression, decision tree, and random forest (RF) in the RAC compressive strength prediction. Of these models, the XGBoost model showed superior testing performance.

Based on the provided information, it is clear that there is an increasing focus on the application of machine learning models to forecast the compressive strength of RAC. AI-

though previous studies have proposed machine learning-based models, such as traditional single models, hybrid models, and ensemble models for predicting the RAC compressive strength, limited studies have been undertaken on utilizing stacking learning models for forecasting the strength outcomes of RAC [32]. Stacking models, due to their combinations of predictions from multiple strong learners and their advantages, often achieve better results. Therefore, it is valuable to investigate and promote stacking learning models to provide more robust and powerful solutions for predicting the compressive strength of RAC. This research aims to fill this gap by exploring the potential of stacking models in predicting RAC compressive strength. In addition, compared to previous studies [33,34] that only considered using recycled coarse aggregate, this work has been improved to include samples where recycled fine aggregates are utilized to replace their natural counterpart.

Besides accurately predicting the compressive strength of RAC, understanding the impact of individual input parameters on the output and adjusting the parameter values to achieve the desired compressive strength is both necessary and meaningful. A thorough examination of the significance of input parameters is highly important, yet such studies are scarce. Therefore, leveraging the fusion strategy of stacking models, this study proposes a stacking model-based method for estimating the RAC compressive strength and suggests Shapley (SHAP) additive explanations to examine how the input feature affects the output, guiding the design of RAC mix proportion.

2. Methodology

2.1. Stacked Model

Stacking is an ensemble technique that operates in parallel, amalgamating the outputs of multiple learners through a multi-layer learning framework. The core idea behind the stacking model involves employing base learners in the initial layer to train on the original data. Subsequently, the output from each base learner is utilized to construct a new dataset. Ultimately, the meta-learner in the second layer generates the final predictions, being trained on this new dataset with the output from the base learners as input. Figure 1 shows the framework for a two-layer stacked model.

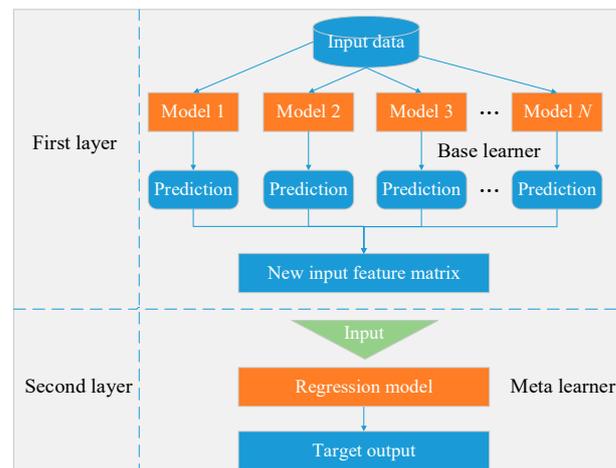


Figure 1. Framework strategy for a two-layer stacked model.

2.2. Learners

When constructing the stacking model, it is essential to consider both the accuracy and diversity of the models when selecting the base learners. By combining predictions from several models, ensemble learning produces forecasts that are often more accurate and reliable than those from any one model alone. Thus, four effective ensemble learning models, i.e., RF, gradient boosting decision tree (GBDT), XGBoost, and LightGBM (LGBM), were chosen to serve as base learners in this study. A model with great generalization capacity is usually favored for the role of the meta-learner, and the RF model was selected

for this role. A synopsis of the four learners may be found below. For more detailed information about these four models, please refer to the references [35,36], and this study will not further elaborate on them.

2.2.1. RF

A classical and popular ensemble learning model is RF. The core idea of this algorithm is to randomly select samples and feature parameters in the model. As a representative algorithm of the bagging method, the RF model adopts bootstrap resampling from the original data, randomly selecting multiple samples, modeling each sample into a decision tree, and then combining them into multiple decision trees. Every decision tree has its independent training procedure, allowing the training of random forest to be conducted concurrently, greatly improving efficiency. The prediction result for input samples is the average output of multiple trees. This method has advantages such as robust capacity for generalization and quick training.

2.2.2. GBDT

The GBDT model is a decision tree model that operates iteratively based on the ensemble learning “Boosting” concept, which combines additive modeling and the forward distribution algorithm. It utilizes decision trees as lesser-capacity models, training current weak learner with the negative gradient of the loss function from the preceding lesser-capacity models, and updating the weights of the training set. Finally, it obtains the strong learner by summing the weighted results of each round of training weak learners. The GBDT model has the benefit of being able to produce superior prediction outcomes for a variety of features.

2.2.3. XGBoost

The ensemble learning technique GBDT has been expanded into XGBoost, which includes improvements to the loss function and the optimization process for minimizing the loss. Compared to the GBDT model, XGBoost provides integration within the loss function through a second-order Taylor series expansion, which improves the model’s accuracy. XGBoost also improves the loss function’s regularization term, preventing overfitting and enhancing the model’s generalization capability.

2.2.4. LGBM

LGBM is another decision tree-based algorithm released by Microsoft Research in 2017, which is a gradient boosting framework that utilizes decision trees based on histogram methods. It is also the GBDT model’s optimization. Two strategies are used by the LGBM model: exclusive feature bundling (EFB) and gradient-based one-side sampling (GOSS). Samples with bigger gradients are given more weight in the GOSS strategy. The EFB technique reduces dimensionality within features by grouping together some mutually incompatible features. The LGBM model has the following benefits: it requires less memory, trains models quickly, and supports parallel training.

2.3. Feature Analysis Using SHAP

In addition to accurate output predictions, understanding how input parameters affect the result is essential. It can effectively guide the design of the RAC mix proportions, and adjusting the mix proportions for decreasing the importance of input parameters will significantly save time and costs. Therefore, the model interpretation based on SHAP analysis is introduced to analyze the importance of features and their impact on the output. SHAP was proposed by Lundberg and Lee in 2017 [37]. It presents the idea of additive explanations, offering a unified method for comprehending predictions made by models.

The additive feature attribution technique known as SHAP, which derives from cooperative game theory, yields an explainable output for the model. The Shapley value, as

referred to by [38], quantifies the contribution of each characteristic. Consequently, the explanation of the model $g(x')$ can be characterized as follows:

$$g(x') = \varphi_0 + \sum_{i=1}^M \varphi_i x'_i \tag{1}$$

The vector of simplified input variables, denoted by x' , is derived from the dataset's original input variables, x . M denotes the dataset's feature count. When there are no inputs at all, φ_0 is a constant; φ_i represents the attribution value of each feature i . For each individual model, it meets the following condition:

$$\varphi_i(\phi, x) = \sum_{z' \subseteq x'} \frac{|z'|!(M - |z'| - 1)!}{M!} [\phi_x(z') - \phi_x(z' \setminus i)] \tag{2}$$

where $|z'|$ denotes the non-zero entries of z' and z' represents the potential subsets of x' . Clearly, directly solving the above equation would be computationally burdensome as a result of the various feature subset options. Therefore, since TreeSHAP matches tree-based ensemble models naturally and offers effective Shapley value computation for all features, it is used here. Once determined, the Shapley values can be employed to easily interpret the model's output. The procedural flowchart for the method employed in this paper is illustrated in Figure 2. First, the dataset is normalized to eliminate scale effects and then is randomly divided into training and test sets. Next, machine learning models undergo parameter tuning based on cross-validation and grid search methods. The performances of different models are compared using evaluation metrics on the test set. Finally, the SHAP analysis is introduced to analyze the importance and impact of input features on the prediction output.

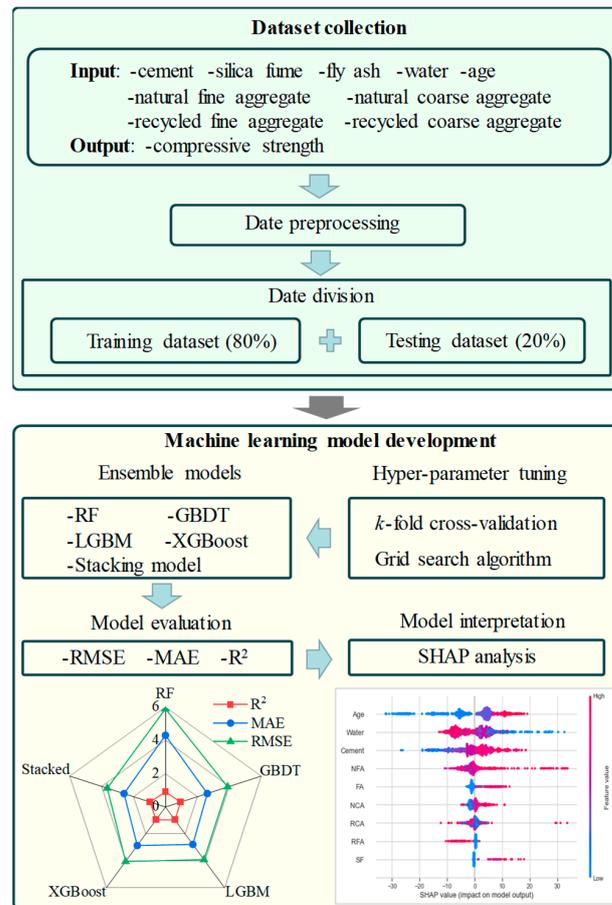


Figure 2. Methodology framework of this work.

3. Dataset Description and Analysis

To establish an accurate compressive strength prediction model, a comprehensive experimental database was constructed [39], containing 1100 samples. The nine input feature variables were selected, including recycled fine aggregate (RFA), water content, cement content, silica fume (SF) content, natural coarse aggregate (NCA), fly ash (FA) content, natural fine aggregate (NFA), recycled coarse aggregate (RCA), and age. The output feature variable was the RAC compressive strength. Table 1 provides a summary of the statistical information pertaining to the dataset, which includes the maximum, minimum, mean, standard deviation, kurtosis, and skewness based on all samples. In Figure 3, the distributions of these variables are depicted. The compressive strength (output variable) ranged from 5.96 MPa to 121.42 MPa. Additionally, it can be observed that for SF, FA, NFA, NCA, RFA, and RCA, the ranges of these component content variations are quite large, indicating a certain representativeness of the dataset.

Table 1. Attributes, range, and statistical features of variables.

Parameter	Cement	SF	FA	Water	NFA	NCA	RFA	RCA	Age	CS
Unit	kg/m ³	days	MPa							
Max	600	50	227.5	271	1065	1366	1000	1632	180	121.42
Min	140	0	0	120	0	0	0	0	1	5.96
Mean	356.00	1.09	27.29	193.32	650.11	529.91	32.98	518.49	32.66	41.50
SD	72.72	6.14	58.07	26.01	222.20	447.16	134.73	427.67	36.41	18.66
Kurtosis	1.77	37.72	2.71	−0.29	2.64	−1.55	21.12	−1.10	4.84	2.10
Skewness	−0.88	6.07	1.97	−0.32	−1.42	−0.03	4.52	0.27	2.07	1.01
Type	Input	Input	Output							

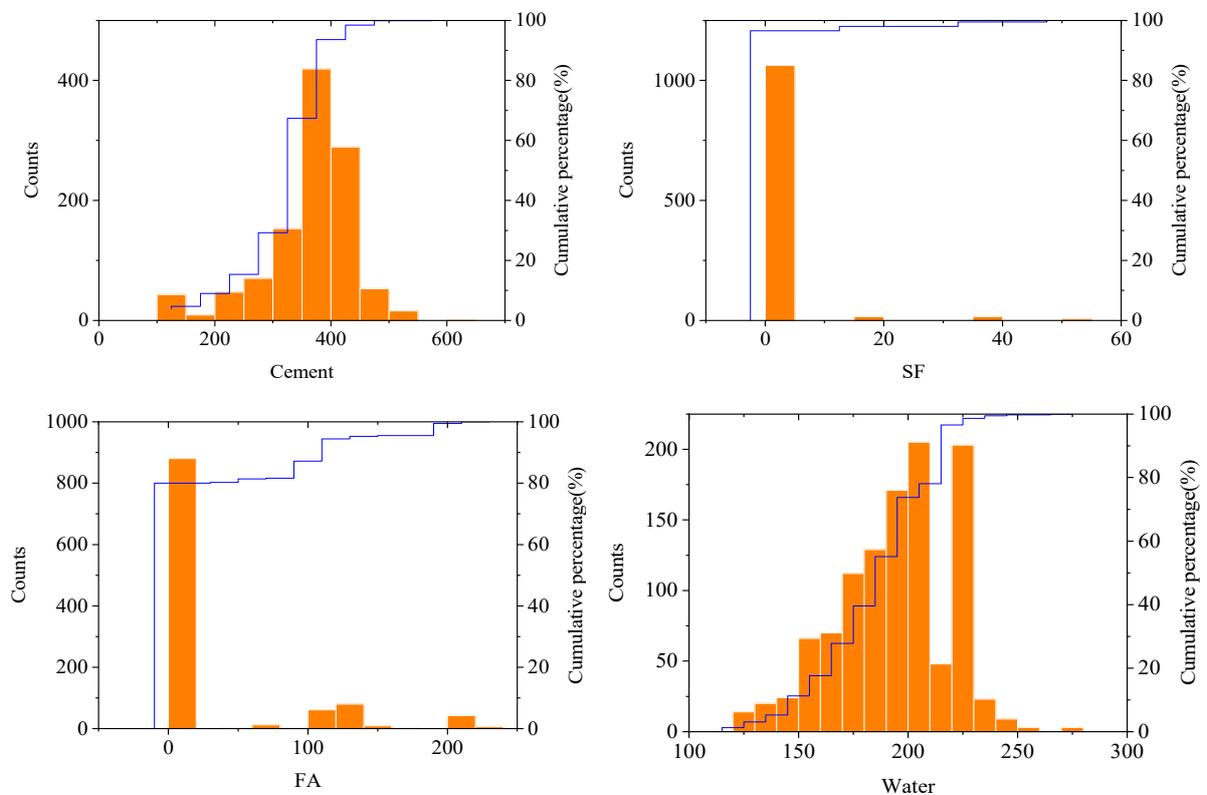


Figure 3. Cont.

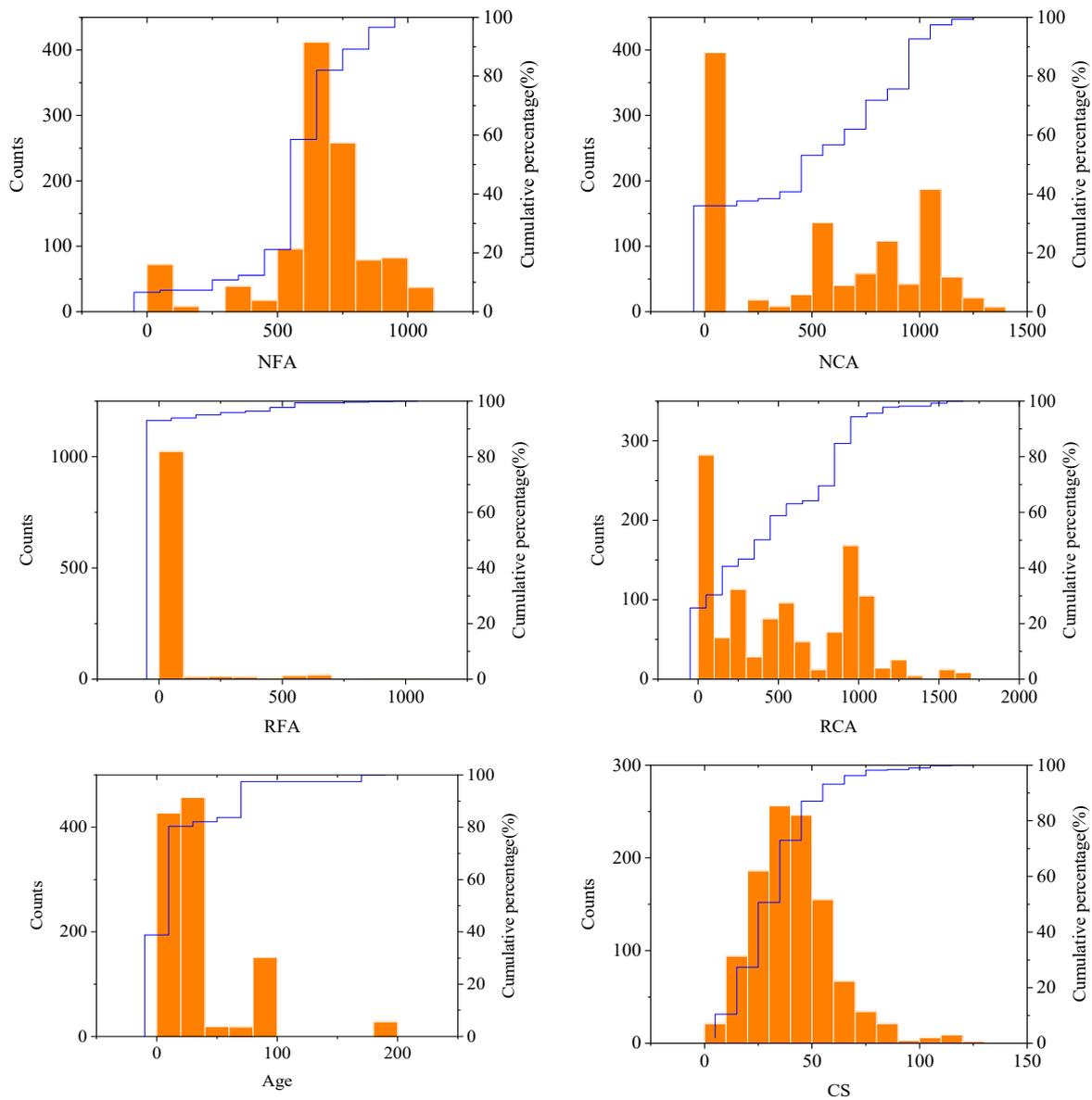


Figure 3. Frequency distribution histograms of the counts and cumulative percentages for individual variables.

To examine the relationships between input factors and output results, Figure 4 depicts the Pearson correlation coefficient matrix. Coefficients close to 1 in the matrix indicate significant positive associations between two variables, while those close to -1 suggest strong negative correlations. From Figure 4, the two variables with the strongest linear correlation to CS are age and water, with coefficients of 0.37 and -0.36 , respectively. These values represent the positive and negative effects on CS, respectively. Nevertheless, the linear relationships among individual input factors and the output results can be largely disregarded, suggesting that navigating the intricate non-linear correlation among numerous factors is a challenging task. It is the primary motivation behind the study's adoption of machine learning techniques to accomplish this objective. The 880 samples (training set) and 220 samples (test set) were randomly selected from the whole dataset.

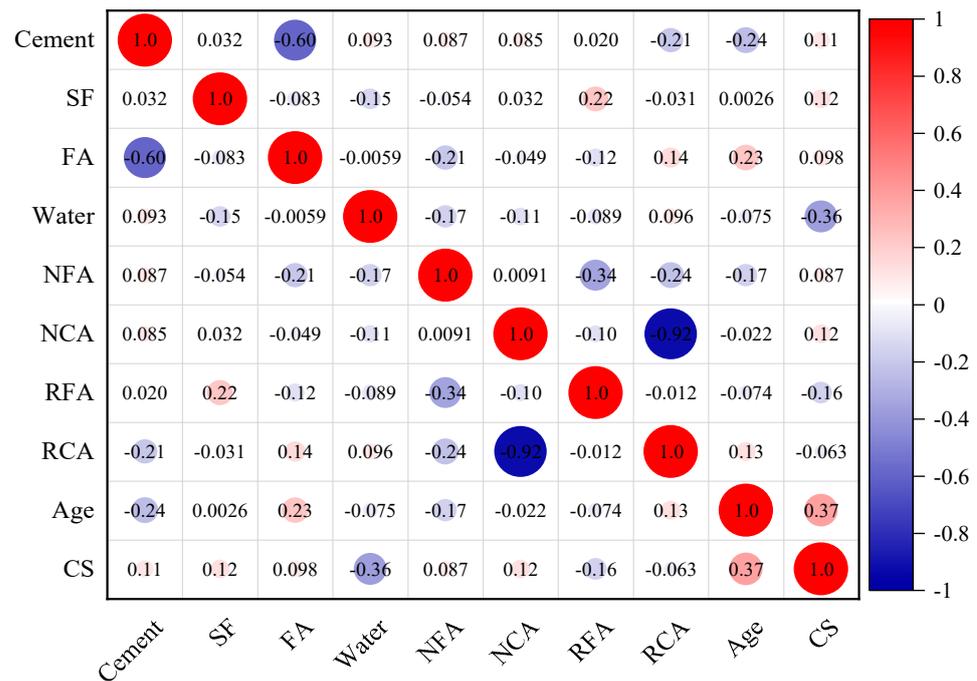


Figure 4. Pearson correlation coefficients among the variables.

4. Model Building and Prediction Results

4.1. Model Implementation

The variable data was preprocessed and standardized during the training process. Cross-validation and grid search were utilized for hyper-parameter tuning, and the final setting values are presented in Table 2. All other parameters not listed are set to the default values of the model. Subsequently, the model's performance was assessed using the test set, and its effectiveness was gauged through the utilization of three widely recognized indicators: the coefficient of determination (R^2), the root mean square error (RMSE), and the mean absolute error (MAE). The definitions of these statistical indexes are shown in Equations (3)–(5).

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_{pre,i} - y_{exp,i})^2}{\sum_{i=1}^N (y_{pre,i} - \bar{y}_{exp})^2} \quad (3)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_{pre,i} - y_{exp,i}| \quad (4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{pre,i} - y_{exp,i})^2} \quad (5)$$

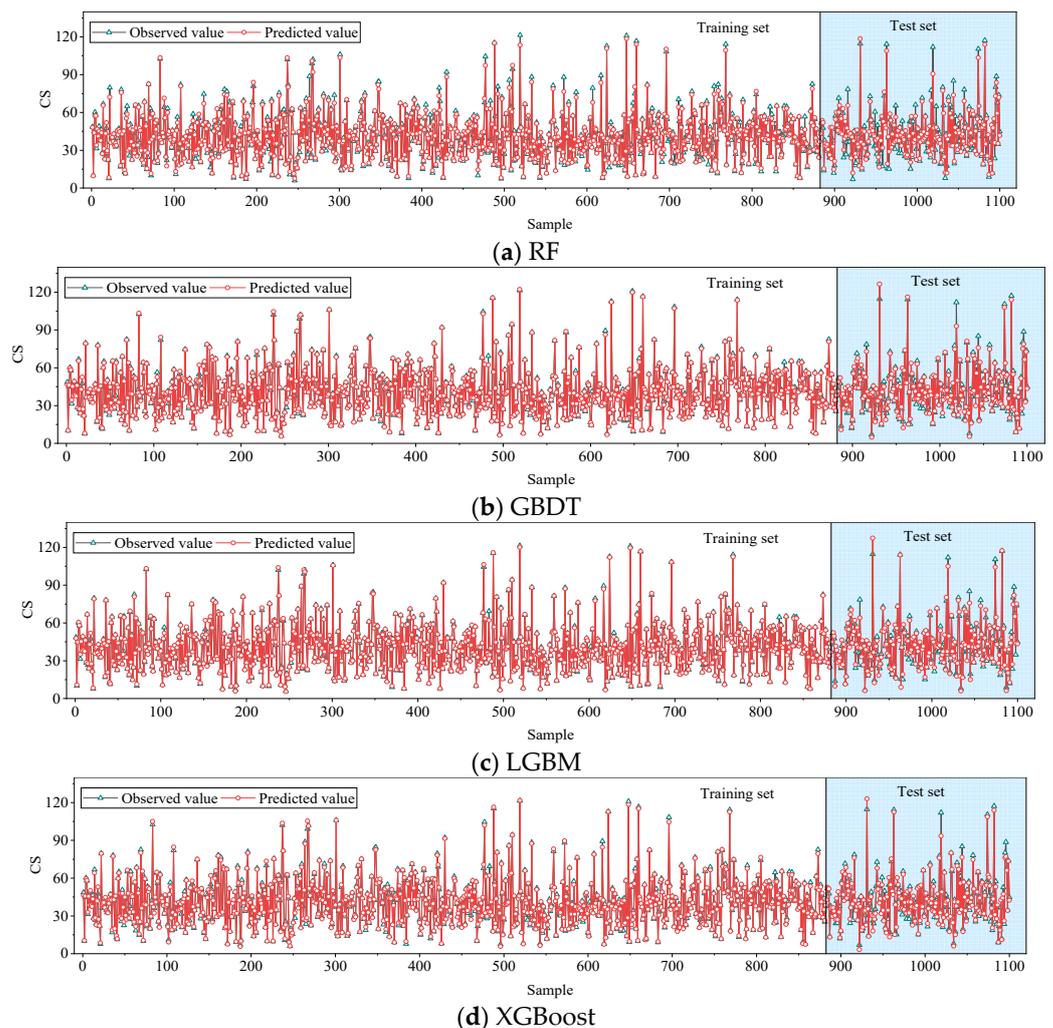
where $y_{exp,i}$ and $y_{pre,i}$ are the actual value and predicted value of the sample and \bar{y}_{exp} is the mean for N samples. A value for R^2 that is closer to 1 denotes a greater degree of accuracy since it shows a stronger connection in the context of the predicted and real results. For RMSE and MAE, smaller values indicate smaller prediction errors and more accurate predictions [40–42].

Table 2. Optimal parameter selections of different ensemble models.

Model	Optimal Value of Hyper-Parameters
RF	n_estimators = 20, criterion = 'squared_error', max_features = 5, max_depth = 25
GBDT	n_estimators = 1050, learning_rate = 0.2, max_depth = 3
LGBM	n_estimators = 2000, learning_rate = 0.3, max_depth = 4
XGBoost	n_estimators = 800, learning_rate = 0.15, max_depth = 3, min_samples_leaf = 1, min_samples_split = 1

4.2. Prediction Comparison of Different Models

Figure 5 displays the predictions obtained from the four ensemble models over all training and testing datasets. For most samples, observations show a high degree of agreement between the expected and actual values. For a more thorough comparison of the model's predictive capacity, the linear correlations between each model's predicted and actual values on the test samples are shown in Figure 6. In contrast to the RF model, the other three models out of the four, which utilize boosting trees and more integrated improvement strategies, exhibit significantly enhanced performance. The distribution of most test samples is positioned closer to the diagonal line, indicating a substantial agreement between the predicted and observed values. Specifically, the LGBM model achieves the strongest correlation coefficient R^2 , which was 0.959, closely trailed by GBDT with 0.958, XGBoost with 0.955, and RF with 0.908.

**Figure 5.** Discrepancies between the observed values and predicted outcomes for four models.

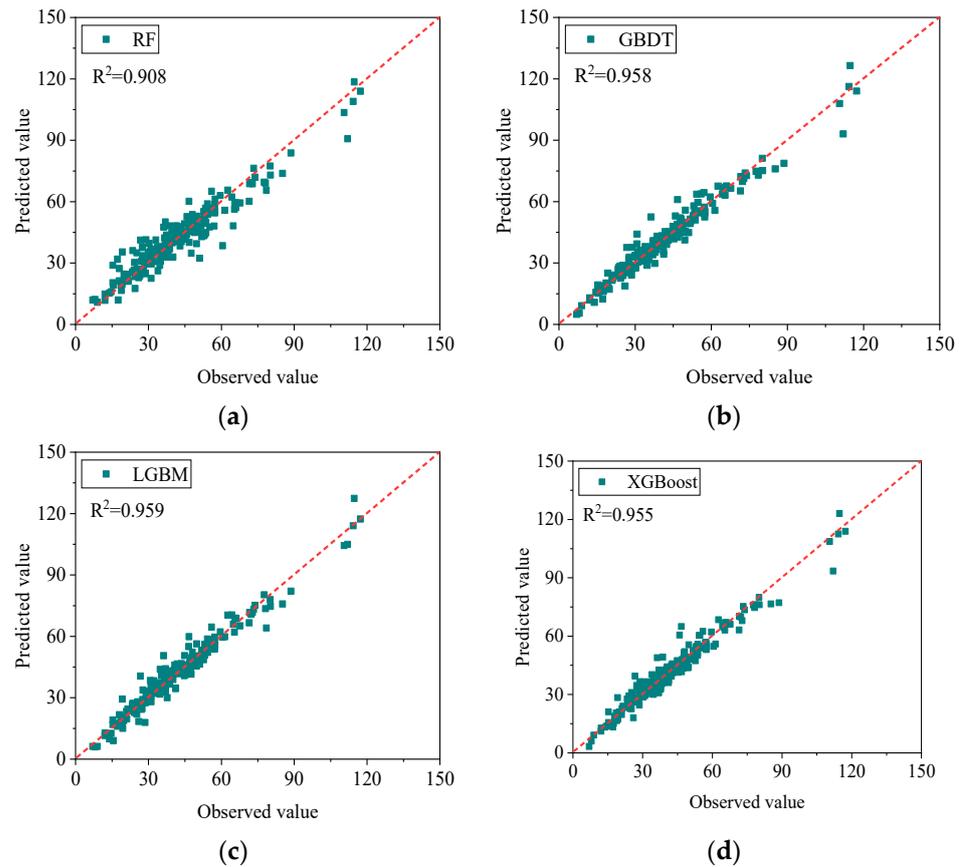


Figure 6. Linear correlations between the predicted results and actual values for four models: (a) RF; (b) GBDT; (c) LGBM; (d) XGBoost.

Figure 7 displays the stacking model testing results adopting RF as the second-layer meta-learner. Compared to the four individual ensemble models, there is a superior linear connection between the stacking model's predictions and the true results, with $R^2 = 0.964$. However, relying solely on a single metric to evaluate the model's performance is not always reliable. Therefore, the error metrics MAE and RMSE are also calculated and plotted for the test set, as depicted in Figure 8. Of the four ensemble models, the GBDT model exhibits the most favorable predictive performance, yielding MAE = 2.624 MPa and RMSE = 3.899 MPa. However, the stacking model exhibits even smaller prediction errors, with MAE = 2.599 MPa and RMSE = 3.645 MPa.

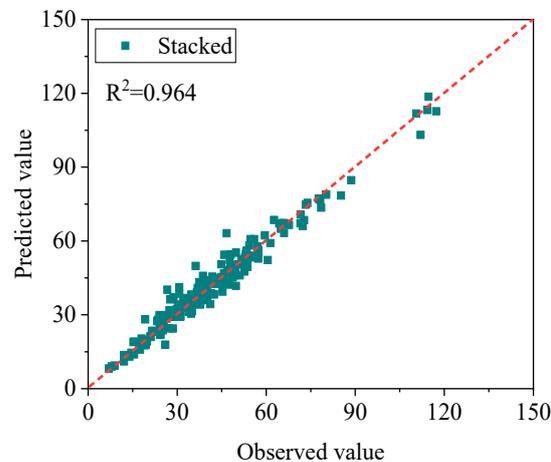


Figure 7. Linear correlation between the predicted values and actual values for the stacked model.

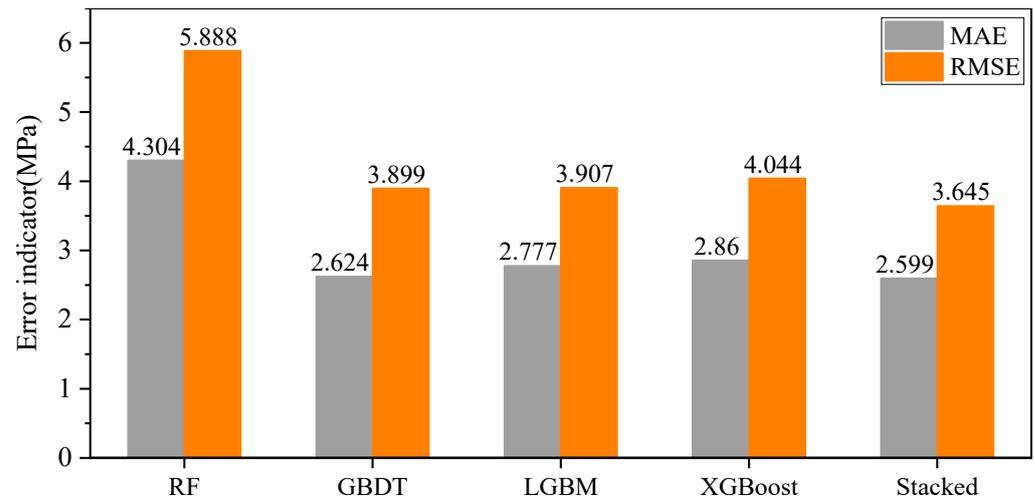


Figure 8. Error statistical indexes for test set.

Figure 9 provides a statistical analysis of the relative errors for testing samples. Among the 220 test samples, 173 (173/220), 163 (163/220), 163 (163/220), 171 (171/220), and 132 (132/220) samples have prediction errors within 10% for RF, GBDT, LGBM, XGBoost, and the stacked model, respectively. Moreover, the stacked model has the fewest samples with prediction errors exceeding 20%, which amounts to only 12 samples. These statistics further illustrate the superiority and reliability of the stacking model.

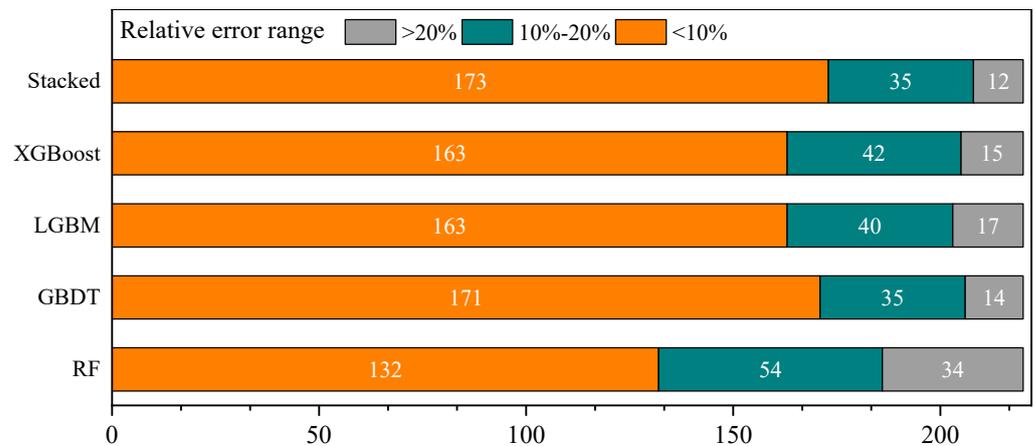
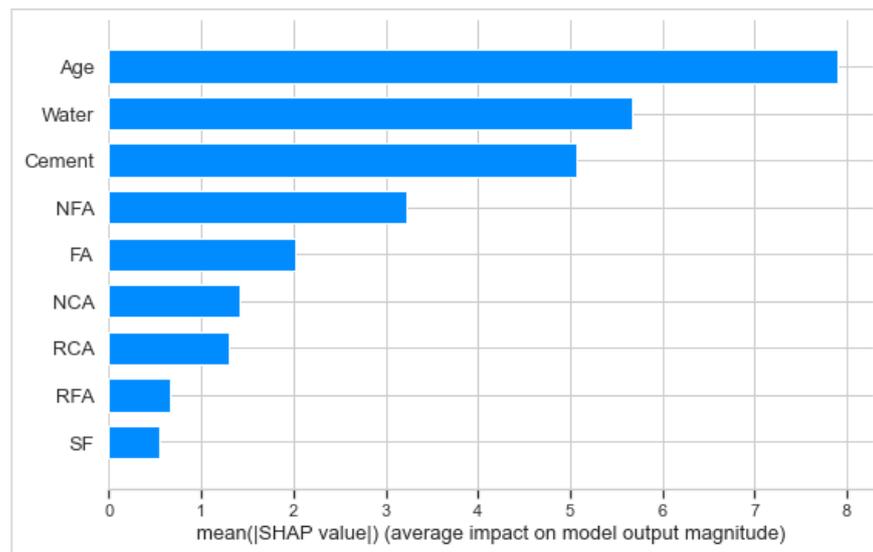


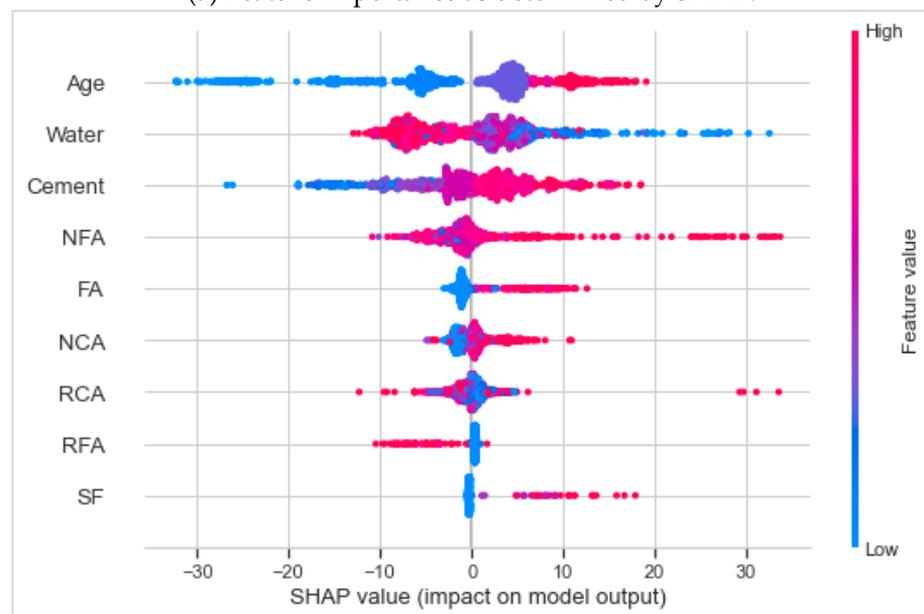
Figure 9. Range distributions of the prediction relative errors.

5. Feature Analysis Using SHAP Interpretation

The predictive outcomes of machine learning models can be explained through different approaches using SHAP. First, as shown in Figure 10a, the feature relevance reveals their overall influence on the prediction. In truth, it is computed as the mean of all the datasets' absolute Shapley values. From the graph, it can be observed that in general, "Age" exerts the greatest impact on compressive strength, making it the most crucial variable, followed by "Water" and "Cement", while "SF" has the least effect. "RCA" and "RFA" also influence the compressive strength, albeit with lower importances compared to "NFA", "FA", and "NCA".



(a) Feature importance as determined by SHAP.



(b) SHAP summary plot.

Figure 10. Global interpretation based on SHAP values.

The feature summary plots are shown in Figure 10b, which also illustrates the effect trends that coincide with each feature's distribution of SHAP values. In Figure 10b, importance is presented along the y -axis, with individual SHAP values represented along the x -axis. Each point represents a sample, with the color of the point indicating a specific value. Ranging from blue to red, the color scale represents values from lower to higher magnitudes. The clustering of points horizontally represents the density of the samples. It can be observed that "Age", "Cement", "FA", "NCA", and "SF" positively influence the model's results. On the contrary, "Water" and "RFA" have negative effects, leading to reductions in the compressive strength of RAC. The "NFA" seems to have a positive effect, but it is not very prominent, requiring further investigation and analysis, similar to the case with "RCA".

6. Graphics User Interface Development

For the convenience of estimating the compressive strength under different mix proportions, a simple graphics user interface (GUI) has been designed in this study for easy

use. As shown in Figure 11, input the actual values of the mix proportions in each text box, then click “Predict”, and the interface will provide a prediction based on the stacking model. For the next set of mix proportion tests, input the content of each component first, then click “Update”, and then click “Predict” to proceed.

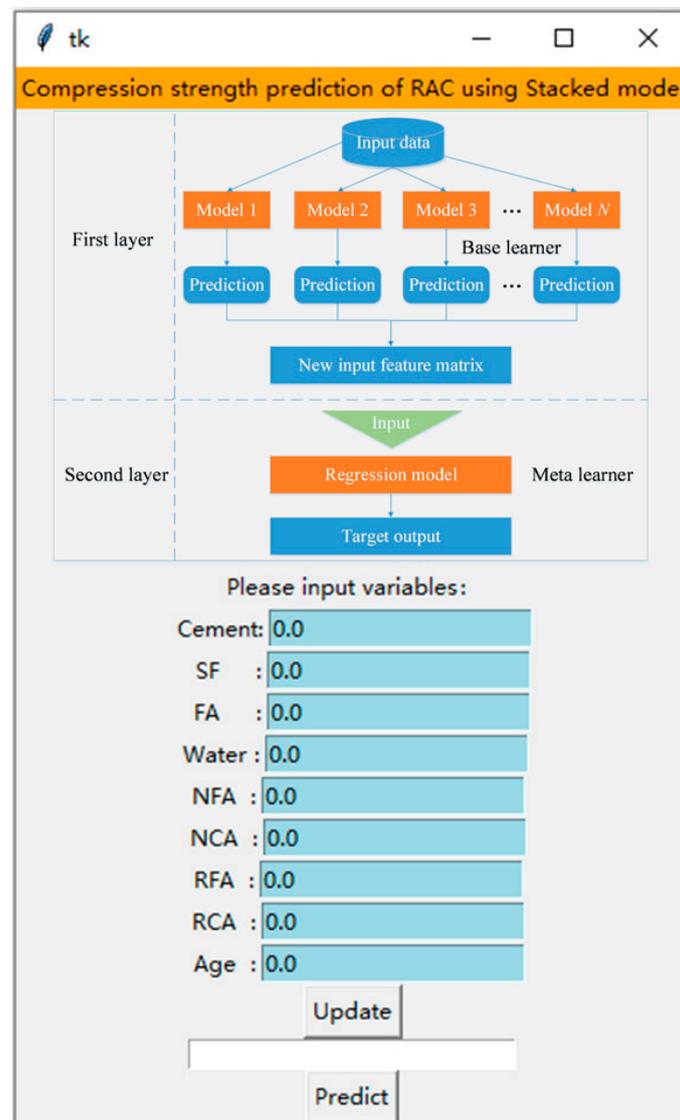


Figure 11. Graphic user interface design.

It is important to highlight that the effectiveness of machine learning models in prediction is significantly influenced by the dataset. As this study is based on a limited dataset (1100 samples), where the data processing and standardization involved the highest and lowest values of each feature, significant deviations may occur in the model’s predictions when the input feature values of new test samples fall outside the range of the dataset used in this study. Therefore, for new data samples that are not within the range of the dataset, it is necessary to retrain the overall dataset with the addition of these new samples. The incorporation of more new datasets is crucial and urgently needed to boost the utility and generalization of the proposed model.

7. Conclusions

This study introduces a two-layer stacked model for estimating the compressive strength of RAC. A dataset containing 1100 samples was collected for training and testing,

and the model's ability to predict the outcomes was compared. The SHAP method was utilized to assess the significance of input variables and their influence on the output.

Compared with four individual models (RF, GBDT, LGBM, and XGBoost), the RAC compressive strength on the test set was predicted more accurately by the suggested stacked model, with $R^2 = 0.964$, MAE = 2.599 MPa, and RMSE = 3.645 MPa. The performance surpassed that of individual base learners, rendering it a valuable tool for predicting the RAC compressive strength. The effects of integrating more base learners and multiple-layer stacking models on improving prediction accuracy will be explored in the future.

SHAP analysis indicates that among the current dataset samples, the input feature "Age" has the largest impact on compressive strength and is the most critical, succeeded by "Water" and "Cement", while the impact of "SF" is the smallest. "RCA" and "RFA" also influence the compressive strength, but their importance is slightly lower than that of "NFA", "FA", and "NCA". These results provide comprehensive insights into how features influence the prediction of compressive strength. The feature importance and interactions between variables can serve as a guide for the mix proportioning of RAC. Modifying the dosage of each component according to the significance of input attributes in relation to output results will significantly reduce the adjustment time for mix proportion and minimize the additional costs associated with production and testing.

It is worth noting that this study is limited by the dataset, which includes only a restricted number of variables as input features. Future research will incorporate additional factors, such as the quality and porosity of the recycled aggregate, into the model to enhance its predictive performance and generalizability. Furthermore, comparing the performance of the proposed model with other models will also be a key focus of future research.

Author Contributions: F.Q.: methodology, writing—review and editing, software, validation. H.L.: formal analysis, methodology, writing—original draft. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

References

1. Olsson, J.A.; Miller, S.A.; Alexander, M.G. Near-term pathways for decarbonizing global concrete production. *Nat. Commun.* **2023**, *14*, 4574. [[CrossRef](#)] [[PubMed](#)]
2. Wu, T.; Ng, S.T.; Chen, J. Deciphering the CO₂ emissions and emission intensity of cement sector in China through decomposition analysis. *J. Clean. Prod.* **2022**, *352*, 131627. [[CrossRef](#)]
3. Li, N.; Unluer, C. Enhancement of the wet carbonation of artificial recycled concrete aggregates in seawater. *Cement. Concr. Res.* **2024**, *175*, 107387. [[CrossRef](#)]
4. García, G.; Cabrera, R.; Rolón, J.; Pichardo, R.; Thomas, C. Systematic review on the use of waste foundry sand as a partial replacement of natural sand in concrete. *Constr. Build. Mater.* **2024**, *430*, 136460. [[CrossRef](#)]
5. Ferrández, D.; Saiz, P.; Zaragoza-Benzal, A.; Zúñiga-Vicente, J.A. Towards a more sustainable environmentally production system for the treatment of recycled aggregates in the construction industry: An experimental study. *Heliyon* **2023**, *9*, e16641. [[CrossRef](#)]
6. Burdier, M.; Anshassi, M.; Guo, Y.; Laux, S.J.; Townsend, T.G. Enhancing the beneficial reuse properties of construction and demolition debris fines using lab-scale washing. *Resour. Conserv. Recycl.* **2022**, *183*, 106361. [[CrossRef](#)]
7. Kabirifar, K.; Mojtahedi, M.; Wang, C.; Tam, V.W.Y. Construction and demolition waste management contributing factors coupled with reduce, reuse, and recycle strategies for effective waste management: A review. *J. Clean. Prod.* **2020**, *263*, 121265. [[CrossRef](#)]
8. Zawal, D.; Grabiec, A.M. Influence of selected mineral additives on properties of recycled aggregate concrete (RAC) considering eco-efficiency coefficients. *Case Stud. Constr. Mater.* **2022**, *17*, e01405. [[CrossRef](#)]
9. Sabbrojjaman, M.; Liu, Y.; Tafsirojjaman, T. A comparative review on the utilisation of recycled waste glass, ceramic and rubber as fine aggregate on high performance concrete: Mechanical and durability properties. *Dev. Built Environ.* **2024**, *17*, 100371. [[CrossRef](#)]
10. Kazemi, M.; Madandoust, R.; Brito, J.D. Compressive strength assessment of recycled aggregate concrete using Schmidt rebound hammer and core testing. *Constr. Build. Mater.* **2019**, *224*, 630–638. [[CrossRef](#)]

11. Datta, S.D.; Sobuz, M.H.R.; Akid, A.S.M.; Islam, S. Influence of coarse aggregate size and content on the properties of recycled aggregate concrete using non-destructive testing methods. *J. Build. Eng.* **2022**, *61*, 105249. [[CrossRef](#)]
12. Mata, R.; Ruiz, R.O.; Nuñez, E. Correlation between compressive strength of concrete and ultrasonic pulse velocity: A case of study and a new correlation method. *Constr. Build. Mater.* **2023**, *369*, 130569. [[CrossRef](#)]
13. 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](#)]
14. Khademi, F.; Jamal, S.M.; Deshpande, N.; Londhe, S. Predicting strength of recycled aggregate concrete using Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System and Multiple Linear Regression. *Int. J. Sustain. Built Environ.* **2016**, *5*, 355–369. [[CrossRef](#)]
15. Deshpande, N.; Londhe, S.; Kulkarni, S. Modeling compressive strength of recycled aggregate concrete by Artificial Neural Network, Model Tree and Non-linear Regression. *Int. J. Sustain. Built Environ.* **2014**, *3*, 187–198. [[CrossRef](#)]
16. Duan, Z.H.; Kou, S.C.; Poon, C.S. Prediction of compressive strength of recycled aggregate concrete using artificial neural network and cuckoo search method. *Mater. Today Proc.* **2021**, *46*, 8480–8488.
17. Duan, Z.H.; Kou, S.C.; Poon, C.S. Prediction of compressive strength of recycled aggregate concrete using artificial neural networks. *Constr. Build. Mater.* **2013**, *40*, 1200–1206. [[CrossRef](#)]
18. Ridho, B.K.A.M.A.; Ngamkhanong, C.; Wu, Y.; Kaewunruen, S. Recycled aggregates concrete compressive strength prediction using Artificial Neural Networks (ANNs). *Infrastructures* **2021**, *6*, 17. [[CrossRef](#)]
19. de-Prado-Gil, J.; Martínez-García, R.; Jagadesh, P.; Juan-Valdés, A.; González-Alonso, M.-I.; Palencia, C. To determine the compressive strength of self-compacting recycled aggregate concrete using artificial neural network (ANN). *Ain. Shams Eng. J.* **2024**, *15*, 102548. [[CrossRef](#)]
20. 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](#)]
21. Salimbahrami, S.R.; Shakeri, R. Experimental investigation and comparative machine-learning prediction of compressive strength of recycled aggregate concrete. *Soft Comput.* **2021**, *25*, 919–932. [[CrossRef](#)]
22. Omer, B.; Jaf, D.K.I.; Abdalla, A.; Mohammed, A.S.; Abdulrahman, P.I.; Kurda, R. Advanced modeling for predicting compressive strength in fly ash-modified recycled aggregate concrete: XGboost, MEP, MARS, and ANN approaches. *Innov. Infrastruct. Solut.* **2024**, *9*, 61. [[CrossRef](#)]
23. Nunez, I.; Marani, A.; Nehdi, M.L. Mixture optimization of recycled aggregate concrete using hybrid machine learning model. *Materials* **2020**, *13*, 4331. [[CrossRef](#)] [[PubMed](#)]
24. 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](#)]
25. Pal, A.; Ahmed, K.S.; Hossain, F.M.Z.; Alam, M.S. Machine learning models for predicting compressive strength of fiber-reinforced concrete containing waste rubber and recycled aggregate. *J. Clean. Prod.* **2023**, *423*, 138673. [[CrossRef](#)]
26. de-Prado-Gil, J.; Palencia, C.; Silva-Monteiro, N.; Martínez-García, R. To predict the compressive strength of self compacting concrete with recycled aggregates utilizing ensemble machine learning models. *Case Stud. Constr. Mater.* **2022**, *16*, e01046. [[CrossRef](#)]
27. Hosseinzadeh, M.; Dehestani, M.; Hosseinzadeh, A. Prediction of mechanical properties of recycled aggregate fly ash concrete employing machine learning algorithms. *J. Build. Eng.* **2023**, *76*, 107006. [[CrossRef](#)]
28. 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](#)]
29. Nguyen, X.H.; Phan, Q.M.; Nguyen, N.T.; Tran, V.Q. Interpretable machine learning model for evaluating mechanical properties of concrete made with recycled concrete aggregate. *Struct. Concr.* **2023**, *25*, 2890–2914. [[CrossRef](#)]
30. Tran, V.Q.; Viet, Q.D.; Ho, S.L. Evaluating compressive strength of concrete made with recycled concrete aggregates using machine learning approach. *Constr. Build. Mater.* **2022**, *323*, 126578. [[CrossRef](#)]
31. Biswal, U.S.; Mishra, M.; Singh, M.K.; Pasla, D. Experimental investigation and comparative machine learning prediction of the compressive strength of recycled aggregate concrete incorporated with fly ash, GGBS, and metakaolin. *Innov. Infrastruct. Solut.* **2022**, *7*, 242. [[CrossRef](#)]
32. Golafshani, E.M.; Kim, T.; Behnood, A.; Ngo, T.; Kashani, A. Sustainable mix design of recycled aggregate concrete using artificial intelligence. *J. Clean. Prod.* **2024**, *442*, 140994. [[CrossRef](#)]
33. Moghaddas, S.A.; Nekoei, M.; Golafshani, E.M.; Behnood, A.; Arashpour, M. Application of artificial bee colony programming techniques for predicting the compressive strength of recycled aggregate concrete. *Appl. Soft. Comput.* **2022**, *130*, 109641. [[CrossRef](#)]
34. 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](#)]
35. Zhang, T.; Zhang, Y.; Wang, Q.; Aganyira, A.K.; Fang, Y. Experimental study and machine learning prediction on compressive strength of spontaneous-combustion coal gangue aggregate concrete. *J. Build. Eng.* **2023**, *71*, 106518. [[CrossRef](#)]
36. Kaloop, M.R.; Kumar, D.; Samui, P.; Hu, J.W.; Kim, D. Compressive strength prediction of high-performance concrete using gradient tree boosting machine. *Constr. Build. Mater.* **2020**, *264*, 120198. [[CrossRef](#)]
37. Lundberg, S.M.; Lee, S.-I. A Unified approach to interpreting model predictions. *arXiv* **2017**, arXiv:1705.07874.

38. Wu, Y.; Zhou, Y. Hybrid machine learning model and Shapley additive explanations for compressive strength of sustainable concrete. *Constr. Build. Mater.* **2022**, *330*, 127298. [[CrossRef](#)]
39. Hoang, N.-D. A novel ant colony-optimized extreme gradient boosting machine for estimating compressive strength of recycled aggregate concrete. *Multiscale Multidiscip. Model. Exp. Des.* **2024**, *7*, 375–394. [[CrossRef](#)]
40. Erdogan, B.; Kocar, O.; Topal, H.I. Measurement of the dynamic viscosity of water-based nanofluids containing Al₂O₃, TiO₂, and ZnO using the Artificial Neural Network method. *Sci. Iran.* **2023**. [[CrossRef](#)]
41. Dizdar, E.N.; Koçar, O. Artificial neural network-based risk assessment for occupational accidents in the shipbuilding industry in Turkey. *Neural Comput. Appl.* **2024**. [[CrossRef](#)]
42. Wu, Y.; Zhou, Y. Prediction and feature analysis of punching shear strength of two-way reinforced concrete slabs using optimized machine learning algorithm and Shapley additive explanations. *Mech. Adv. Mater. Struct.* **2023**, *30*, 3086–3096. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.