# Prediction of Compressive Strength of Eco-friendly Concrete using Polynomial Regression Method

**N. Ramanjaneyulu[1], B. Javheri[2], Dr. M. Nithya[3], Ommi Suresh[4], Amruta Jagdish Killol[5], Manjunatha[6]**

[1]Sr. Assistant Professor, Department of Civil Engineering, CVR College of Engineering, Hyderabad, Telangana, India

[2]Assistant Professor, Department of Civil Engineering, Walchand Institute of Technology, Solapur, Maharashtra, India.

[3]Professor, Department of Civil Engineering, Yashoda Technical Campus, Satara – 415011, Maharashtra, India.

[4]Assistant Professor, Department of Civil Engineering, Dadi institute of Engineering and technology, Anakapalle, Andhra Pradesh, India

[5]Research scholar, Department: Civil engineering, G.H.Raisoni college of Engineering, Nagpur, Maharashtra, India, Email Id: killolamruta9@gmail.com

[6]Assistant Professor, Department of Civil Engineering, B V Raju Institute of Technology, Narsapur, Telangana, India.

**Abstract:** The most popular construction material, concrete, is also a known pollutant that has a negative impact on sustainability due to resource depletion, energy consumption, and greenhouse gas emissions. As a result, in order to boost concrete's long-term viability, efforts should be focused on minimising its environmental effects. This study aimed to develop a prediction model for the compressive strength of various mixes in order to construct ecologically acceptable concrete mixtures. The concrete mixes that were employed in this study to construct our suggested prediction model are concrete mixtures that comprise both ground granulated blast-furnace slag (GGBFS) and recycled aggregate concrete (RAC). To forecast the compressive strength of eco-friendly concrete, the multivariate polynomial regression (MPR) white-box machine learning model was created. The model was contrasted with the other two machine learning models: support vector machine (SVM), a black-box machine learning model, and linear regression (LR), a white-box machine learning model. In terms of R2 (coefficient of determination) and RMSE (root mean absolute error) measures, the newly proposed model beats the previous two models and demonstrates robust estimate capabilities.

**Keywords:** machine learning; compressive strength of concrete; ground granulated blast-furnace slag; recycled concrete aggregate; multivariate polynomial regression (MPR)

## 1. Introduction

Recent decades have seen a rapid growth of industrialisation and urbanisation in the majority of rising nations, which has led to a sharp rise in the demand for natural raw materials. The construction industry has a higher detrimental impact on the environment. It consumes a significant amount of energy and natural resources and generates a lot of garbage from construction and demolition (C&D), for which there are no practical alternatives. Additionally, the by-products produced by industrial operations and others, such as blast-furnace slag, silica fume, fly ash, and ferrochrome slag, add to issues such a lack of land for garbage disposal and growing prices for waste treatment before the dumping. Therefore, economic endeavours ought to be made in order to maintain the ecological balance of the planet [1]. Recycled coarse aggregates (RCA) is the name given to them. Recycled aggregate concrete (RAC) performs less mechanically and sustainably than natural aggregate concrete (NAC), according to study [1,2]. This is due to RCA's lower quality than NCA's. Industry-grade jaw crushers are often used to process recycled concrete aggregate. This most popular technique for preparing aggregate is linked to the development of microcracks in aggregate grains, which are made up of natural aggregate and a cement matrix (which may also contain other additives besides cement). Because linked mortar is present on the RCA, a weaker interfacial transition zone is created, making it porous. The literature has recommended a number of enhancement strategies for enhancing the features of RAC. Adding more cementitious materials is one of the most efficient ways to improve RCA's properties and produce RAC that resembles NAC [3,4]. For many years, blended cement or Portland cement concretes have included additional cementitious components such metakaolin (MK), ground granulated blast-furnace slag (GGBS), silica fume (SF), and fly ash (FA) [5-7]. The use of GGBFS in lieu of regular Portland cement results in a denser matrix, which boosts the strength and durability of concrete and lengthens the service life of concrete buildings. Even though it is generally accepted that adding FA and GGBS to concrete with standard fineness reduces its early strength, these additives have advantages over time, including decreased alkali-silica reaction expansion, decreased porosity and permeability, and increased workability [8,9].

The concrete compressive strength test, which evaluates the concrete's working stress after 3 days, 7 days, 28 days, or 90 days [10], reveals the concrete's characteristic strength. To determine the maximum working stress of concrete during structural design, a common test called the compressive strength test is employed. It is a gauge for quality assurance in a factory or workshop producing concrete.

The flexural and compressive strength of eco-friendly concrete cannot be predicted using typical linear regression since it includes more forecasting characteristics, such as GGBFS or RCA. Forecasting the strength of eco-friendly concrete may be done using machine learning techniques. Using multivariable statistical methods can help you better comprehend and evaluate challenging data [11]. For many years, multivariate polynomial regression (MPR) has been utilised to address a variety of civil engineering challenges, particularly in the domains of building materials [12-14]. Such MPR models have an advantage over black-box models like ANNs since they can be investigated more simply utilising methods like graphical approaches, sensitivity analysis, and the use of variable significance ratios. These qualities allow the MPR to be a very helpful tool. Numerous studies have created various prediction models over the past few years to gauge the compressive strength of eco-friendly concrete [15–22]. Two types of hybridised machine learning techniques—the interval type-2 fuzzy inference system (IT2FIS) and the type-1 fuzzy inference system (T1FIS)—were employed to create a prediction model for the CS that comprises RAC [15]. The results showed that the IT2FIS model performed better than the T1FIS model. In addition, four artificial intelligence approaches were used to study the 28-day RAC concrete compressive strength based on a meta-heuristic search of sociopolitical algorithms (i.e., ICA) [16]. The outcomes show that the proposed ICA-XGBoost model performed better than the other models. In this work [17], machine learning algorithms were used to estimate RAC compressive strength and its ideal mixture design. The findings showed that all of the developed models, including deep learning, Gaussian processes, and gradient boosting regression, produced accurate prediction performance, with gradient boosting regression trees surpassing the others. In a another study, a convolutional neural network was used to construct a prediction model to calculate the compressive strength of RAC [18]. Experimental work was done alongside the deep learning model's creation. The back propagation neural network and the support vector machine were compared with the convolutional neural network model, showing that the convolutional neural network can predict the compressive strength of RAC better than the other two models.

The article employed three distinct machine learning models, including an artificial neural network (ANN), a support vector machine (SVM), and multiple linear regression (MLR), to estimate the compressive strength of environmentally friendly concrete that contains GGBFS. The ANN and SVM approaches were compared with MLR using k-fold cross-validation, and it was shown that the artificial intelligence techniques outperformed MLR. Without include RCA in the concrete mixes, the study [20] employed a random forest method to forecast ground granulated blast-furnace slag concrete (GGBFSC). The curing temperature (T), superplasticizer, fine aggregate (FA), coarse aggregate (CA), water-to-binder ratio (w/b), water content (W), GGBFSC-to-total-binder ratio (GGBFSC/B), and so on (SP) were used to build the prediction model. The RF model performed exceptionally well in predicting CS with limited input parameters, depending on the level of prediction accuracy attained. A method based on random forest (RF) was also recommended in article [21] for predicting concrete compressive strength by adding GGBFS to concrete mixes. According to the results, the RF algorithm demonstrated a high level of prediction accuracy and may be applied to cut down on the expense of experiments. For the purpose of predicting the long-term effects of ground granulated blast-furnace slag on the compressive strength of concrete under wet curing conditions, fuzzy logic and artificial neural network models have been created in this study [22]. In training and testing models, artificial fuzzy logic systems and neural networks have shown great potential for predicting long-term impacts of ground granulated blast-furnace slag on concrete compressive strength.

## 2. Research Significance

The majority of the research described above used artificial intelligence and machine learning algorithms to predict the compressive strength of eco-friendly concrete. Although they do a great job of forecasting the mechanical characteristics of concrete, they operate as a complicated system that is challenging to use and understand. Contrary to current black-box algorithms like artificial neural networks, the MPR model proposed in this paper is substantially more transparent and simple for academics to use. In addition, earlier researchers developed a prediction model for compressive strength for environmentally friendly concrete that uses either GGBFS or RCA as a substitute for conventional cement, but not both. Unlike earlier studies, this one includes a prediction model for concrete's compressive strength.

## 3. Materials and Methods

### *3.1. System Methodology*

We developed an effective method that we may use in this study to predict the compressive strength of eco-friendly concrete. Figure 1 shows the system approach in action. In this framework, we evaluated the performance of the LR (white-box algorithm) and SVM (black-box algorithm) models with the MPR approach to predict the compressive strength of eco-friendly concrete. The system's earliest stage involves gathering experimental data from earlier investigations, as shown in Figure 1. The second part of the process involves splitting the data into training and testing datasets. The final stage involves dividing our training dataset into fivefold cross-validation to see how well the three models performed. The fourth stage involves using k-fold cross-validation to apply all three models. The fifth stage involves comparing the MPR model's performance to that of the other two machine learning models using accepted performance criteria. In order to verify our model, we created a regression equation using a training dataset and applied it to hypothetical data.

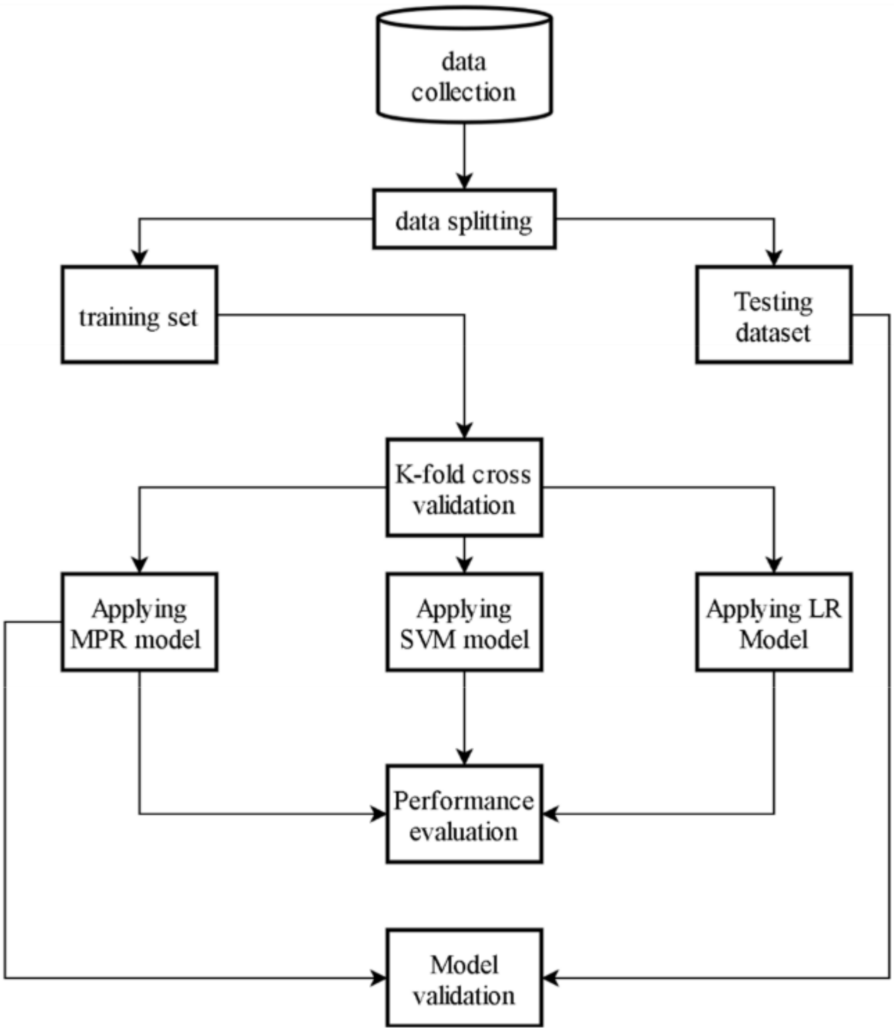
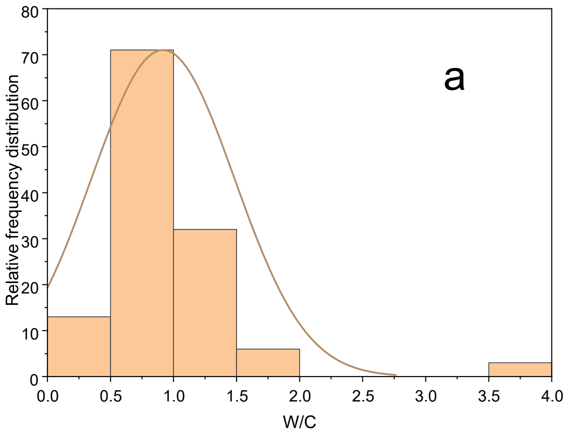
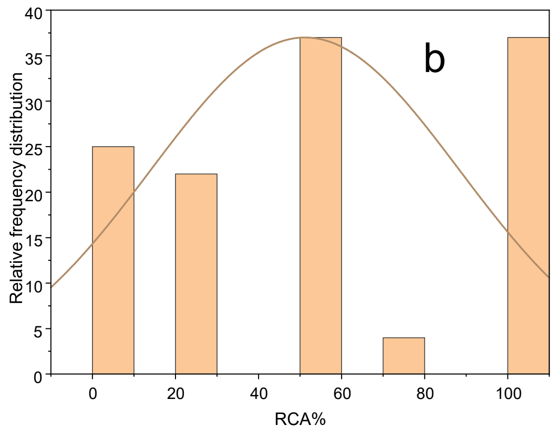



**Figure 1.** Proposed system methodology for eco-friendly concrete compressive strength prediction.

*3.2. Dataset*

The components and number of data samples used fully affect the model's performance. The variables used to build the models used to forecast concrete strength were taken from the literature [23-28], resulting in a total of 125 mix proportions. The following predictors were drawn from the literature and used to construct the models: the ratio of water to cement (W/C), the percentage of recycled aggregate used to replace conventional aggregate in the mixture (RAC%), the percentage of GGBFS used to replace OPC in the binder (GGBFS%), the superplasticizer, and the age (days). Compressive strength (CS) of environmentally friendly concrete was the key variable. The total distribution of all variables in terms of relative frequency is shown in Figure 2. The statistical analysis of the variables is presented in Table 1.



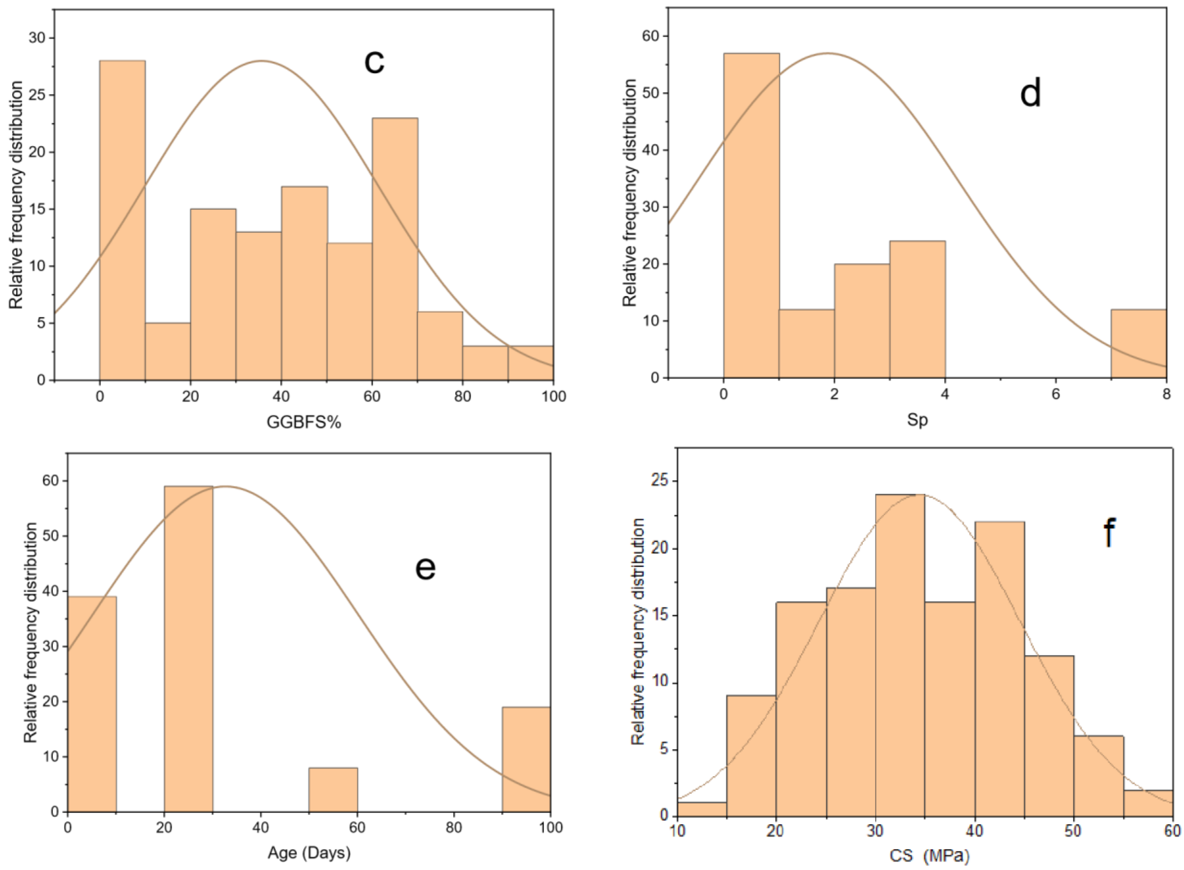



**Figure 2.** Relative frequency distribution of variables: W/C (**a**), RCA% (**b**), GGBFS% (**c**), Sp (kg) (**d**), Age (**e**), CS (**f**).

**Table 1.** Statistical measures on variables.

| Statistics | W/C | RCA% | GGBFS% | Sp (kg) | Age (days) | CS (MPa) |
|---|---|---|---|---|---|---|
| Median | 0.71 | 50 | 40 | 1.15 | 28 | 34.05 |
| Mean | 0.92 | 51.20 | 35.64 | 1.88 | 32.66 | 34.44 |
| Minimum | 0.4 | 0 | 0 | 0 | 7 | 12.4 |
| Maximum | 3.7 | 100 | 90 | 7.8 | 90 | 56.63 |
| Range | 3.3 | 100 | 90 | 7.8 | 83 | 44.23 |
| Standard deviation | 0.57 | 37.14 | 25.82 | 2.36 | 27.50 | 10.15 |

W/C = water-to-cement ratio; RCA = recycled aggregate replacement percentage; GGBFS = ground granulated blast-furnace slag replacement percentage; Sp = superplasticizers quantity in kg; Age = the time of compressive test after pouring; CS = compressive strength.

### 3.3 Data-Splitting Procedure

After acquiring the dataset, we randomly split it into two parts: the training set and the testing set. In this study, the ratio of training to testing sets was 8:2. We trained our model using 120 observations from the training set, and we tested its effectiveness using 25 observations from the testing set. The training sample was selected using a stratified sampling strategy from the initial datasets, ensuring that the outcome distribution across the board is comparable.

### 3.4. MPR Model Development

Polynomial regression was used to fit the connection between dependent and independent variables. The generic equation for the sth-order ($s > 1$) polynomial regression is represented by the mathematical formula shown below [29]:

$$\hat{y} = w_0 + w_1 x + w_2 x^2 + w_3 x^3 + \ldots + w_s x^s \qquad (1)$$

where x is the input variable, y is the output variable, w0 is the intercept, and w1, w2,... ws are the polynomial regression coefficients. MPR is the term for polynomial regression that incorporates multiple variables. For a system with n input variables and the sth-order ($s > 1$), the MPR is represented by the following formula [29]:

$$\hat{y} = w_0 + \sum_{l_1=1}^{n} w_{l_1} x_{l_1} + \sum_{l_1=1}^{n} \sum_{l_2=l_1}^{n} w_{l_1 l_2} x_{l_1} x_{l_2} + \ldots + \sum_{l_1=1}^{n} \sum_{l_2=l_1}^{n} \ldots \sum_{l_k=l_{k-1}}^{n} w_{l_1 l_2 \ldots l_k} x_{l_1} x_{l_2} \ldots x_{l_k} \qquad (2)$$

Even though the MPR applies a non-linear model to the data, the multivariate function (Equation (2)) is linear with regard to its coefficients. The MPR model therefore has the same solution as the MLR issue when the least-squares method is applied. Least-squares approaches are used to identify the polynomial regression coefficients by reducing the sum of squared errors of the predicted vs. actual outcome.

Exponents for MPR predictors may be integers or fractions. This research only considered integer exponents between 3 and +3. The preliminary testing of fractions and integers did not provide any noticeably better models outside of this range.

We employed a TaylorFit [30] programme to offer a conventional mathematical explanation of MPR in the form of a general equation (Equation 2). MPR is simply a multilinear regression extension (MLR) that integrates interaction and other nonlinearities. In addition, the quantity of words expands as the multiplicands and exponents do. The most multiplicands that could be used for any candidate phrase in this inquiry was three. The model was created using the following stepwise algorithm:

- The first step in every model was usually an intercept, which was the average of the values of the dependent variables. The programme generated terms that interacted with the existing model terms in the best way possible depending on the user-selected permitted exponents and multiplicands. The best t-statistics from the fit data were used to order the words.
- 2 Two requirements must be satisfied for a word to be included in the model. First, the candidate term has to have variables of fit that are statistically significant. Second, it is

important to reduce the total RMSE value across all cross-correlation datasets. This approach improved the generalizability of the model and decreased the chance of overfitting.

- The statistical significance of the previously included items was evaluated after each item was added to the model; if they were not statistically significant, they were removed.
- The aforementioned process was repeated for other potential keywords.
- The model was built through an iterative process of adding and removing potential terms from a list of statistically significant terms based on the fit dataset, which also improved the RMSE of the test dataset, up until the point where the model could no longer be improved by adding or removing any individual term.

*3.5. Cross-Validation*

Use of the cross-validation resampling approach to assess machine learning models in a small data sample is one of the best practises. The procedure only has one parameter, k, which determines how many groups should be created from a given data sample. K-fold cross-validation is a common name for the procedure as a result. When an exact k value is given, it can be used in place of k in the reference model, such as k = 5 for 5-fold cross-validation. Cross-validation is typically used in machine learning to assess how well a model performs on fresh data [31]. This would include using a small sample to see how the model will perform in practise when used to provide projections on data that.

## 4. Model Result

4.1. K-Fold Cross-Validation

For each ML model's training performance, the boxplots (Figure 3) show the ranges and fluctuation of the performance metrics, R2, and RMSE. 100 actual values from the 5-fold cross-validation of the training datasets were used to construct these boxplots. Table 2 shows the RMSE and R2 for three machine learning models using the fivefold cross-validation's lowest, maximum, median, and average values.

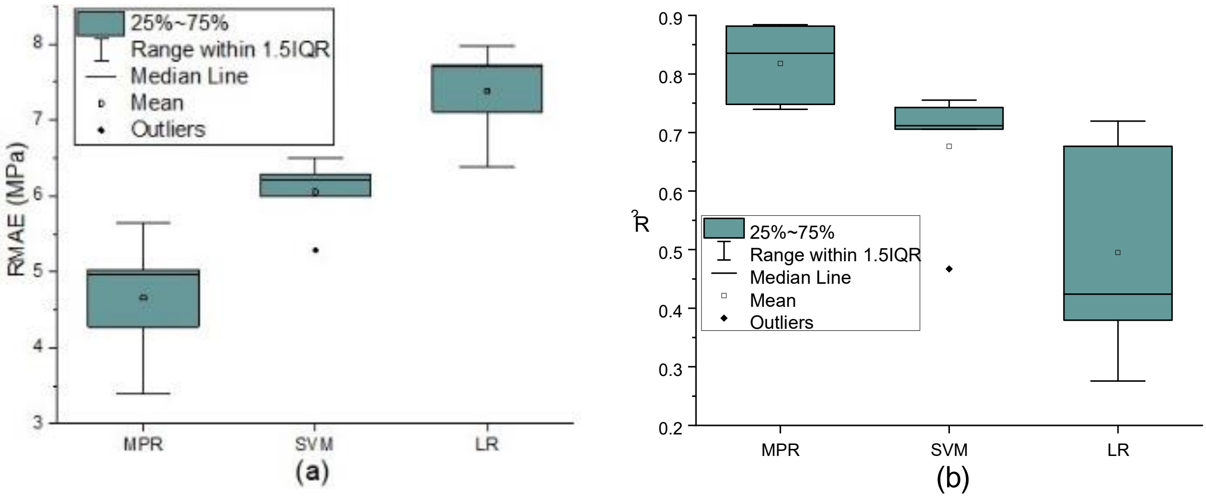


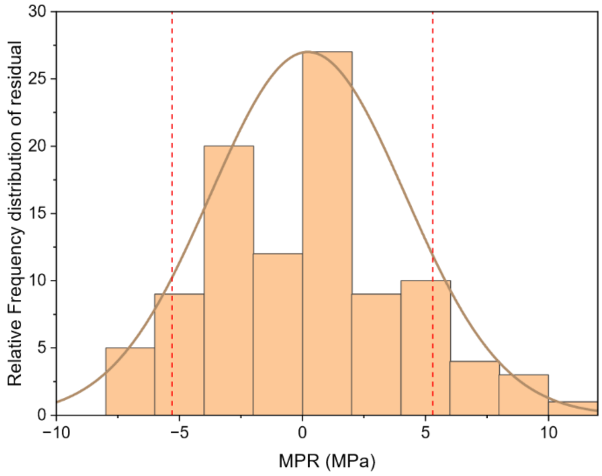

**Figure 3.** Comparing the performance of MPR, SVM, and LR for five-fold cross-validation training dataset with boxplots: RMSE (**a**), $R^2$ (**b**).

**Table 2.** Cross-validation measurement results.

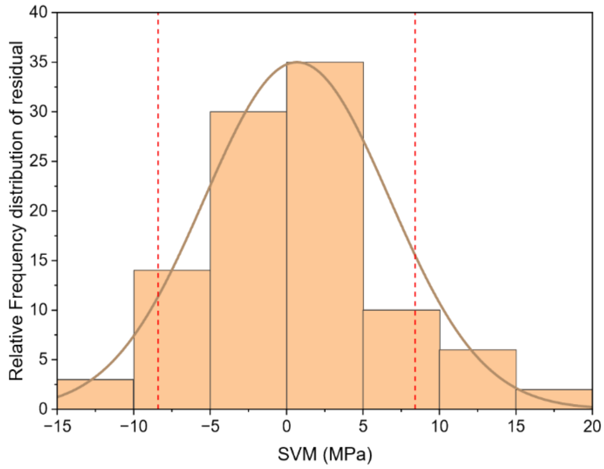
| Statistics | MPR | | SVM | | LR | |
|---|---|---|---|---|---|---|
| | ***R*2** | **RMSE** | ***R*2** | **RMSE** | ***R*2** | **RMSE** |
| Median | 0.835 | 4.958 | 0.712 | 6.202 | 0.424 | 7.7083 |
| Mean | 0.818 | 4.659 | 0.676 | 6.053 | 0.495 | 7.38138 |
| Minimum | 0.740 | 3.391 | 0.467 | 5.290 | 0.276 | 6.3806 |
| Maximum | 0.884 | 5.638 | 0.755 | 6.493 | 0.720 | 7.9816 |

The median and mean values of the two measures show that our MPR model outperforms LR and SVM. Because of the nonlinear relationships between compressive strength and the input variables, nonlinear models performed better than linear ones. In addition to evaluating the model's predictive power, more research on the stability and reliability of the model is needed. A good model's random error should have a normal distribution, with the bulk of mistakes concentrated in the centre and symmetric behaviour following N(,0). Figure 4 displays the fitted Gaussian functions and residual distributions for the three models. The residuals of three models often have a normal distribution.

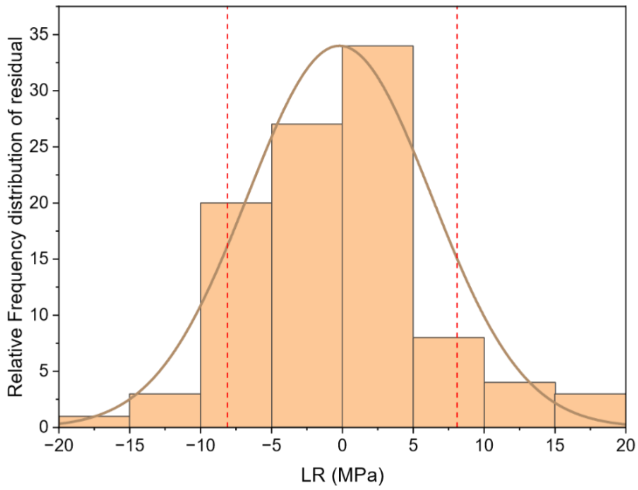
The 10th and 90th percentiles of the residual's normal distribution are shown by the dashed red lines, respectively. Based on the distribution patterns, the MPR models' 80% residuals range from [5.3, 5.3], demonstrating that they are more accurate predictors than the SVM and LR approach. Additionally, the MPR's continuity is far better than that of the SVM and LR models, as can be shown.




(a)




**(b)**




**(c)**

**Figure 4.** Normal distribution of the residual errors of five-fold cross-validation training dataset.

(**a**) MPR model (**b**) SVM model (**c**) LR model.

4.2. Model Validation

Using the training dataset, we created a prediction formula to calculate the compressive strength of eco-friendly concrete after comparing the performance of our MPR model with the results of the other two models using five-fold cross-validation. The following is the final prediction formula:

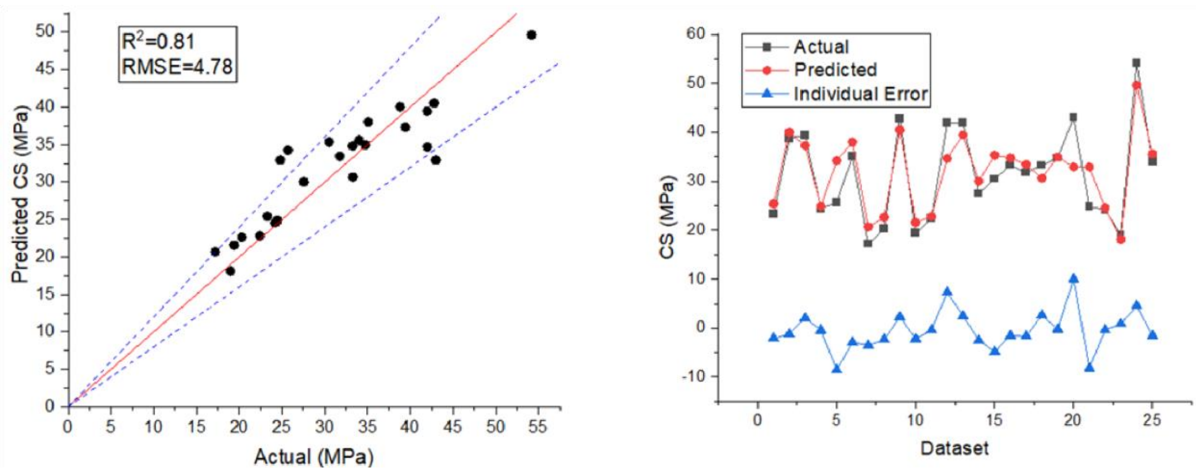
$$CS = (0.001110429806568298 \times (W/C)^2 \times (GGBFS\%)^3 \times (Age)^{-3}) + (-4267.442949761137 \times W/C \times (Age)^{-3}) + (34.78562674882411) + (-0.04850685145004892 \times (W/C)^{-1} \times RCA\%) + (7.39980447082911 \times 10^{-8} \times (W/C)^{-1} \times$$

$$^{-3} \times GGBFS\%) + (0.000005453293244435137 \times (W/C)^{-2} \times (Age)^3) + (0.21181781054751617 \times (W/C)^{-3} \times (RCA\%)$$

$(Sp)^2)$;

The testing dataset's compressive strength was estimated using the final prediction model for the compressive strength derived from the MPR model. Figure 5a compares all anticipated concrete compressive strength values to actual values for the testing data using pair plots. The projected data points using the MPR model are close to the 45th degree, as seen in Figure 4a. This demonstrates the MPR model's ability to properly generalise and forecast for samples outside of the train set. Between the observed and predicted MPR model values for the test data, the RMSE and R2 were 4.78 and 0.81, respectively.

The testing dataset's compressive strength was estimated using the final prediction model for the compressive strength derived from the MPR model. Figure 5a compares all anticipated concrete compressive strength values to actual values for the testing data using pair plots. The projected data points using the MPR model are close to the 45th degree, as seen in Figure 4a. This demonstrates the MPR model's ability to properly generalise and forecast for samples outside of the train set. Between the observed and predicted MPR model values for the test data, the RMSE and R2 were 4.78 and 0.81, respectively.




(a) (**b**)

**Figure 5.** Actual versus predicted values of concrete compressive strength using testing dataset (**a**); model errors between targets and predictions from MPR model technique (**b**).

**Table 3.** The independent variable values and the corresponding actual vs. predicted values using the MPR model.

| Point | W/C | RCA% | GGBFS% | Sp (kg) | Age (days) | Actual CS (MPa) | Predicted CS (MPa) | Individual Error (MPa) |
|---|---|---|---|---|---|---|---|---|
| 1 [28] | 0.568 | 50 | 25 | 1.15 | 7 | 23.30 | 25.41 | −2.11 |
| 2 [27] | 0.714 | 25 | 30 | 3.42 | 28 | 38.80 | 40.03 | −1.23 |
| 3 [26] | 1.250 | 0 | 60 | 0 | 90 | 39.42 | 37.32 | 2.10 |
| 4 [26] | 0.651 | 25 | 20 | 0 | 7 | 24.45 | 24.88 | −0.43 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 5 [28] | 0.852 | 50 | 50 | 1.15 | 56 | 25.70 | 34.23 | −8.53 |
| 6 [28] | 0.426 | 50 | 0 | 1.15 | 56 | 35.10 | 37.99 | −2.89 |
| 7 [27] | 1.250 | 25 | 60 | 3.42 | 7 | 17.20 | 20.68 | −3.48 |
| 8 [28] | 0.852 | 50 | 50 | 1.15 | 7 | 20.30 | 22.64 | −2.34 |
| 9 [26] | 0.833 | 0 | 40 | 0 | 90 | 42.78 | 40.51 | 2.27 |
| 10 [27] | 1.250 | 100 | 60 | 3.8 | 7 | 19.40 | 21.56 | −2.16 |
| 11 [26] | 0.868 | 25 | 40 | 0 | 7 | 22.41 | 22.81 | −0.39 |
| 12 [23] | 1.111 | 0 | 55 | 0 | 28 | 42.00 | 34.68 | 7.32 |
| 13 [24] | 0.464 | 75 | 15 | 2.28 | 28 | 41.99 | 39.45 | 2.54 |
| 14 [26] | 0.689 | 100 | 20 | 0 | 28 | 27.54 | 30.01 | −2.47 |
| 15 [25] | 3.700 | 100 | 90 | 7.8 | 28 | 30.50 | 35.32 | −4.82 |
| 16 [26] | 0.833 | 0 | 40 | 0 | 28 | 33.26 | 34.80 | −1.54 |
| 17 [26] | 0.868 | 25 | 40 | 0 | 28 | 31.75 | 33.43 | −1.68 |
| 18 [27] | 0.714 | 100 | 30 | 3.8 | 7 | 33.30 | 30.66 | 2.64 |
| 19 [26] | 0.625 | 0 | 20 | 0 | 28 | 34.76 | 34.97 | −0.21 |
| 20 [23] | 1.111 | 50 | 55 | 0 | 28 | 43.00 | 32.95 | 10.05 |
| 21 [28] | 0.852 | 50 | 50 | 1.15 | 28 | 24.80 | 32.94 | −8.14 |
| 22 [26] | 0.833 | 0 | 40 | 0 | 7 | 24.19 | 24.56 | −0.38 |
| 23 [26] | 1.327 | 50 | 60 | 0 | 7 | 19.00 | 18.10 | 0.90 |
| 24 [24] | 0.400 | 25 | 0 | 2.28 | 28 | 54.17 | 49.63 | 4.54 |
| 25 [25] | 3.700 | 100 | 90 | 7.8 | 56 | 34.00 | 35.57 | −1.57 |

### 4.3. Parametric Study

In this work, a parametric analysis was also carried out to assess how well the built-in MPR model predicted the trend in compressive strength as the input variables varied. To do this, we varied the amount of GGBFS present in the mixes while maintaining the water-to-binder ratio at 0.5. The complete amount of cement combined with the total amount of GGBFS constitutes the binder in this instance. Additionally, we altered the ratio of recycled aggregate to regular aggregate in the combinations. GGBFS was utilised in percentages of (0%, 20%, 40%, 60%, 80%), whereas recycled aggregate was used in percentages of (0%, 25%, 50%, 75%, 100%). Per one cubic metre of space, there were 175 kg of water and 2 kilogramme of superplasticizer. In both the early (7 days) and late (28 days) stages of the test (Age), the compressive strength generally increased both GGBFS and RCA. This is shown in Figure 6. As we can see, early stage increases in the GGBFS and RCA resulted in lower compressive strength values. A mixture with 20% GGBFS and 25% RCA had a compressive strength of 28.55 MPa, but a mixture with 60% GGBFS and 75% RCA had a compressive strength of 17.85 MPa. The picture also shows that the disparities in compressive strength across mixes with various RCA values decrease when the amount of GGBFS is increased in the mixture. A mixture with 80% GGBFS and 50% RCA had a compressive strength of 13.12 MPa, whereas a mixture with 80% GGBFS and 100% RCA had a compressive strength of 12.15 MPa. The behaviour of the concrete mixes was slightly different in the late stage compared to the early stage. The compressive strength of the combination remained almost constant across various GGBFS percentages at higher RCA % levels. For instance, the compressive strength of mixes containing 75% RCA and 20% GGBFS was 32.62 MPa, whereas the compressive strength of mixtures containing 75% RCA and 60% GGBFS was 32.15 MPa.

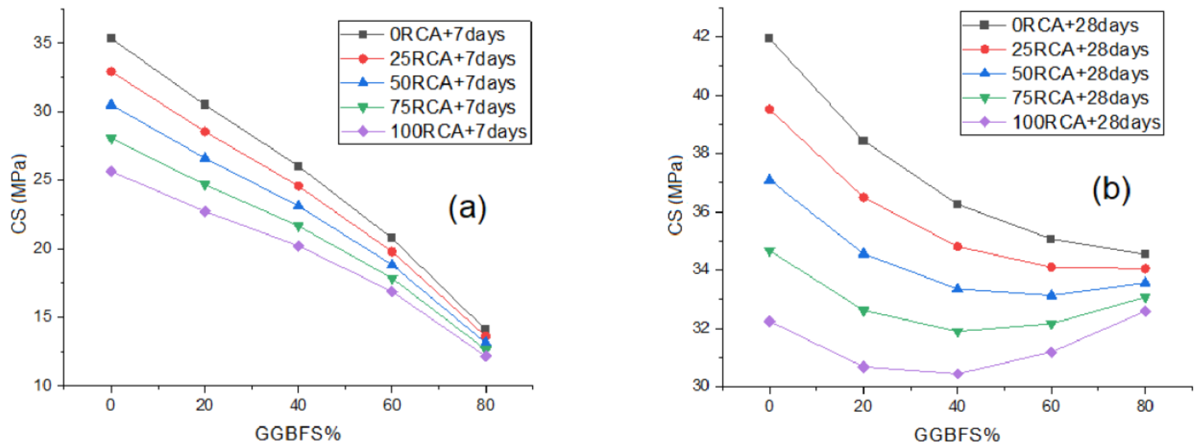



**Figure 6.** The effect of GGBFS and RAC on compressive strength at early stage (**a**), effect on GGBFS and RAC on compressive strength at late stage (**b**).

4.4. Sensitivity Study

This analysis's objective was to determine how input variables impact CS predictions. Figure 7 displays the impact of each input parameter on CS prediction. According to this study, Sp was the most significant factor, making up 38.77% of the total, followed by Age, which made up 37.04%. While W/C ratio accounted for 19.01 percent of the total input variables, RCA contributed 5.03 percent, and GGBFS contributed 0.45 percent, the remaining input factors had a less significant impact in the prediction of CS. The sensitivity value (percent) of the dependent variable with regard to each independent variable was determined using equations (6) and (7) [34].

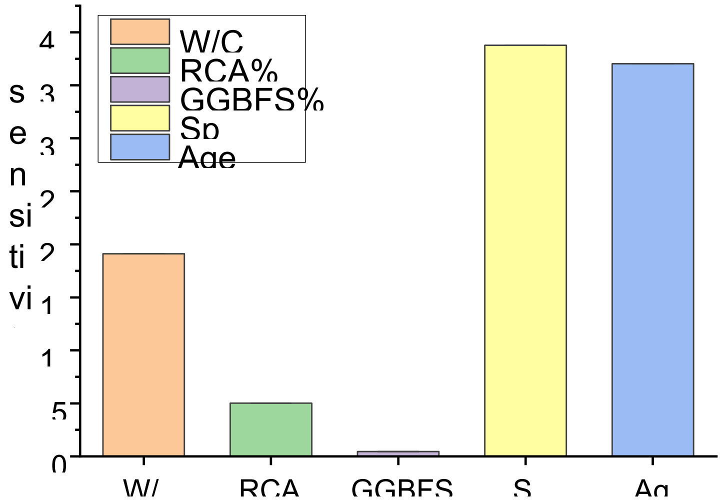



**Figure 7.** Relative importance input variables in sensitivity analysis with MPR.

## 6. Conclusions

This research centred on the notion of developing prediction equations for the compressive strength of eco-friendly concrete using a white-box machine learning technique, such as MPR. To train and validate the MPR machine learning system, 125 recordings of concrete mix specimens were acquired as data. The predictors from the literature that were used to build the models (days) were the water-to-cement ratio (W/C), recycled aggregate percentage as a replacement for normal aggregate in the mixture (RAC percent), GGBFS percentage as a replacement for OPC in the binder (GGBFS percent), superplasticizer, and age. The outcome variable was the compressive strength (CS) of green concrete. The MPR model performed better than the LR (white-box model) and SVM (black-box model) when tested using the cross-validation approach in terms of R2 and RMSE. In the five-fold cross-validation, the MPR model's average prediction performances were 0.818 for R2 and 4.659 for RMSE, whereas the SVM model's average prediction performances were 0.676 for R2 and 6.053 for RMSE. Out of all the models, LR performed the worst. The testing dataset was used to validate the MPR model formula that was produced from the training dataset. Using the MPR model, the predicted values of the compressive strength of environmentally friendly concrete for the testing dataset were quite similar to the experimental values. Additionally, a parametric analysis was carried out to investigate the impact. According to the study, the compressive strength of concrete with a high percentage of GGBFS and a low cement content is comparable to concrete with a high content of cement and a low percentage of GGBFS with a larger proportion of RCA in the mixture at the late stage.

## References

1. Hammoudi, A.; Moussaceb, K.; Belebchouche, C.; Dahmoune, F. Comparison of artificial neural network (ANN) and response surface methodology (RSM) prediction in compressive strength of recycled concrete aggregates. *Constr. Build. Mater.* **2019**, *209*, 425–436. [CrossRef]

2. Zhang, Q.; Feng, X.; Chen, X.; Lu, K. Mix design for recycled aggregate pervious concrete based on response surface methodology. *Constr. Build. Mater.* **2020**, *259*, 119776. [CrossRef]
3. Younis, K.H.; Jirjees, F.F.; Khoshnaw, G.; Ali, B.H. Experimental study on performance of recycled aggregate concrete: Effect of reactive mineral admixtures. *Int. J. Civ. Eng. Technol.* **2019**, *10*, 2566–2576.
4. Tiwari, P.K.; Sharma, P.; Sharma, N.; Verma, M. An experimental investigation on metakaoline GGBS based concrete with recycled coarse aggregate. *Mater. Today Proc.* **2021**, *43*, 1025–1030. [CrossRef]
5. Nedunuri, S.S.S.A.; Sertse, S.G.; Muhammad, S. Microstructural study of Portland cement partially replaced with fly ash, ground granulated blast furnace slag and silica fume as determined by pozzolanic activity. *Constr. Build. Mater.* **2020**, *238*, 117561. [CrossRef]
6. Chore, H.S.; Joshi, M.P. Strength characterization of concrete using industrial waste as cement replacing materials for rigid pavement. *Innov. Infrastruct. Solut.* **2020**, *5*, 1–9. [CrossRef]
7. Aydın, S.; Baradan, B. Sulfate resistance of alkali-activated slag and Portland cement based reactive powder concrete. *J. Build. Eng.* **2021**, *43*, 103205. [CrossRef]
8. Karthik, S.; Mohan, K.S.R. A Taguchi Approach for Optimizing Design Mixture of Geopolymer Concrete Incorporating Fly Ash, Ground Granulated Blast Furnace Slag and Silica Fume. *Crystals* **2021**, *11*, 1279. [CrossRef]
9. Ahmad, S.; Kumar, A.; Kumar, K. Axial performance of GGBFS concrete filled steel tubes. In Proceedings of the Structures, St. Louis, MO, USA, 5–8 April 2020; pp. 539–550.
10. Al-Abdaly, N.M.; Al-Taai, S.R.; Imran, H.; Ibrahim, M. Development of prediction model of steel fiber-reinforced concrete compressive strength using random forest algorithm combined with hyperparameter tuning and k-fold cross-validation. *East.Eur. J. Enterp. Technol.* **2021**, *5*, 113. [CrossRef]
11. Su, S.; Zhi, J.; Lou, L.; Huang, F.; Chen, X.; Wu, J. Spatio-temporal patterns and source apportionment of pollution in Qiantang River (China) using neural-based modeling and multivariate statistical techniques. *Phys. Chem. Earth Parts A/B/C* **2011**, *36*, 379–386. [CrossRef]
12. DeRousseau, M.; Laftchiev, E.; Kasprzyk, J.; Rajagopalan, B.; Srubar, W., III. A comparison of machine learning methods for predicting the compressive strength of field-placed concrete. *Constr. Build. Mater.* **2019**, *228*, 116661. [CrossRef]
13. Arshian, A.H.; Morgenthal, G. Probabilistic assessment of the ultimate load-bearing capacity in laterally restrained two-way reinforced concrete slabs. *Eng. Struct.* **2017**, *150*, 52–63. [CrossRef]
14. Ouyang, B.; Song, Y.; Li, Y.; Wu, F.; Yu, H.; Wang, Y.; Yin, Z.; Luo, X.; Sant, G.; Bauchy, M. Using machine learning to predict concrete's strength: Learning from small datasets. *Eng. Res. Express* **2021**, *3*, 015022. [CrossRef]
15. Golafshani, E.M.; Behnood, A.; Hosseinikebria, S.S.; Arashpour, M. Novel metaheuristic-based type-2 fuzzy inference system for predicting the compressive strength of recycled aggregate concrete. *J. Clean. Prod.* **2021**, *320*, 128771. [CrossRef]

16. 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]
17. Nunez, I.; Marani, A.; Nehdi, M.L. Mixture Optimization of Recycled Aggregate Concrete Using Hybrid Machine Learning Model. *Materials* **2020**, *13*, 4331. [CrossRef] [PubMed]
18. Topçu, I.B.; Sarıdemir, M. Prediction of mechanical properties of recycled aggregate concretes containing silica fume using˙ artificial neural networks and fuzzy logic. *Comput. Mater. Sci.* **2008**, *42*, 74–82. [CrossRef]
19. 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]
20. Mohana, M.H. The determination of ground granulated concrete compressive strength based machine learning models. *Period. Eng. Nat. Sci. (PEN)* **2020**, *8*, 1011–1023.
21. Mai, H.-V.T.; Nguyen, T.-A.; Ly, H.-B.; Tran, V.Q. Prediction Compressive Strength of Concrete Containing GGBFS using Random Forest Model. *Adv. Civ. Eng.* **2021**, *2021*. [CrossRef]
22. Sarıdemir, M.; Topçu, I.B.; Özcan, F.; Severcan, M.H. Prediction of long-term effects of GGBFS on compressive strength of concrete˙ by artificial neural networks and fuzzy logic. *Constr. Build. Mater.* **2009**, *23*, 1279–1286. [CrossRef]
23. Kou, S.-C.; Poon, C.-S.; Agrela, F. Comparisons of natural and recycled aggregate concretes prepared with the addition of different mineral admixtures. *Cem. Concr. Compos.* **2011**, *33*, 788–795. [CrossRef]
24. Moghadam, A.S.; Omidinasab, F.; Goodarzi, S.M. Characterization of concrete containing RCA and GGBFS: Mechanical, microstructural and environmental properties. *Constr. Build. Mater.* **2021**, *289*, 123134. [CrossRef]
25. Elchalakani, M.; Basarir, H.; Karrech, A. Green concrete with high-volume fly ash and slag with recycled aggregate and recycled water to build future sustainable cities. *J. Mater. Civ. Eng.* **2017**, *29*, 04016219. [CrossRef]
26. Majhi, R.; Nayak, A.; Mukharjee, B. Development of sustainable concrete using recycled coarse aggregate and ground granulated blast furnace slag. *Constr. Build. Mater.* **2018**, *159*, 417–430. [CrossRef]
27. Tüfekçi, M.M.; Çakır, Ö. An investigation on mechanical and physical properties of recycled coarse aggregate (RCA) concrete with GGBFS. *Int. J. Civ. Eng.* **2017**, *15*, 549–563. [CrossRef]
28. Subarkah, M.; Sjah, J.; Maknun, I. Effects of Ground Granulated Blast Furnace Slag and Recycled Coarse Aggregates in Compressive Strength of Concrete. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Changchun, China, 21–23 August 2020; p. 012045.
29. Wei, J.; Chen, T.; Liu, G.; Yang, J. Higher-order multivariable polynomial regression to estimate human affective states. *Sci. Rep.* **2016**, *6*, 1–13.
30. Vaccari, D.A. TaylorFit Response Surface Analysis- with Stepwise Multivariate Polynomial Regression. Available online: http://www.taylorfit-rsa.com/ (accessed on 15 August 2018).

31. Ling, H.; Qian, C.; Kang, W.; Liang, C.; Chen, H. Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment. *Constr. Build. Mater.* **2019**, *206*, 355–363. [CrossRef]
32. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 112.
33. Yang, K.; Xu, X.; Yang, B.; Cook, B.; Ramos, H.; Krishnan, N.A.; Smedskjaer, M.M.; Hoover, C.; Bauchy, M. Predicting the Young's modulus of silicate glasses using high-throughput molecular dynamics simulations and machine learning. *Sci. Rep.* **2019**, *9*, 1–11. 34. Mirzahosseini, M.; Najjar, Y.; Alavi, A.H.; Gandomi, A.H. ANN-Based prediction model for rutting propensity of asphalt mixtures. In Proceedings of the 92nd Annual Meeting of Transportation Research Board, Washington, DC, USA, 13–17 January 2013.