



ARTICLE

Advanced Machine Learning for Sustainable Concrete Strength Prediction and Resource Optimization

Nayemuddin Mohammed^{1,2}, Tahar Ayadat^{1,2,*}, Andi Asiz^{1,2} and Nadeem Pasha³

¹Department of Civil Engineering, Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia

²Centre for Sustainable Infrastructure Materials (SIM), Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia

³Department of Civil Engineering, Khaja Bandanawaz University, Kalaburagi, Karnataka, India

*Corresponding Author: Tahar Ayadat. Email: tayadat@pmu.edu.sa

Received: 10 February 2026; Accepted: 17 April 2026; Published: 30 June 2026

ABSTRACT: Significant efforts have been made to increase the strength of concrete by using industrial waste such as fly ash and steel slag as partial substitutes for concrete in concrete. However, predicting the concrete's compressive strength is a challenge as it is influenced by several factors such as the shape and size of the aggregate, the water-ratio balance. This study examines the predictive capability of three deep learning models: Bagging Extreme Gradient Boosted Model (BXGBM), Deep Random Vector Functional Link (DRVFL), and Kernel Extreme Learning Machine (KELM) on the prediction for compressive strength of concrete. The dataset was split into a training and testing set, and the performance measures were analyzed. The statistical metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and coefficient of determination (R^2). According to model BXGBM, the MSE was found to be 3.757, and the R^2 of 0.9864 was as in testing, with MSE of 16.63, $R^2 = 0.941$ performed well with good accuracy. The model DRFVL has a training MSE of 29.858, $R^2 = 0.8922$, and has low overall generalizability of 46.030 and 0.8391 on the testing set. KELM also did well with a training MSE of 13.851, R^2 of 0.950; testing performance declined with an MSE of 31.05 and R^2 of 0.891. The results show that BXGBM is most trustworthy as a model that predicts the compressive strength of concrete, which allows emphasizing its high potential in applying to the practical sphere of concrete technology.

KEYWORDS: Concrete; strength; sustainable; prediction; regression; performance; optimization

1 Introduction

Concrete is a highly popular building material that is used due to its diversity and strength. The strength of concrete is one of the most crucial parameters that determine the lifespan and safety of buildings, so it is vital to focus on its correct prediction in the sphere of civil engineering [1]. Historically, compressive strength of concrete has been determined empirically by the use of empirical models and standardized tests, which may be time-consuming and expensive. With the increasing demand for more efficient and reliable methods, machine learning (ML) has appeared as a prospective substitute with respect to the traditional approach [2]. Paudel et al. investigate the machine learning algorithms, especially different regression models, to identify the compressive strength of concrete with fly ash, and conclude that XGBoost Regressor is better than the other regression models. It emphasizes the significance of such input parameters as concrete age, cement, and water content in increasing prediction quality [3]. Asteris et al. compare a range of machine learning models, such as support vector machines, random forests, etc., to predict compressive strength of cement-based mortars, which have limitations identified in artificial neural networks. It highlights the opportunity

of AdaBoost and RF models in streamlining the design of mortar by examining how parameters like cement grade, age, and water-to-binder ratio relate to compressive strength [4]. Gucluer et al. conduct a comparative evaluation of the performance of different machine learning algorithms, Artificial Neural Networks, Decision Trees, Support Vector Machines, and Linear Regression to predict the compressive strength of concrete in samples at 7 and 28 days of curing time. The Decision Tree algorithm has the most significant correlation coefficient ($R^2 = 0.86$), and the least mean absolute error (2.59), which made the method the most effective in estimating compressive strength in terms of input parameters such as unit weight and water content [5]. Zhang et al. offer an effective machine learning algorithm to forecast the compressive strength of cement-stabilized soft soil, which would deal with the inefficiencies of conventional geotechnical tests. The Extreme Gradient Boosting model was able to predict the cement content, water content, curing age, and fine grain as the most significant influencing factors of compressive strength with a determination coefficient of 0.93, which is a solid reference to the cement soil design in soft foundation works using a dataset of 566 samples [6].

Hadzima-Nyarko et al. analyze the compressive strength of rubberized concrete based on a dataset of 457 samples, using four machine learning models, namely Artificial Neural Network (ANN), k-nearest neighbor (KNN), regression trees (RT), and random forests (RF). Findings showed that the RT model was best in training, but the ANN model actually gave the best predictions in testing when compared to the RT and RF, as well as the traditional expressions. The results indicate that the ANN model can be used to predict the compressive strength of rubberized concrete, providing useful information on the engineering management and safety [7]. Chopra et al. apply machine learning procedures in R in forecasting 28, 56, and 91 days of concrete compressive strength, the application of decision tree, random forest, and neural network models. As observed in the analysis, the neural network model offers the best predictions according to R^2 and RMSE. This proves the neural network to be a useful tool in estimating the concrete strength under controlled conditions [8]. In Feng et al., the approach to machine learning is based on the adaptive boosting algorithm that is used to forecast concrete compressive strength with the improvement of multiple weak learners. The model has a 1030-test results dataset, with 10-fold cross-validation accuracy of more than 95 percent, and better performance than a single method, such as artificial neural networks and support vector machines. The paper also examines the effect of data size of training, selection of the weak learner, and the sensitivity to the input parameter, and makes a conclusion that decision trees are weak learners that are best in this area [9]. Table 1 addresses the recent studies of machine learning applications applied to the strength of concrete.

Table 1: Recent studies of machine learning applications for predicting the strength of concrete.

Study	Structural Element	Behavior	ML Used	Key Findings	Ref.
Adaptive Boosting-Based Concrete Strength Prediction	Concrete	Compressive strength	(AdaBoost), compared with ANN and SVM	Achieved >95% accuracy (R^2) using 10-fold cross-validation.	[9]
Fly Ash-Based Concrete Strength Prediction	Concrete with fly ash	Compressive strength	MLR, SVR, AdaBoost, RF, XGBoost, Bagging	XGBoost showed the best performance ($R^2 = 0.95$, lowest RMSE, MAE).	[3]

(Continued)

Table 1 (continued)

Study	Structural Element	Behavior	ML Used	Key Findings	Ref.
Comparative ML Study for Hardened Concrete Strength	Hardened concrete	Compressive strength (7 and 28 days)	ANN, Decision Tree (DT), SVM, Linear Regression (LR)	DT model ($R^2 = 0.86$, MAE = 2.59). 28-day curing data	[5]
Rubberized Concrete Strength	Rubberized concrete	Compressive strength	ANN, KNN, RT, Random Forest (RF)	ANN performed well in the training and testing phase.	[7]

According to recent research, the American Concrete Institute (ACI 318-19) code ignores such important factors as shear span ratio (a/d) and steel yield strength (F_y), but gene expression programming (GEP) models trained on extensive experimental datasets provide more accurate predictions of the SRCB-WS shear strength. Nevertheless, data-driven techniques are dependent on the quality of the data and are not as simple as a code formula, and thus a combination of ML and conventional design regulations is required in order to estimate shear-strength practically and reliably [10]. Analytical and experimental research demonstrates that FRP reinforced beam shear strength is significantly dependent on concrete strength, shear span/depth ratio, longitudinal ratio, and transverse FRP stirrups. Recent papers use data-driven models (M5, RF, ELM, GEP) to enhance vs. prediction, with tree-based and ensemble approaches tending to trade accuracy and interpretability, whereas neural/ELM approaches need adjustment [11].

Abuodeh et al. explore the compressive strength of Ultra-High-Performance Concrete (UHPC) with the help of Artificial Neural Networks to solve its black-box nature in terms of Sequential Feature Selection (SFS) and Neural Interpretation Diagram (NID) to determine the material constituents that were important. A database of 110 UHPC tests indicated that four major constituents, cement, fly ash, silica fume, and water, could be used to achieve a higher degree of prediction (91% and a normalized mean square error (NMSE) of 0.012), in contrast to the prediction with all eight constituents. The research finds that a combination of SFS and NID is a highly effective method to enhance the accuracy of models and provide information on UHPC mix forecasting [12]. Li et al. perform a literature search of 3135 journal articles on the topic of concrete compressive strength prediction published between 2012 and 2021, and utilize Cite Space in mapping research trends and hotspots. It divides the techniques of prediction into more traditional and machine-learning ones, suggesting a gradient boosting regression tree (GBRT) algorithm to improve predictions. The GBRT model was best and therefore had the highest accuracy in comparison to other machine-learning models, as evidenced by its coefficient of determination R^2 of 0.92, mean square error (MSE) of 22.09 MPa, and root mean square error (RMSE) of 4.7 MPa using a dataset of 1030 samples. A five-fold cross-validation and understanding of the significance of eight input variables were also done in the study [13].

This paper is dedicated to applying three modern advanced machine learning methods that are Deep Random Vector Functional Link (DRVFL), Bagging Extreme Gradient Boosted Model (BXGBM), and Kernel Extreme Learning Machine (KELM) to forecast the strength of concrete depending on its components and curing environment. DRVFL applies random projections to increase the efficiency of learning, whereas BXGBM is a combination of bagging and boosting methods to increase accuracy. KELM uses the functionality of a kernel that has the ability to capture nonlinear relations in data. In this way, comparing the predictive powers of these models using measures of Mean Squared Error (MSE), Root

Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination R^2 statistical metrics. The results of this study provide some practical suggestions to any engineer who wants to maximize concrete mixes to achieve better production. Finally, we hope to determine which model will give the best predictions of concrete strength, thus developing better decision-making in concrete production and concrete construction practice.

2 Research Gap and Novelty

The research paper presents major gaps in the research on predictive modeling of compressive strength of concrete, especially the use of modern machine learning tools. Although conventional models have been extensively exploited, there is no research on the deep learning model, such as Deep Random Vector Functional Link (DRVFL), Bagging Extreme Gradient Boosted Model (BXGBM), and Kernel Extreme Learning Machine (KELM), to this end. The results show that despite a high level of training performance in DRVFL, the training performance is affected by overfitting, which means that better generalizing models are required. Conversely, BXGBM turned out to be the most predictable model with high accuracy and external validity in concrete strength prediction. This shows the newness of the application of ensemble and deep learning methods, and this aspect adds to a more solid comprehension of compressive strength prediction in concrete technology. The research highlights the feasibility of BXGBM in practice, and it is hoped that future research can pay attention to the feasibility of increasing the applicability of the model and other innovative machine learning algorithms in concrete performance prediction.

3 Methodology and Materials

The dataset is taken from the Kaggle site [14]. The input for the design model is cement m^3 , blast furnace slag kg/m^3 , fly ash, water kg/m^3 , superplasticizer kg/m^3 , coarse aggregate kg/m^3 , fine aggregate kg/m^3 , curing age in days, and the output is compressive strength of concrete in MPa. The Total number of datasets used are 1030, observations among from these 824 data 80% were used for training, and the remaining 206 data points 20% for testing were employed in the design model [15]. Fig. 1 displays the network architecture design model for a machine learning algorithm.

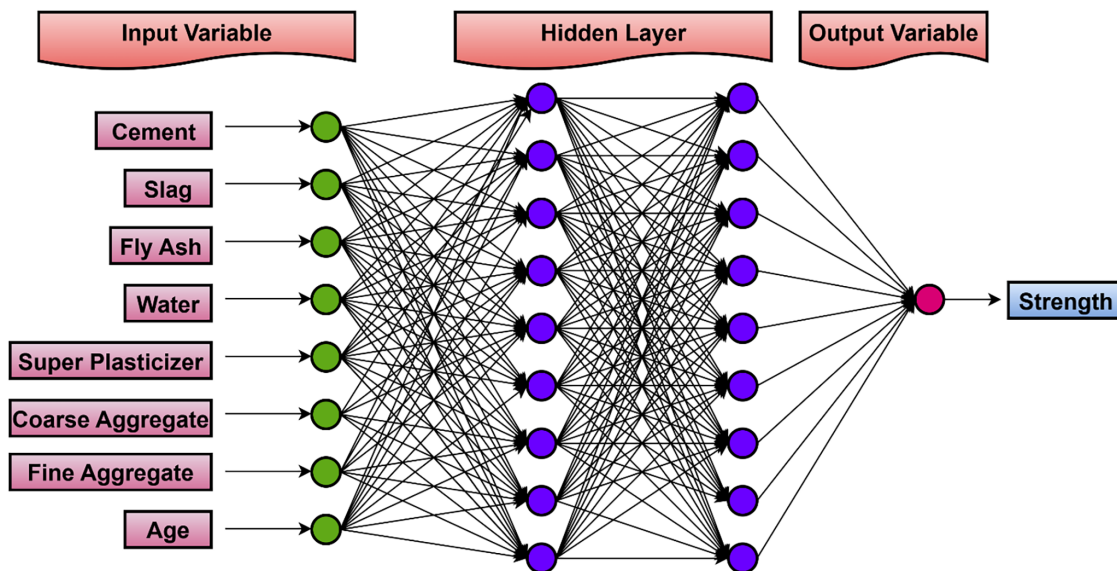


Figure 1: Network architecture for the machine learning model.

Fig. 2 illustrates the normal distribution histogram of different components and characteristics of the concrete with a fitted normal distribution curve shown alongside each one, indicating the statistical characteristics mean (μ) and standard deviation (SD). The cement distribution exhibits a bell-shaped curve, and the mean of the shape is equal to about 281.17, and the standard deviation of the shape is equal to 104.50, indicating a fairly normal distribution with the value concentrating around the mean. Conversely, slag distribution shows a right-skewed distribution with a mean of 73.90 and standard deviation of 86.27, meaning that a good portion of occurrences have low or zero values of slag. The distribution of fly ash is also skewed to the right with a mean of 54.18 and a standard deviation of 63.99, indicating a difference in the use of fly ash among mixtures of concrete. Water content demonstrates a more balanced distribution at a mean of 181.57 and a standard deviation of 21.35, which implies uniformity in the water usage in concrete mixes. On the other hand, the distribution of the superplasticizer is highly skewed to the right, with the central tendency of the distribution being 6.20 and the area of distribution being 5.97, which indicates that the majority of the samples contain a low level of superplasticizer [16].

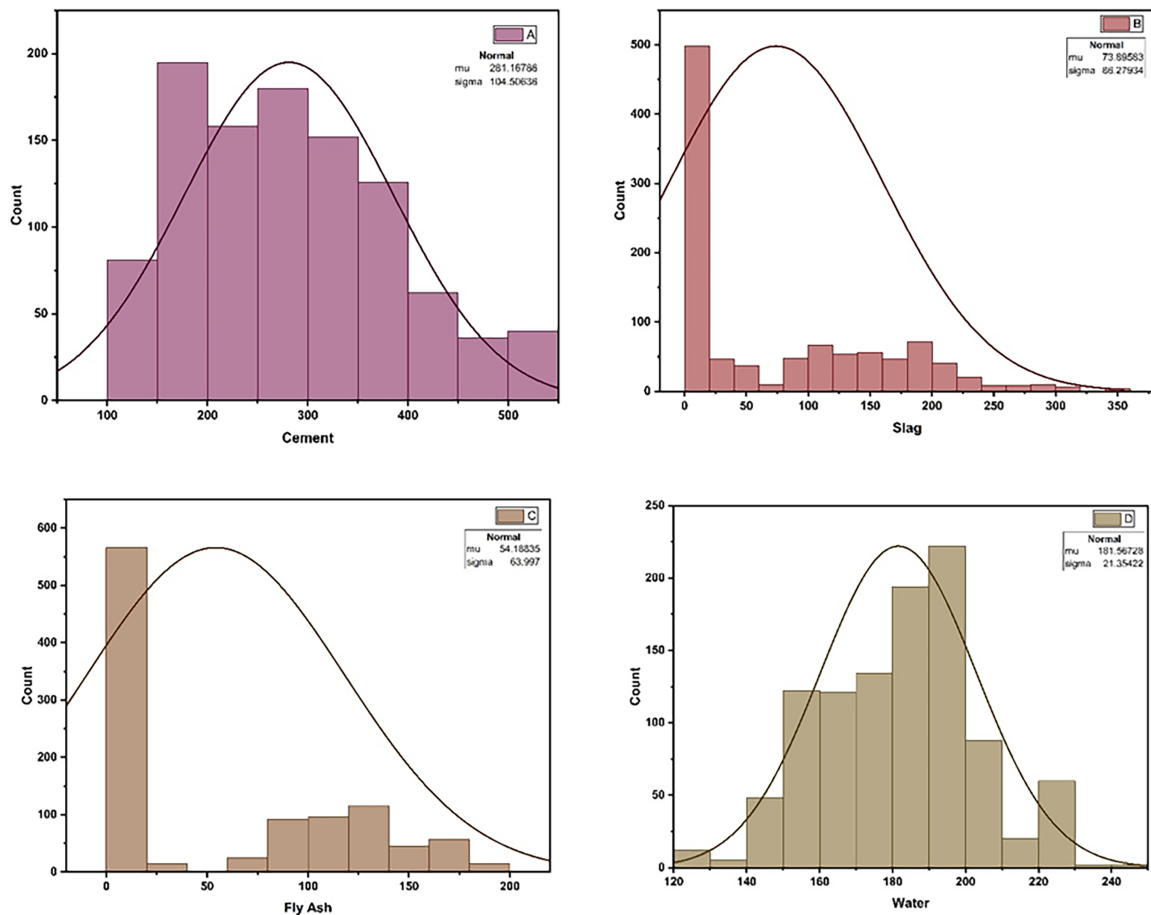


Figure 2: (Continued)

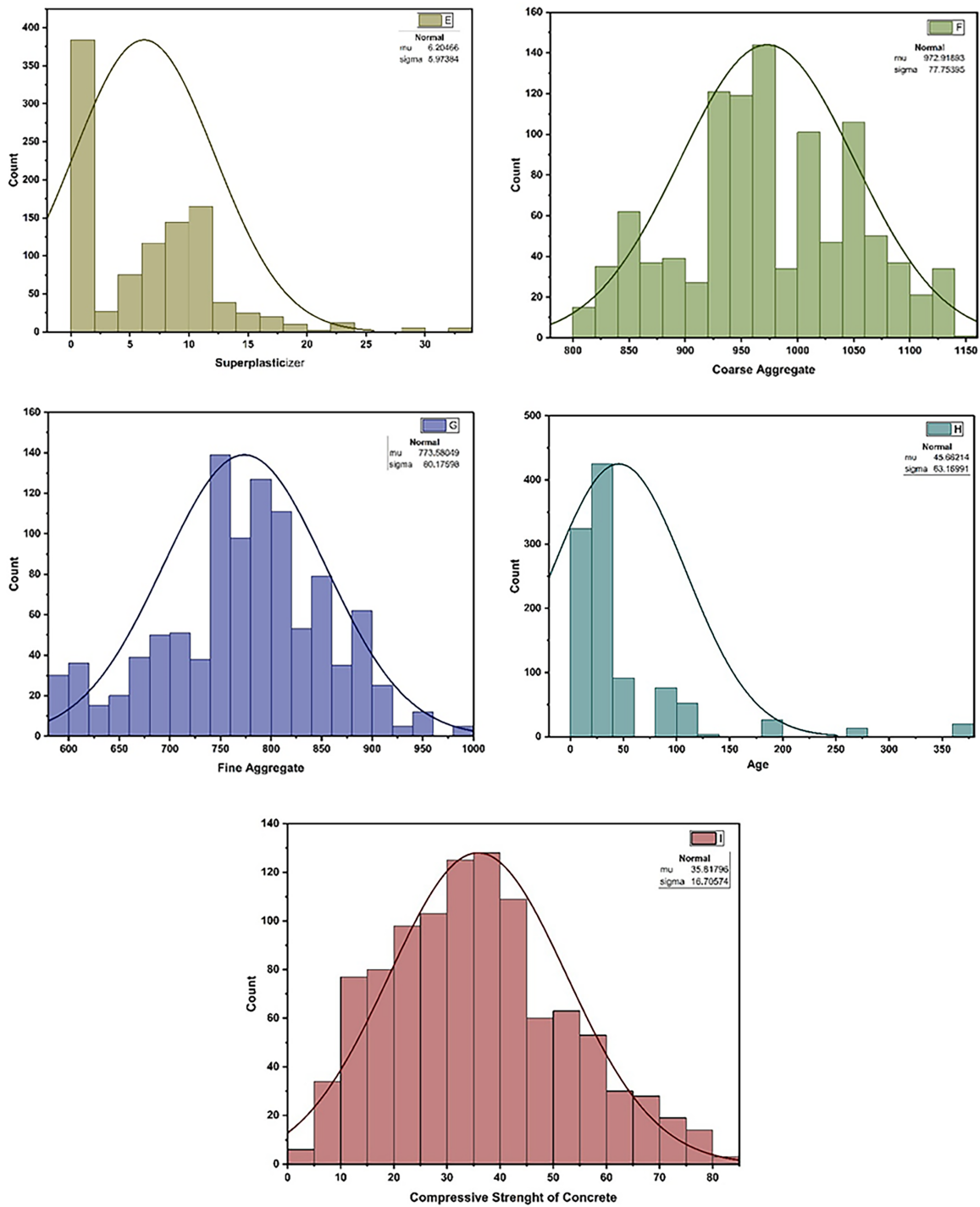


Figure 2: Normal distribution for the input and output variables.

Fig. 3 shows a heatmap visualizing the correlation matrix of different variables. The correlation coefficient is -1 to 1 , which represents the correlation between two variables, such as the inputs, which are concrete production, including cement, slag, ash, water, superplasticizer, coarse aggregate, fine aggregate in kg/m^3 mixture, age in days, and compressive strength in MPa. Compressive strength was strongly and

positively correlated with input cement, with a moderate $r = +0.33$ higher cement content in the tested input cement in the range was more likely to augment load-bearing capacity at an early age. Compressive strength shows the best positive relationship with age, $r = +0.45$, which identifies the significance of curing time and progressive hydration in reading the results of the strength. Conversely, water exhibits a negative correlation with strength, -0.21 , indicating that increased water content may weaken the concrete.

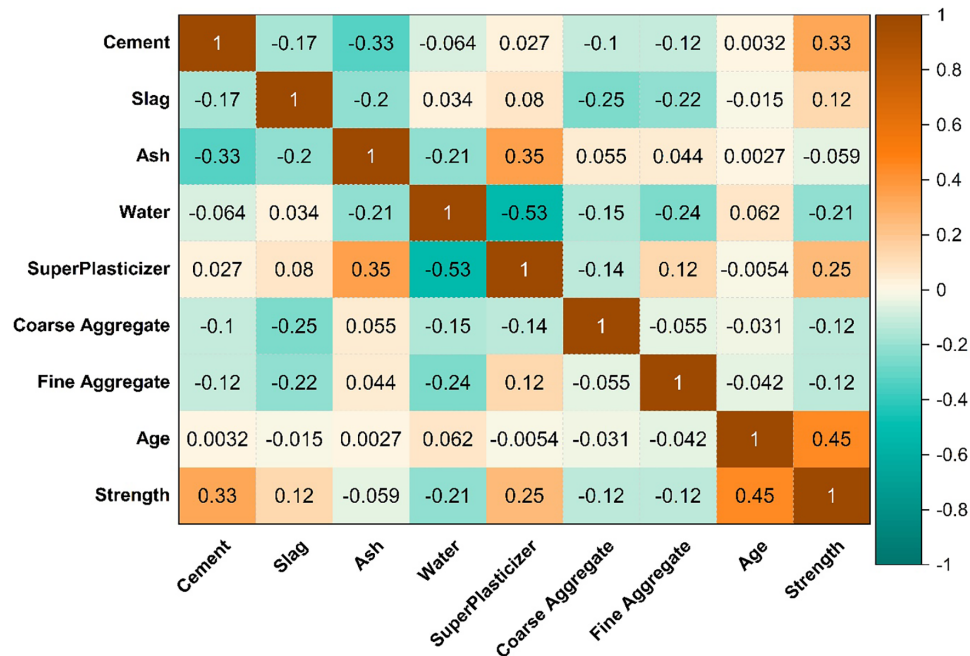


Figure 3: Correlation matrix for the input and output variables.

Other significant associations are the positive association between age and strength, $+0.25$, which implies that concrete keeps on acquiring strength with time. Rough and fine aggregate materials have weak, largely negative relationships to strength, less than 0.15, as suggested by the composition ranges examined. The paste matrix (cementitious content, water, admixture, and age) was the leading predictor of compressive strength. The heatmap aids in understanding the relationships between these variables, which can inform decisions about concrete mix design for optimal performance. This heatmap is very effective in demonstrating the correlation between the variables, and this can offer information that can be used in subsequent analysis or modeling.

4 Results and Discussion

Fig. 4 displays the scatterplot matrix, which shows the connections between the different components present in the production of concrete, including cement, slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, age, and compressive strength of concrete are depicted. In addition, it indicates the relationship between pairs of variables, and the plots against the variables show their distributions. Boxplots on the diagonal show that cement, coarse aggregate, fine aggregate, and strength have fairly broad value distributions and skewness, whereas superplasticizer and some SCMs have more limited and concentrated distributions, indicating variances of different scales and the need to normalize them before multivariate modeling. Scatter panels indicate a positive relation, though scattered, between the cement content and the strength due to the role of cement in the paste matrix, and a more apparent positive relation between age and strength due to the continued hydration and strength increase.

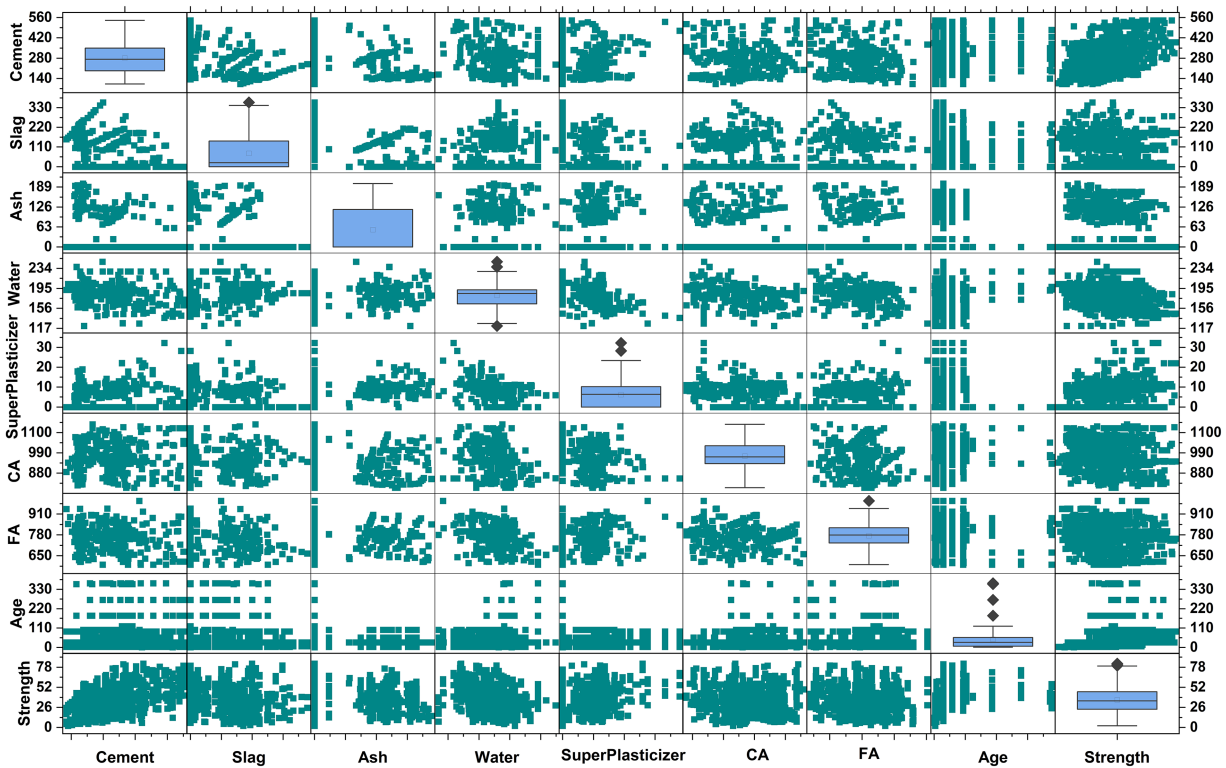


Figure 4: Scatter plot for the input and output parameters.

Water exhibits a negative relationship with strength throughout its range, which confirms the anticipated deleterious impact of a higher water-to-binder ratio on compressive performance, whereas superplasticizer has a more dispersed pattern with numerous low-water, higher-admixture points that concur with higher strength in line with its purpose of preserving workability at the cost of making it possible to use less water. There are also mixed interactions between SCMs and strength: slag-related scattered clouds indicate that there are minimal positive tendencies, whereas fly ash is more distributed and may experience early age strength dilution at the higher replacement levels. Aggregate variables generate thicker and lower-gradient scatter clouds as compared to strength, implying the lack of direct impact in the observed composition ranges; however, the existence of clustered vertical bands in multiple columns is an indication of repeated or discrete dosage ranges in the dataset and thus should be controlled, such as through categorical encoding or through cautious cross-validation to prevent bias. This matrix is a good representation of how components are related to each other and can be used in understanding concrete mix design optimization.

4.1 Machine Learning Techniques

4.1.1 Deep Random Vector Functional Link (DRVFL)

Deep Random Vector Functional Link is a type of neural network that can use random feature mapping in order to map input data to a high-dimensional space. DRVFL uses random projections to retain the key characteristics of the input and makes training of traditional deep networks much simpler. This leads to quicker convergence, hence DRVFL is efficient with real-time applications. Also, the model is not easily overfit, especially when using smaller datasets [17]. Fig. 5 displays the representation of a deep random vector functional link. A Deep random vector functional link (DRVFL) network architecture starts with the input vector that reflects the features of the processed data. This input vector is converted by a random projection,

using a randomly initialized weight matrix. This dimensionality reduction brings randomness to the model, with the hope that it will learn complicated relations in the data because the activation function is nonlinear. Conventional deep learning algorithms are based on gradient descent learning, which has the demerits such as improper learning rate, overfitting issues, and tuning of weights and biases. To prevent these challenges, the DRVFL feed forward deep learning models were used to categorize the weights where the weights, which are randomly initialized, and give a better result at a faster speed [18].

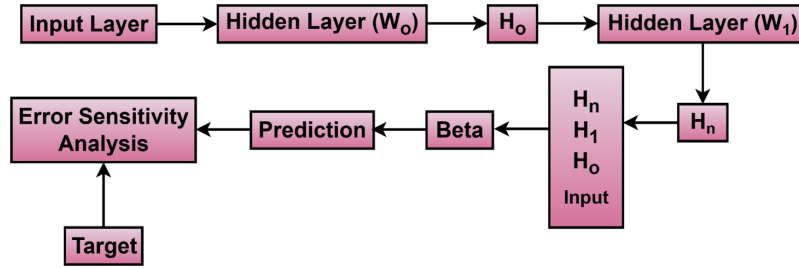


Figure 5: Representation of steps involved in the DRVFL machine learning model.

The output layer is identified on the basis of the activations, which give the predicted output, indicating the estimates of the network on the input data. In order to check the performance of the model, a loss function is employed, and it measures the difference between the actual and the predicted outputs. This role evaluates the performance of the model, and it acts as a guide on the optimization procedure when training, where the aim of minimizing the error in prediction is going to be considered [19]. Eqs. (1) to (5) were employed for the DRVFL machine learning model.

$$\text{Input Vector } \{x\} = [x_1, x_2, \dots, x_n]^T \quad (1)$$

$$\text{Random Projection } \{z\} = \{W\} \cdot \{x\} \quad (2)$$

where, W is the random weight matrix of size $m \times n$, m is the number of hidden units and n is the input dimension.

$$\text{Activation Function } \{a\} = f \{z\} \quad (3)$$

f is the nonlinear activation function sigmoid, ReLU to the projected vector.

$$\text{Output Layer } \hat{y} = \{V\} \cdot \{a\} \quad (4)$$

V is the weight matrix connecting the hidden units to the output.

$$\text{Loss Function, } L = \frac{1}{N \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (5)$$

N is the number of training samples, y_i is the true output value and \hat{y}_i is the predicted output value [20].

4.1.2 Bagging Extreme Gradient Boosted Model (BXGBM)

BXGBM, which stands for Bagging Extreme Gradient Boosted Model combines the two ideas of bagging and extreme gradient boosting (XGBoost), which is an ensemble level learning framework that improves predictive ability. BXGBM can be used to enhance the accuracy and strength by synthesizing many models

and summing the results. XGBoost is especially known to be efficient and scalable, which is why BXGBM is a potent method of approaching a complex regression and classification [21]. Fig. 6 shows the steps involved in the design model of the Bagging extreme gradient boosted model. Eqs. (6) to (8) are utilized for the prediction of the BXGBM machine learning model.

$$\text{Individual XGBoost Prediction, } y_i^{(b)} = \sum_{k=1}^K f_k^{(b)}(x_i) \quad (6)$$

$$\text{Bagged Ensemble Prediction, } \widehat{y}_i^{\text{BXGB}} = \frac{1}{B} \sum_{b=1}^B \widehat{y}_i^{(b)} \quad (7)$$

$$\text{Combined Formulation, } \widehat{y}_i^{\text{BXGB}} = \frac{1}{B} \sum_{b=1}^B \left(\sum_{k=1}^K f_k^{(b)}(x_i) \right) \quad (8)$$

where, B = number of bagged XGBoost models, K = number of trees in each XGBoost model, $f_k^{(b)}$ = k^{th} tree in the b^{th} XGBoost model, x_i = input sample and $\widehat{y}_i^{\text{BXGB}}$ = final averaged prediction.

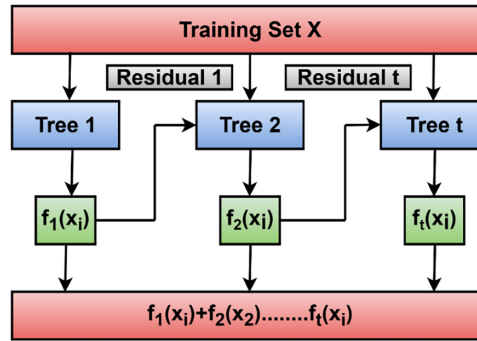


Figure 6: Representation of steps involved in the BXGBM machine learning model [22].

4.1.3 Kernel Extreme Learning Machine (KELM)

KELM or Kernel Extreme Learning Machine is a machine learning algorithm that is an improvement to the original extreme learning machine (ELM) but uses kernel functions. This enables KELM to learn the nonlinear data relationships better. The strong sides of KELM are that it can reach broad learning speed and sound performance of learning generalization, and thus, it can be applied to diverse tasks, such as time-series forecasting and classification. Altogether, the models each provide a distinct way of approaching to machine learning to suit various features of prediction precision, computational capability, and model resilience. Fig. 7 illustrates the details of the steps involved in the design model of the Kernel Extreme Learning Machine. Equations from 9 to 13 were employed for the prediction of the kernel extreme learning machine algorithm.

$$\text{Hidden Layer Model, } H\beta = T \quad (9)$$

$$\text{Output Weight Estimation, } \beta = H^T \left(\frac{I}{C} + HH^T \right)^{-1} T \quad (10)$$

$$\text{Kernel Matrix Definition, } \Omega_{ij} = K(x_i, x_j) \quad (11)$$

$$\text{Dual Coefficient Computation, } \alpha = \left(\frac{I}{C} + \Omega \right)^{-1} T \quad (12)$$

$$\text{Final Prediction Equation, } \hat{y}(x) = \sum_{i=1}^N \alpha_i K(x, x_i) \quad (13)$$

where, H = Hidden layer output matrix, β = Output weight vector, T = Target/Output matrix, I = Identity matrix, C = Regularization coefficient, Ω_{ij} = Kernel value between x_i and x_j , $K(x_i, x_j)$ = Kernel function, α_i = Dual weight vector, x_i, x_j = Training input samples, N = Number of training samples, $\hat{y}(x)$ = Predicted output for input x .

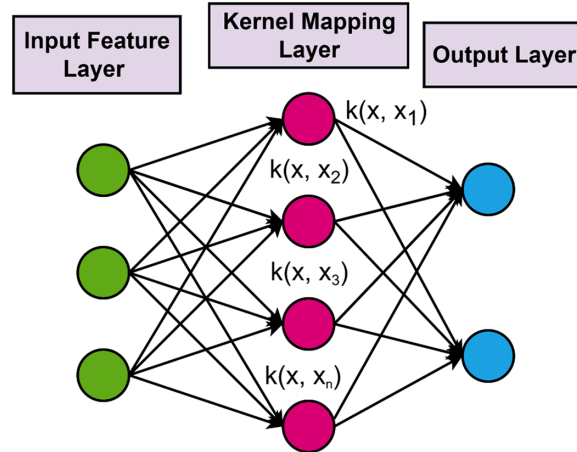


Figure 7: Representation of steps involved in the KELM machine learning model [23].

4.2 Sensitivity Analysis for the Model Performance

The predictive model performance in the present study is determined through a variety of statistical measures such as Coefficient of Determination R^2 , Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) as per Eqs. (14) to (17).

$$\text{Coefficient of Determination, } R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (14)$$

$$\text{Mean Squared Error, } \text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (15)$$

$$\text{Root Mean Squared Error, } \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

$$\text{Mean Absolute Error, } \text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (17)$$

where, y_i = Actual experimental values, \hat{y}_i = Predicted values, \bar{y} = Mean of actual value, n = Number of data samples.

4.3 Time Line Series for the Best Performance Model

Initially, the models were trained using default parameters. The model configuration for the initial phase is presented in Table 2. When models were trained using their parameters, the results showed that models faced an overfitting problem. Therefore, to solve the overfitting issue, a rigorous regularization strategy was adopted. The model tuning focused on minimizing the gap between the training and test data's R^2 score performance.

Table 2: Model configuration for initial and phase parameters.

Initial Phase			Tune Parameters		
KELM	DRVFL	BXGBM	KELM	DRVFL	BXGBM
Kernel = RBF	N_Hidden = 300	Number of Estimator = 300	Kernel = RBF	N_Hidden = 200	Number of Estimator = 400
Alpha = 0.01	Alpha = 0.001	Max Depth = 5	Alpha = 0.05	Alpha = 0.01	Max Depth = 4
Gamma = 0.5	Activation = tanh	Learning Rate = 0.05	Gamma = 0.2	Activation = tanh	Learning Rate = 0.04

For the BXGBM method, the learning rate is reduced to 0.04 from 0.05, and the maximum tree depth is fixed to 4. This forced the model to learn incremental, stable patterns rather than memorizing noise. In the KELM and DRVFL models, the regularization parameters are adjusted to 0.05 and 0.01, respectively. It adds a penalty to large weight values, effectively smoothing the decision boundary and improving test-set reliability. The corresponding results are summarized in [Table 3](#) for both methods. Additionally, a 10-fold validation approach is used in the parameter tuning. The dataset was divided into 10 equal subsets. In each iteration, nine out of ten subsets were used to train the models, and one set was used to test the model's performance [24]. This process is repeated 10 times, ensuring that all data samples are utilized for both training and validation. Finally, the average of 10 folds was used to present the quantitative performance of all models [25].

Table 3: (a) Results obtained using parameter tuning and 10-fold validation; (b) Results obtained using initial parameters.

Model	R ² Train	R ² Test	MSE Train	MSE Test	RMSE Train	RMSE Test	MAE Train	MAE Test
(a)								
BXGBM	0.986	0.941	3.757	16.630	1.938	4.078	1.141	2.640
DRVFL	0.892	0.839	29.858	46.030	5.464	6.784	4.157	4.823
KELM	0.950	0.891	13.851	31.059	3.721	5.573	2.717	3.786
(b)								
KELM	0.982	0.868	4.96	37.57	2.22	6.12	1.27	3.58
DRVFL	0.958	0.761	11.51	68.29	3.39	8.26	2.5	5.19
BXGBM	0.985	0.925	4.07	21.37	2.01	4.62	1.3	1.27

The results show that the tune parameters and 10-fold validation mitigated overfitting and improved the model's performance on the test data. KELM and DRVFL overfitting during the initial phase, with a large deviation in the test data's R² score relative to the training data. In both approaches, the BXGBM performed best. In the initial phase, the R² score for BXGBM was 0.9864 and 0.941 on the training and testing datasets, respectively. After parameter tuning, the model's performance on unknown data (i.e., test data) improves to 0.941. Furthermore, the gap between the training and test data's R² scores is reduced. This shows that parameter tuning and 10-fold cross-validation eliminated the overfitting issues.

The performance analysis of a machine learning model on the BXGBM dataset is presented in [Fig. 8](#) and split into training and testing. The scatter plot of the training dataset demonstrates a comparison of target values and model outputs. Although the correlation is evident, there are some differences that seem to show

that there is room to improve. The residual analysis indicates the mean squared error (MSE) of 3.757, and the root mean squared error (RMSE) is 1.938, indicating that the model is good on the training data. The error distribution histogram shows that the error profile is rather balanced, with a mean of 0 to 0.020 and a standard deviation of 1.9395. On the other hand, the testing data is more varied, with the MSE of 16.630 and RMSE of 4.078, indicating that the model might not be able to predict unseen data as strongly as it was tested. The mean error of the testing data is 2.64, and the standard deviation of the error is 4.074, indicating that the error differs more than the error of the training set. This variation requires additional adjustments to improve the robustness of the model and its use on unseen samples. The red line curve indicates the Gaussian distribution, and green lines show the sample error.

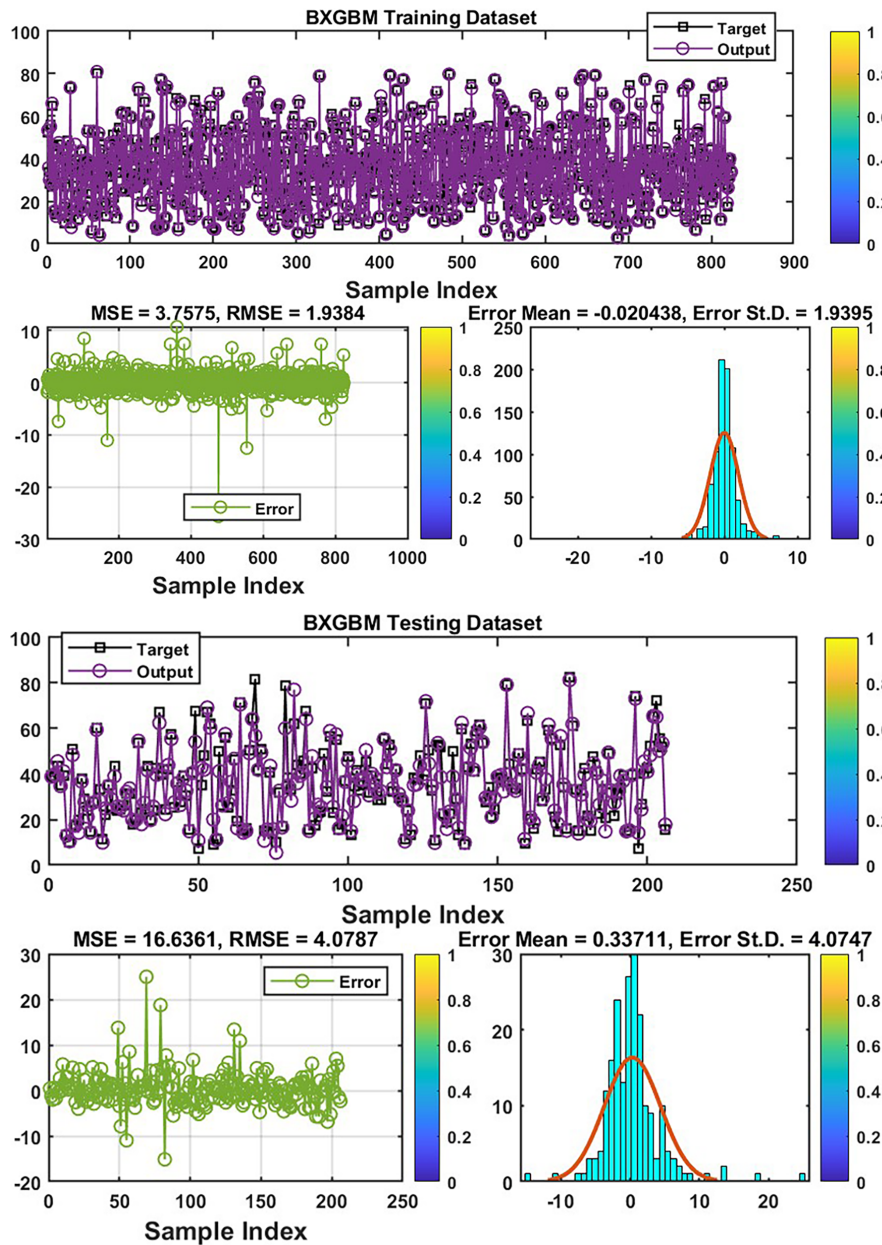


Figure 8: Training and testing dataset for BXGBM machine learning.

Fig. 9 demonstrates the performance of three methods, DRVFL, BXGBM, and KELM, in terms of training and testing data with the metrics of Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R^2 . Regarding the MSE, BXGBM has the greatest error of the training set at 3.757; its average error is large with regard to the values it predicts, whilst KELM and DRVFL work well on both training and testing sets. The values of RMSE indicate the same pattern, BXGBM has the highest RMSE (1.938) to be used in training, and the lower RMSE (4.078) of KELM and DRVFL is indicative of the high prediction capacity. These results are supported by the values of MAE, with the highest value of BXGBM of 1.141 in the training data and 2.640 in the testing data, indicating that they have a significant number of prediction errors. Conversely, KELM records the lowest values of MAE, namely 2.717 during training, and with a more stable performance.

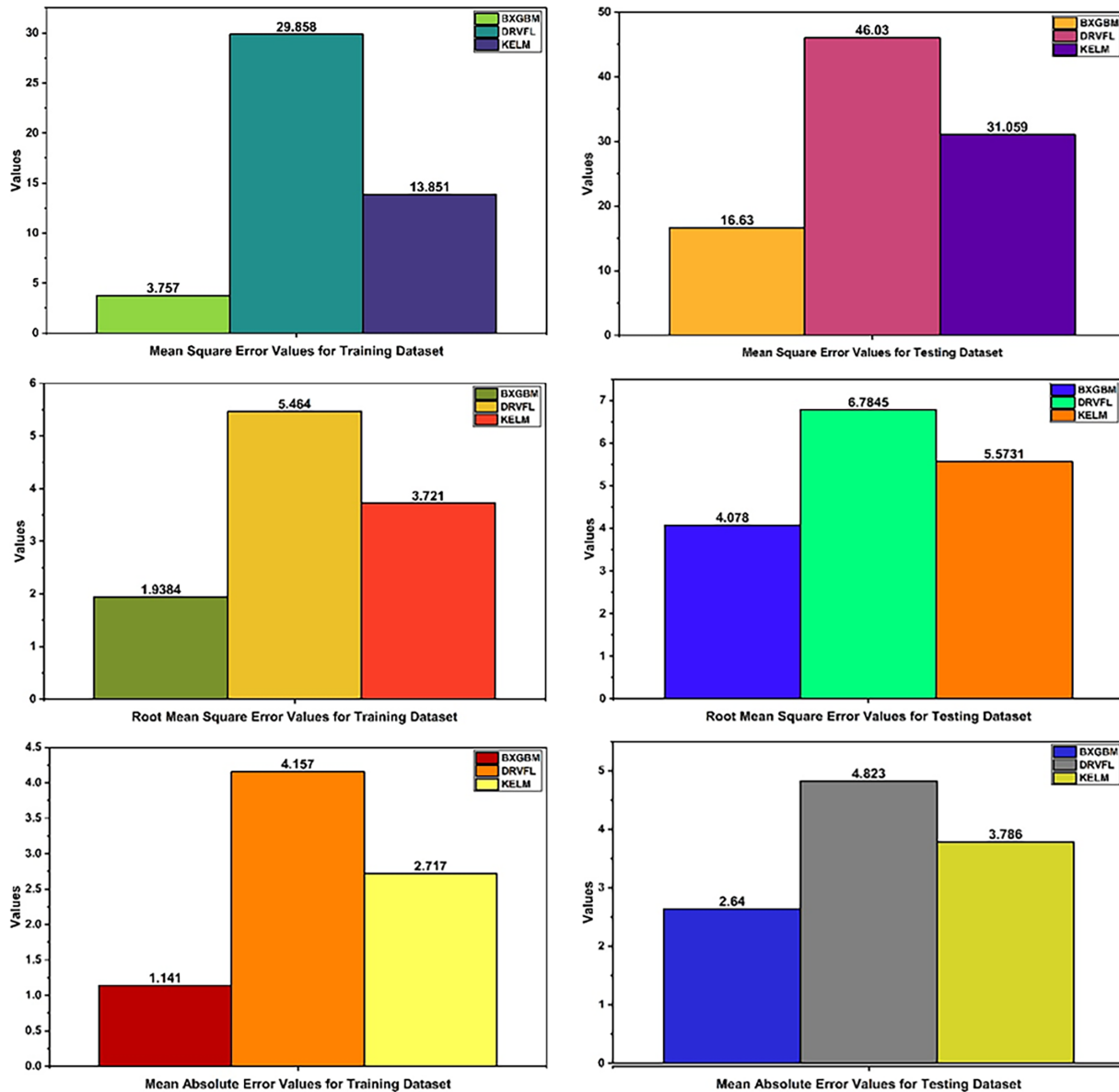


Figure 9: Metric values for various machine learning models.

Fig. 10 represents the regression performance analysis of three machine learning models, namely, DRVFL, BXGBM, and KELM, in terms of their accuracy of prediction using actual and predicted values of both training and testing data. In the case of the BXGBM training dataset, the model has a high coefficient of determination ($R^2 = 0.986$). It has the metric MSE = 3.757, RMSE = 1.938, and MAE = 1.141, and has a strong performance. On the other hand, the testing dataset illustrates $R^2 = 0.941$ with a greater MSE = 16.63, RMSE = 4.078, and MAE of 2.640. The analysis of KELM gives an R^2 of 0.891, which shows that the model accounts for 89.10 percent of the variance with an MSE of 31.059, RMSE of 5.573, and MAE of 3.786, indicating high accuracy [26]. Fig. 11 shows the R^2 value for the various machine learning models.

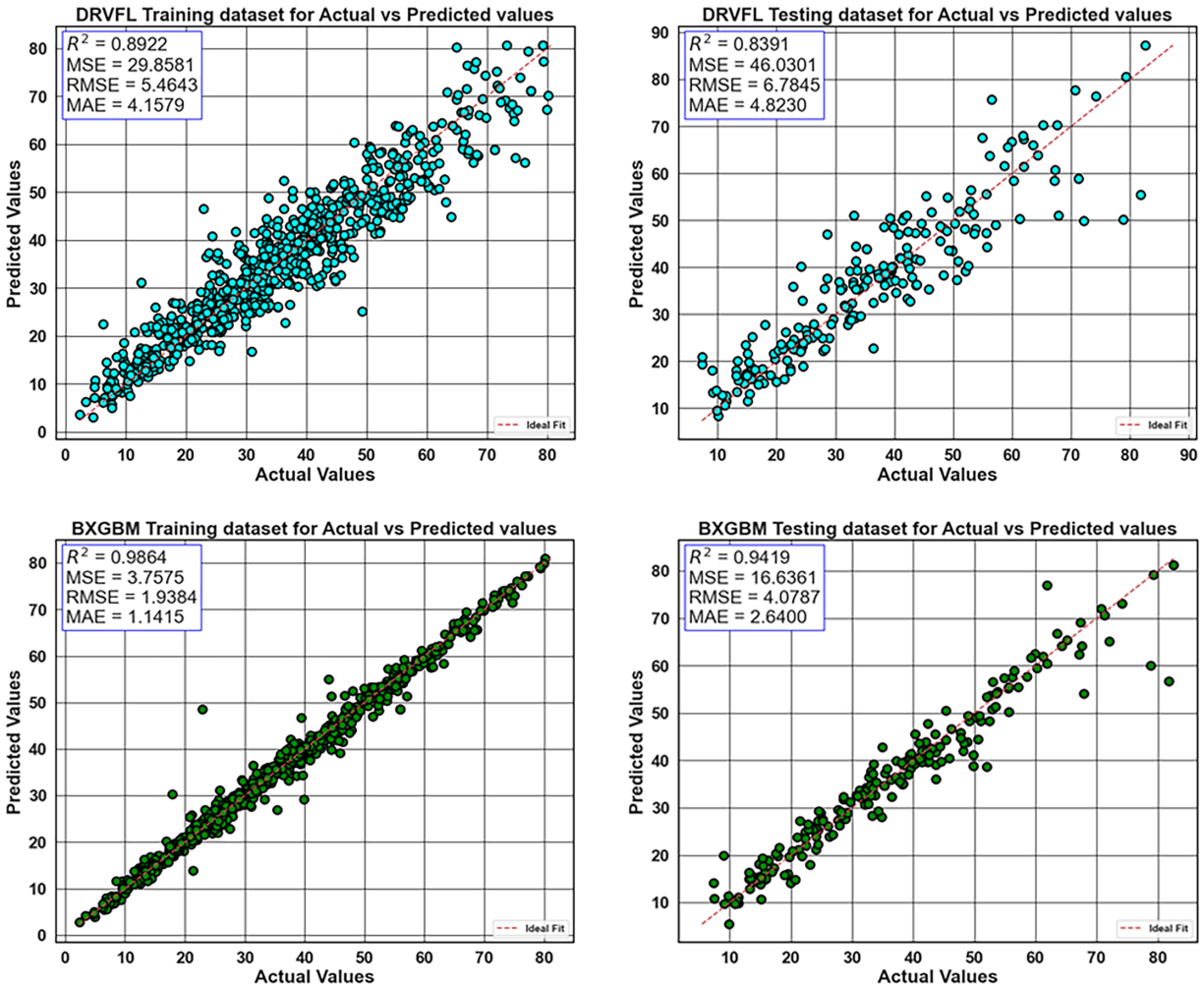


Figure 10: (Continued)

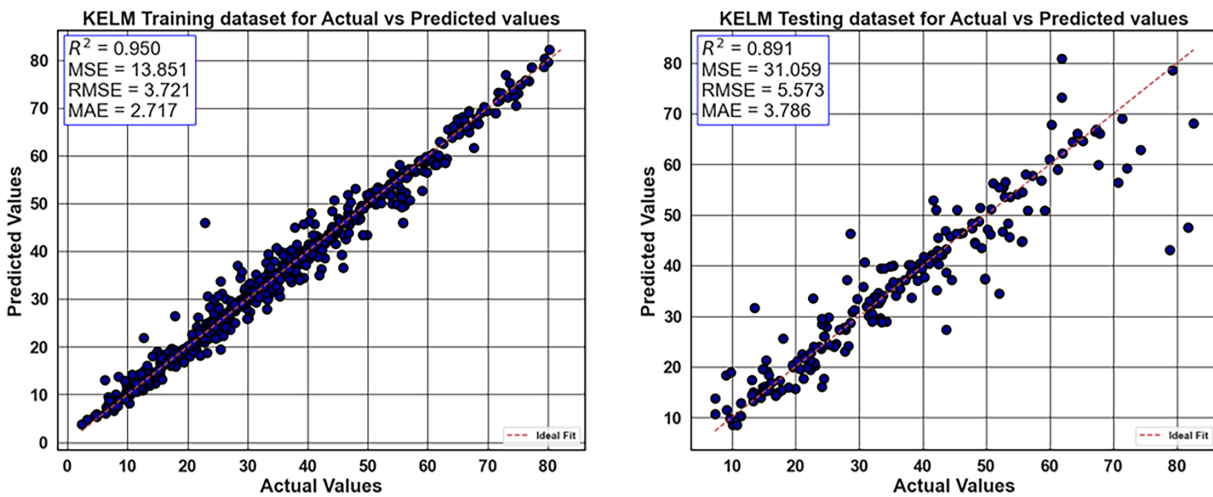


Figure 10: Regression analysis for the various machine learning algorithms.

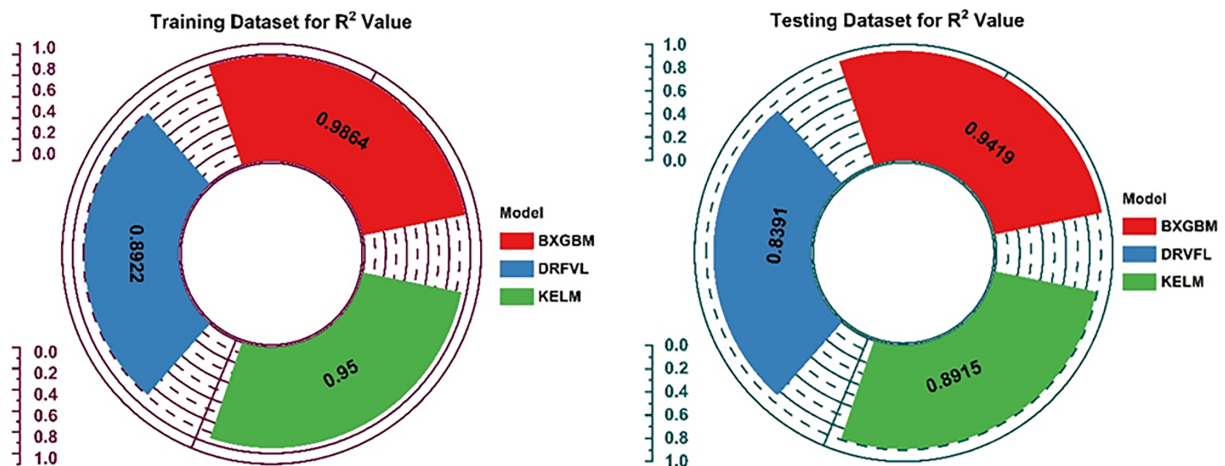


Figure 11: Training and testing dataset R^2 score values for various models.

On the whole, the three models share good training performances, but BXGBM has a balance between training and testing results as it presents good predictive capabilities of concrete strength. KELM and DRVFL have lower performance compared to BXGBM. KELM struggles to match the gradient-boosting efficiency of BXGBM, as it attempts to solve a global optimization problem in a projected feature space, which can be less stable when dealing with the specific noise patterns in concrete mix designs. The use of fixed and randomly generated weights allows faster training for DRVFL; the model's reliance on random projections may not always capture the most relevant feature interactions effectively. While in BGXBM, the use of the boosting approach corrects errors in each iteration, leading to a more refined fit for the readings. Though the model performed well, there are certain limitations. The model primarily focuses on mixed proportions (Cement, Water, Ash, etc.) and Age. It does not currently account for external factors like curing temperature, humidity, or mineralogical composition of the aggregates, which can influence long-term strength development. The current study relies on a dataset of 1031 readings, which, while statistically significant, may not capture the full variance of environmental conditions or specialized chemical admixtures used in global construction

practices. Table 4 illustrates the performance metric analysis for the different machine learning models, and Table 5 displays the details of the previous study with the present work.

Table 4: Performance metric analysis for the different machine learning models.

Model	Dataset	MSE	RMSE	MAE	R ²
BXGBM	Training	3.757	1.938	1.141	0.986
	Testing	16.630	4.078	2.640	0.941
DRVFL	Training	29.858	5.464	4.157	0.892
	Testing	46.030	6.784	4.823	0.839
KELM	Training	13.851	3.721	2.717	0.950
	Testing	31.059	5.573	3.786	0.891

Table 5: Comparison of machine learning model performance with the literature.

Machine Learning Models	Application	Data for Study	R ² Score for Prediction	Ref.
DNN, MOPSO	Concrete mix design optimization	Concrete mix and curing parameters	0.936	[27]
DNN, Classical ML models	Digital twin for material design	Online dataset	0.90 and 0.63	[28]
ANN	Predicting SCC compressive strength	Ultrasonic measurement	0.99	[29]
XG Boost model	Strength of FRP confined columns	Database of 366 samples	0.97	[30]
BXGBM, DVRFL KELM	Predicting the compressive strength of concrete	Database of 1030 samples	0.986, 0.892, and 0.950	Present Study

4.4 Feature Importance and Impact Analysis

To explain the underlying physical mechanisms leading to concrete strength, an interpretative analysis was performed using the optimized BXGBM framework. Firstly, a global feature importance metric using a Gain-based ranking was established to quantify the fractional contributions of constituents such as Cement, Water, and Age. Fig. 12 shows the ranking of these features. It is observed that age and cement are the most significant predictors, followed by Water and Superplasticizer. While this identifies the primary drivers of the model, it does not account for the directionality of the influence. Therefore, a Shapley Additive Explanations (SHAP) was employed to identify complex, non-linear interactions among predictors. SHAP values disclose how specific concentrations of predictors, such as the water-to-cement ratio, positively or negatively shift the predicted strength. Fig. 13 shows the SHAP-based analysis plot. The SHAP plot confirms that large values of Age and Cement (represented by red points) consistently have a positive impact on the model output, thereby increasing predicted strength. The water concentration shows a negative correlation. A higher water content (i.e., red points) shifts the SHAP value to the left of the center line, indicating a reduction in strength.

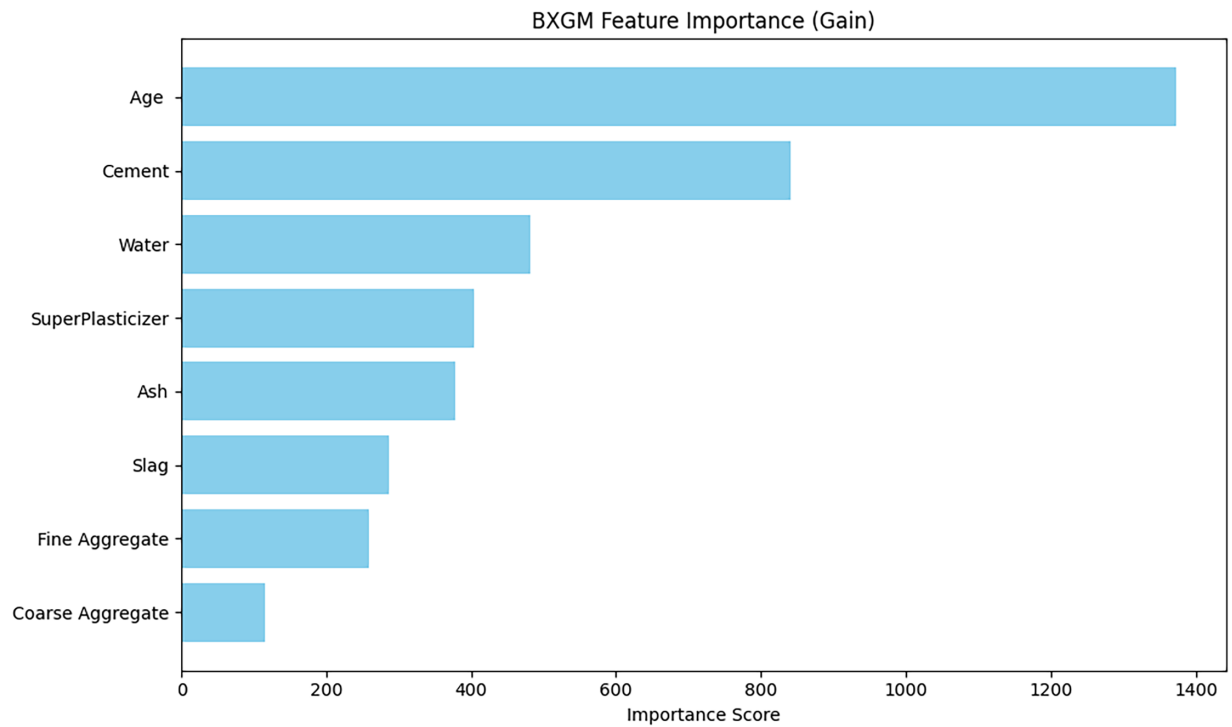


Figure 12: Feature ranking using gain.

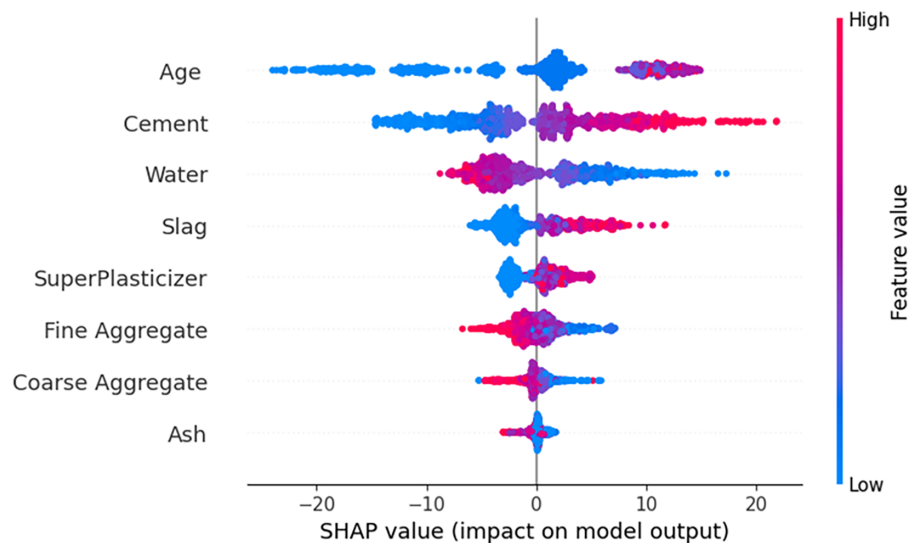


Figure 13: Feature importance using SHAP values.

5 Conclusion

The analysis of the three predictive models, BXGBM, DRVFL, and KELM, brings into view the unique performance features in the predictive strength of concrete. The BXGBM model results in good training performance with a Mean Squared Error (MSE) of 3.757 and a greatly impressive value of R^2 of 0.986. Its test performances, however, show high overfitting with an MSE of 16.63 and an R^2 of 0.941, indicating good accuracy. On the other hand, the DRVFL model in terms of training and testing R^2 of 0.892, 0.839, and MSE

of 29.858, 46.030. Competitiveness is practiced by KELM, found to be training MSE of 13.851 with R^2 of 0.950. Nonetheless, its testing performance, which is represented by an MSE of 31.059 and an R^2 of 0.891, depicts a certain deterioration as compared to its training performance. All in all, BXGBM is the best model to use considering its accuracy and generalizability in predicting concrete strength. These results indicate that it would be useful in the work of engineers and researchers in the field of concrete technology, and further work can be aimed at refining these models or the development of hybrid methods to strengthen prediction capabilities even more.

6 Future Recommendation

To improve the quality of the concrete mix, an additional concrete mix design parameter might be considered during the concrete preparation. The current model uses inputs such as cement, waste slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, and curing age. For further refinement, this can be achieved by including the inputs in the mixing procedure, such as curing conditions like temperature, humidity, compaction methods, and sustainable waste materials. In addition, increasing the dataset beyond the current dataset from 1030 to more, with different mix proportions, can improve the model generalization. Utilizing advanced deep learning models can improve the predictive strength of the prediction.

Acknowledgement: Not Applicable.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Nayeemuddin Mohammed: Conceptualization, Formal Analysis, Visualization, Investigation, Software, Writing—Original Draft. Tahar Ayadat: Proofreading. Andi Asiz: Investigation, Proofreading. Nadeem Pasha: Writing. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the Corresponding Author, upon reasonable request.

Ethics Approval: Not applicable.

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

References

1. Khan M, McNally C. A holistic review on the contribution of civil engineers for driving sustainable concrete construction in the built environment. *Dev Built Environ.* 2023;16(9):100273. doi:10.1016/j.dibe.2023.100273.
2. Mouawad F, Homsy F, Geara F, Mina R. Predicting compressive strength of sustainable concrete using machine learning and artificial neural networks. *Constr Mater.* 2025;5(3):56. doi:10.3390/constrmater5030056.
3. Paudel S, Pudasaini A, Shrestha RK, Kharel E. Compressive strength of concrete material using machine learning techniques. *Clean Eng Technol.* 2023;15(4):100661. doi:10.1016/j.clet.2023.100661.
4. Asteris PG, Koopialipoor M, Armaghani DJ, Kotsonis EA, Lourenço PB. Prediction of cement-based mortars compressive strength using machine learning techniques. *Neural Comput Appl.* 2021;33(19):13089–121. doi:10.1007/s00521-021-06004-8.
5. Güçlüer K, Özbeyaz A, Göymen S, Günaydın O. A comparative investigation using machine learning methods for concrete compressive strength estimation. *Mater Today Commun.* 2021;27:102278. doi:10.1016/j.mtcomm.2021.102278.
6. Zhang C, Zhu Z, Liu F, Yang Y, Wan Y, Huo W, et al. Efficient machine learning method for evaluating compressive strength of cement stabilized soft soil. *Constr Build Mater.* 2023;392(8):131887. doi:10.1016/j.conbuildmat.2023.131887.

7. Hadzima-Nyarko M, Nyarko EK, Lu H, Zhu S. Machine learning approaches for estimation of compressive strength of concrete. *Eur Phys J Plus*. 2020;135(8):682. doi:10.1140/epjp/s13360-020-00703-2.
8. Chopra P, Sharma RK, Kumar M, Chopra T. Comparison of machine learning techniques for the prediction of compressive strength of concrete. *Adv Civ Eng*. 2018;2018(1):5481705. doi:10.1155/2018/5481705.
9. Feng DC, Liu ZT, Wang XD, Chen Y, Chang JQ, Wei DF, et al. Machine learning-based compressive strength prediction for concrete: an adaptive boosting approach. *Constr Build Mater*. 2020;230(3):117000. doi:10.1016/j.conbuildmat.2019.117000.
10. Alshboul O, Almasabha G, Shehadeh A, Al Mamlook RE, Almuflih AS, Almakayeel N. Machine learning-based model for predicting the shear strength of slender reinforced concrete beams without stirrups. *Buildings*. 2022;12(8):1166. doi:10.3390/buildings12081166.
11. Yaseen ZM. Machine learning models development for shear strength prediction of reinforced concrete beam: a comparative study. *Sci Rep*. 2023;13(1):1723. doi:10.1038/s41598-023-27613-4.
12. Abuodeh OR, Abdalla JA, Hawileh RA. Assessment of compressive strength of ultra-high performance concrete using deep machine learning techniques. *Appl Soft Comput*. 2020;95(2):106552. doi:10.1016/j.asoc.2020.106552.
13. Li D, Tang Z, Kang Q, Zhang X, Li Y. Machine learning-based method for predicting compressive strength of concrete. *Processes*. 2023;11(2):390. doi:10.3390/pr11020390.
14. Civil engineering: cement manufacturing dataset. [cited 2025 Jan 1]. Available from: <https://www.kaggle.com/datasets/vinayakshanawad/cement-manufacturing-concrete-dataset/data>.
15. Al-Jamimi HA, Al-Kutti WA, Alwahaishi S, Alotaibi KS. Prediction of compressive strength in plain and blended cement concretes using a hybrid artificial intelligence model. *Case Stud Constr Mater*. 2022;17(2):e01238. doi:10.1016/j.cscm.2022.e01238.
16. Chen B, Wang L, Feng Z, Liu Y, Wu X, Qin Y, et al. Optimization of high-performance concrete mix ratio design using machine learning. *Eng Appl Artif Intell*. 2023;122(5):106047. doi:10.1016/j.engappai.2023.106047.
17. Hu M, Heng Chion J, Suganthan PN, Katuwal RK. Ensemble deep random vector functional link neural network for regression. *IEEE Trans Syst Man Cybern Syst*. 2023;53(5):2604–15. doi:10.1109/TSMC.2022.3213628.
18. Sharma R, Goel T, Tanveer M, Dwivedi S, Murugan R. FAF-DRVFL: fuzzy activation function based deep random vector functional links network for early diagnosis of Alzheimer disease. *Appl Soft Comput*. 2021;106(3):107371. doi:10.1016/j.asoc.2021.107371.
19. Asiful Islam M, Al Muzaddid MA, Jahin Prema A, Vuske SR. Comparative assessment of concrete compressive strength prediction at industry scale using embedding-based neural networks, transformers, and traditional machine learning approaches. *arXiv:2601.09096*. 2026.
20. Ni HG, Wang JZ. Prediction of compressive strength of concrete by neural networks. *Cem Concr Res*. 2000;30(8):1245–50. doi:10.1016/S0008-8846(00)00345-8.
21. Xia Y, Jiang S, Meng L, Ju X. XGBoost-B-GHM: an ensemble model with feature selection and GHM loss function optimization for credit scoring. *Systems*. 2024;12(7):254. doi:10.3390/systems12070254.
22. Deng X, Ye A, Zhong J, Xu D, Yang W, Song Z, et al. Bagging-XGBoost algorithm based extreme weather identification and short-term load forecasting model. *Energy Rep*. 2022;8(11):8661–74. doi:10.1016/j.egy.2022.06.072.
23. Kaya Karakutuk A, Ozdemir O. A novel fuzzy kernel extreme learning machine algorithm in classification problems. *Appl Sci*. 2025;15(8):4506. doi:10.3390/app15084506.
24. Lai BL, Bao RL, Zheng XF, Vasdravellis G, Mensinger M. Machine-learning assisted analysis on the seismic performance of steel reinforced concrete composite columns. *Structures*. 2024;68(12):107065. doi:10.1016/j.istruc.2024.107065.
25. Lai BL, Yang L, Xiong MX. Numerical simulation and data-driven analysis on the flexural performance of steel reinforced concrete composite members. *Eng Struct*. 2021;247(9):113200. doi:10.1016/j.engstruct.2021.113200.
26. Zhang X, Dai C, Li W, Chen Y. Prediction of compressive strength of recycled aggregate concrete using machine learning and Bayesian optimization methods. *Front Earth Sci*. 2023;11:1112105. doi:10.3389/feart.2023.1112105.
27. Tipu RK, Rathi P, Pandya KS, Panchal VR. Optimizing sustainable blended concrete mixes using deep learning and multi-objective optimization. *Sci Rep*. 2025;15(1):16356. doi:10.1038/s41598-025-00943-1.

28. Xin P, Isleem HF, Khishe M. A digital twin approach for sustainable construction: predictive optimization of concrete strength using industry 4.0 principles. *Sci Rep.* 2026;16(1):2443. doi:10.1038/s41598-025-32276-4.
29. Benaicha M. AI-driven prediction of compressive strength in self-compacting concrete: enhancing sustainability through ultrasonic measurements. *Sustainability.* 2024;16(15):6644. doi:10.3390/su16156644.
30. Cui R, Yang H, Li J, Xiao Y, Yao G, Yu Y. Machine learning-based prediction of compressive strength in circular FRP-confined concrete columns. *Front Mater.* 2024;11:1408670. doi:10.3389/fmats.2024.1408670.