Scientific Reports

https://doi.org/10.1038/s41598-026-45024-z

Article in Press

Machine learning-based prediction of compressive and flexural strength of wheat straw reinforced sustainable gypsum composites

Received: 26 January 2026

Accepted: 16 March 2026

Published online: 26 March 2026

Cite this article as: Ahmad H., Ejaz M.F., Riaz M.R. *et al.* Machine learning-based prediction of compressive and flexural strength of wheat straw reinforced sustainable gypsum composites. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-45024-z

Haseeb Ahmad, Muhammad Fahad Ejaz, Muhammad Rizwan Riaz, Shaban Shahzad, Sarah El Kadri & Maria G. Kmeid

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Machine learning-based prediction of compressive and flexural strength of wheat straw reinforced sustainable gypsum composites

Haseeb Ahmad[1], Muhammad Fahad Ejaz[2*], Muhammad Rizwan Riaz[3], Shaban Shahzad[4], Sarah El Kadri[5], Maria G. Kmeid[5]

[1] Department of Civil and Environmental Engineering, Saitama University, Sakura, Saitama, 338-8570, Japan

[2] Graduate School of Urban Innovation, Department of Civil Engineering, Yokohama National University, Kanagawa, Japan

[3] Department of Civil Engineering, University of Engineering & Technology, Lahore, 54890, Pakistan

[4] Interdisciplinary research center for Construction and Building Materials, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, 31261, Kingdom of Saudi Arabia

[5] College of Engineering and Technology, American University of the Middle East, Kuwait

---

ABSTRACT

The environmental burden associated with conventional cement-based materials has intensified research for sustainable alternatives with lower carbon footprints. For this, gypsum-based composites reinforced with agricultural waste, such as wheat straw, offer a promising solution. However, their mechanical performance is governed by nonlinear and complex interactions among multiple mixture parameters. This study proposes a comprehensive machine learning (ML) framework to predict the compressive and flexural strength of wheat straw reinforced gypsum composites. A dataset comprising 161 experimental samples was used and five ML models: Artificial Neural Network, Gaussian Process Regression (GPR), Random Forest, Extreme Gradient Boosting, and Support Vector Machine, were used. Model performance was assessed using 10-fold cross-validation with multiple statistical metrics along with Taylor diagram analysis. Among the evaluated models, GPR demonstrated superior predictive capability for both compressive and flexural strength, while providing uncertainty quantification that enhances reliability for engineering applications. Feature importance and SHapley Additive exPlanations analyses were employed to improve model interpretability, revealing gypsum strength as the most influential parameter, with water-related parameters, wheat straw content, and chemical additives contributing secondary effects. The proposed ML-based framework provides acceptable and interpretable predictions, offering the optimization of sustainable gypsum composites while reducing experimental efforts and supporting environment-friendly construction.

Keywords: Gypsum; Wheat straw; Compressive strength; Flexural strength; Machine learning models

---

*Corresponding author: Muhammad Fahad Ejaz
fahad.ejaz1995@gmail.com; ejaz-fahad-hf@ynu.jp

1. Introduction

Concrete was the second most widely used construction material worldwide, with annual production exceeding 14 billion m^3 [1, 2]. This extensive utilization was attributed to high mechanical strength, durability, and relatively low cost [3]. However, the large-scale production of concrete resulted in the consumption of nearly 40% of global natural resources, including timber, sand, water, and stone [4, 5]. Moreover, cement, a primary binder of concrete, was identified as responsible for approximately 7-8% of global carbon dioxide emissions, with nearly 900 kg of CO_2 reported to be released per ton of cement produced [6-8]. These emissions were projected to increase substantially by 2050 in the absence of effective mitigation strategies [9]. In response to these challenges, significant research efforts were directed toward the development of sustainable alternatives to conventional cement-based materials. Various approaches were explored, including the utilization of recycled wastes [10, 11], and the development of alternative binders such as geopolymer concrete [12, 13], lightweight concrete [14, 15], and gypsum-based composites [16].

Among these, gypsum was recognized as a particularly promising alternative due to its comparatively lower CO_2 emissions, reported at approximately 140 kg per ton during natural gypsum production [17]. Beyond its reduced carbon footprint, gypsum was characterized by several favorable properties, including rapid setting time, low density, excellent fire resistance, architectural aesthetics, good acoustic and thermal insulation behavior [18-20]. These properties supported extensive historical use of gypsum in non-load bearing components such as prefabricated plasterboards, ceiling tiles, interior partitions, and building blocks [21-25]. Therefore, it offers advantages over conventional bricks and concrete blocks due to its lightweight nature, reduced material consumption, and lower environmental impact. Despite these benefits, the broader structural application of gypsum remained limited due to inherent drawbacks, including brittleness, lower mechanical strength, and water resistance [26].

To overcome these shortcomings, extensive research has been conducted on gypsum composites reinforced with various materials, including rice husk [27], plastic [28], polystyrene [29], ground bricks [30], banana fibers [31], wood waste [32], and wheat straw [33]. Among different reinforcement options, the incorporation of agricultural waste fibers presented an effective dual-purpose strategy. It enhanced the ductility and flexural performance of gypsum matrices while addressing critical environmental issues related to waste management and resource conservation. Globally, agricultural activities generated approximately 140 billion metric tons of waste annually [34], and common disposal practices such as open-field burning and landfilling were associated with severe air, soil, and water pollution [35]. In agriculture-based economies like Pakistan, large quantities of wheat straw, about 13 million tons, were produced annually [36], with a substantial portion burned openly, exacerbating smog and environmental health concerns [37]. Therefore, incorporation of wheat straw residue into gypsum composites offers a viable pathway for waste valorization. Furthermore, wheat straw was lightweight, renewable, cost-effective, and readily available, which made it suitable as a reinforcing filler in gypsum-based components [38]. However, the mechanical performance of wheat straw reinforced gypsum composites was governed by complex nonlinear interactions among multiple mix design parameters, including pre-treatments, straw dosage, matrix composition, and gypsum strength. This emphasized the need for careful mixture design and systematic optimization to achieve balanced compressive and flexural performance. Conventional experimental approaches provided valuable insights but remained costly, time-consuming, and limited in their ability to explore optimal mixtures efficiently.

In this context, machine learning (ML) has emerged as a powerful, data-driven approach capable of modeling complex nonlinear behaviors and supporting predictive analysis. As a subfield of artificial intelligence, ML algorithms were trained to learn patterns from data and generate predictions without explicit programming. In civil engineering applications, ML techniques were applied to predict material properties [39], estimate construction costs [40], optimize mix designs [41], compression strength [42, 43], assess structural performance [44], and vibrations induced by traffic [45]. Advanced algorithms such as Artificial neural networks (ANN) and random forest (RF) demonstrated superior capability in predicting nonlinear material behavior compared to traditional regression-based methods [46].

To the best of the author's knowledge, despite the growing application of ML in civil engineering, limited attention was directed towards the gypsum composites reinforced with agricultural waste, notably with respect to systematic predictive modeling of mechanical properties. Therefore, this research aimed at developing an ML-based framework to predict the flexural and compressive behavior of wheat straw reinforced gypsum composites. The framework is intended to reduce experimental effort, improve prediction accuracy, and support the development of eco-friendly gypsum composites suitable for sustainable construction applications.

Based on existing research outcomes and the features of different ML approaches, five representative models were selected in the current study. ANN is well-suited for modeling complex nonlinear relationships through layered feature transformation and adaptive weight learning. Gaussian process regression (GPR) provides a modeling framework that not only achieves high prediction accuracy but also offers uncertainty quantification present in data. RF effectively captures nonlinear interactions through the ensemble of several decision trees and demonstrates effectiveness against overfitting. Extreme gradient boosting (XGBoost) incrementally reduces prediction errors through gradient boosting and incorporates regularization mechanisms, enabling an effective balance between predictive performance and generalization capability. Support vector machine (SVM) constructs optimal regression functions by maximizing margins in high-dimensional feature spaces, making it suitable for handling nonlinear problems with limited data. The complementary strengths of these models enable a comprehensive and reliable comparison for predicting the mechanical properties of gypsum composites.

The remainder of the paper is organized as follows. Section 2 details the methodology, including data collection, theoretical background of selected ML algorithms, and the optimization framework. Section 3 presents the results, involving model performance evaluation, feature importance analysis, uncertainty quantification, and explainability of the best model using SHAP analysis. The major conclusions drawn are summarized in Section 4.

2. Research methodology

2.1. General framework overview

This study presents an ML-based framework to predict the compressive and flexural strength of wheat straw reinforced gypsum composites and to interpret the influence of key mixture parameters. The research framework is illustrated in Figure 1. A dataset comprising 161 experimental samples was compiled from the literature [47-51], including gypsum strength, gypsum quantity, water content, water-to-gypsum ratio, wheat straw content, and chemical additives as input parameters, with compressive and flexural strength as output responses. Five ML models: ANN, GPR, RF, XGBoost, and SVM, used 10-fold cross-validation and evaluated multiple performance metrics. Model convergence and stability were further assessed through learning curves and statistical diagnostics. To enhance model interpretability, permutation feature importance was applied to all models, while SHAP analysis was performed for the best-performing model. This integrated framework enables accurate prediction, robust evaluation, and physically meaningful interpretation of the mechanical properties of gypsum composites incorporating wheat straw.

2.2. Data collection

The collected 161 gypsum composite formulations are presented in Table A1. Gypsum strength and quantity ranged from 2 to 16 MPa and 900 to 1000 g, respectively. Water varied from 509 to 980 g, corresponding to water-to-gypsum ratios between 0.55 and 1.0. Fibers of wheat straw ranged between 0 to 100 g, while $CaCl_2$ and $Ca(OH)_2$ were incorporated at 5% in selected mixes. Flexural strength values spanned 0.19 to 7.65 MPa, and compressive strength values spanned 0.4 to 41.62 MPa.

Figure 2 shows the comparative analysis of compressive and flexural strength results for all mixtures with data sourced from individual papers. Higher strengths appear predominantly in the initial sample groups, which correspond to mixes with no wheat straw content. The figure confirms that similar performance trends occur across all sources. The statistical characteristics of the different parameters are presented in Table 1, providing a comprehensive overview of each parameter. In this study, the mix-design parameters (gypsum strength, gypsum quantity, water content, water-to-gypsum ratio, wheat straw content, and chemical additives) were treated as input variables, whereas compressive and flexural strength were selected as the output variables representing the mechanical performance of the composites.

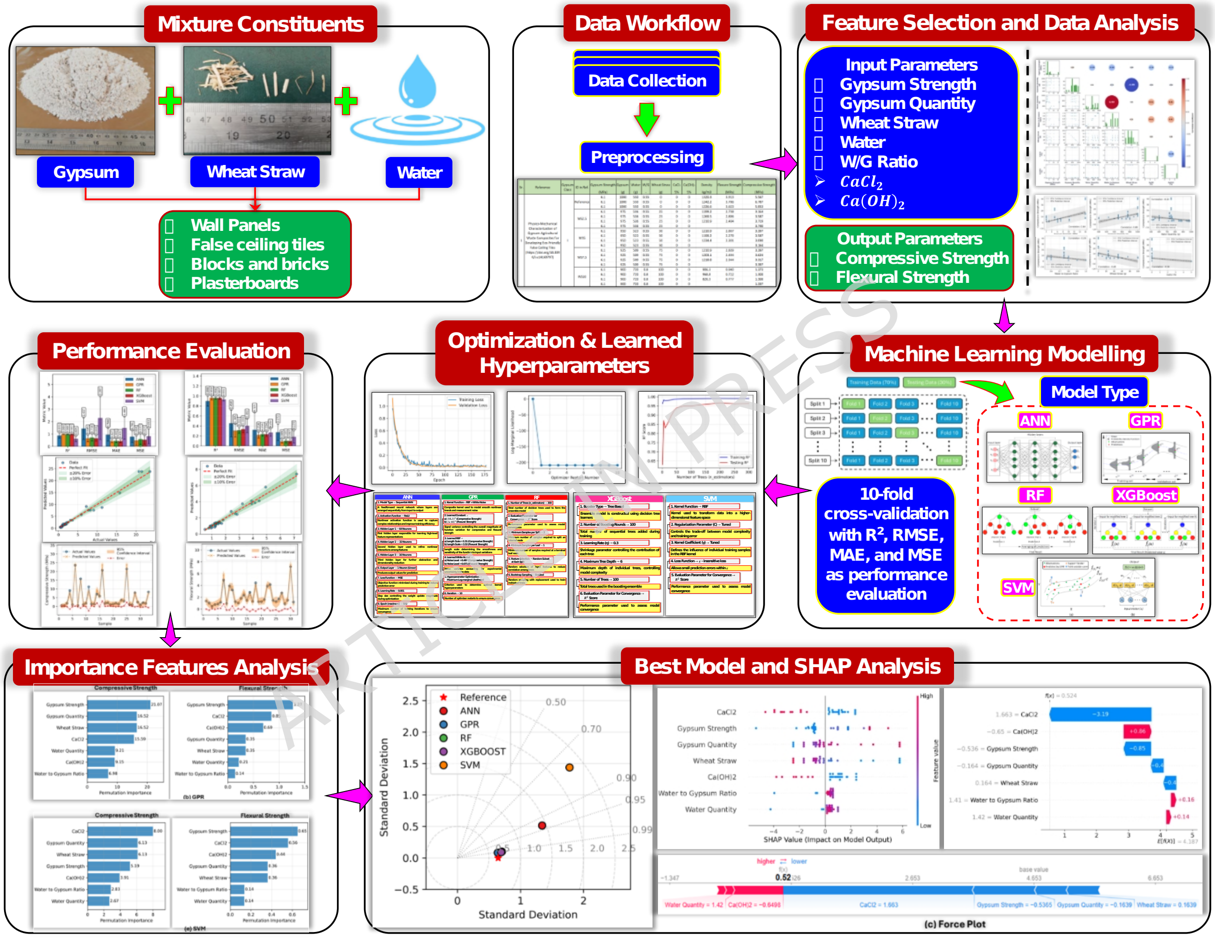


Figure 1. Flowchart of the proposed framework.

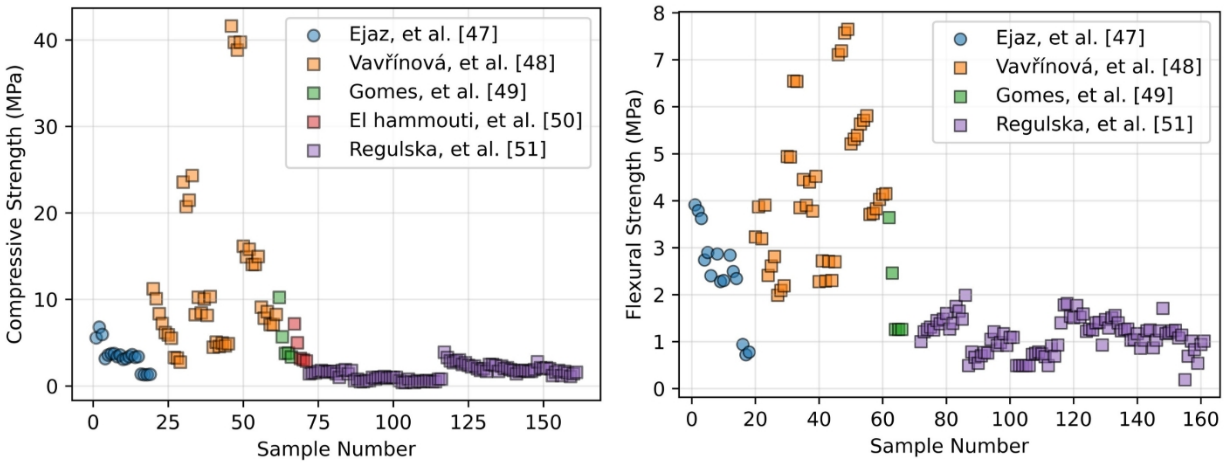


Figure 2. Variation of compressive and flexural strength with sample number from [47-51].

Table 1. Statistical properties of collected data.

Parameter	Parameter Names	Mean	Min	Max	Standard Deviation	Coefficient of Variation (%)
Input Parameters	Gypsum Strength (MPa)	5.05	2	16	4.10	81.11
	Gypsum (g)	962.33	900	1000	21.25	2.20
	Water (g)	793.57	509	980	119.53	15.06
	Water-to-gypsum Ratio	0.82	0.55	1	0.13	15.40
	Wheat Straw (g)	37.66	0	100	21.25	56.44
	$CaCl_2$ (%)	1.48	0	5	2.29	154.71
	$Ca(OH)_2$ (%)	1.48	0	5	2.29	154.71
Output Parameters	Compressive Strength (MPa)	4.96	0.4	41.62	7.38	148.81
	Flexural Strength (MPa)	2.13	0.19	7.65	1.65	77.83

Among the input parameters, the gypsum quantity shows relatively low variability (mean = 962.33 g, COV = 2.20%), indicating consistent material dosage across mixtures, while water content exhibits a mean of 793.57 g and a greater dispersion (COV = 15.06%), with a possible reason being of use of multiple water-to-gypsum ratios. The water-to-gypsum ratio ranges from 0.55 to 1.00, with a mean of 0.82 and a COV of 15.40%, reflecting the variation introduced to investigate water sensitivity. Wheat straw content demonstrates high variability (mean = 37.66 g, COV = 56.44%), consistent with its role as a variable reinforcement parameter. The chemical additives $CaCl_2$ and $Ca(OH)_2$ show the highest coefficients of variation (both 154.71%), as they were presented only in selected mixtures at 5% dosage, while absent in others. For the output parameters, compressive strength exhibits substantial dispersion (mean = 4.96 MPa, range = 0.40 to 41.62 MPa, COV = 148.81%), capturing the broad mechanical response associated with changes in mixture composition. Flexural strength also shows notable variability (mean = 2.13 MPa, COV = 77.83%), but to a lesser extent than compressive strength. Overall, the statistical profile indicates a wide parameter space for training ML techniques.

The distribution of all input and output parameters used in the dataset is presented in Figure 3, illustrating the variability and frequency of parameters across the 161 gypsum composite mixtures. Gypsum strength and gypsum quantity are concentrated between 2 to 7.6 MPa and 940 to 1000 g, respectively, while water quantity exhibits a concentration between 697 to 980 g. For water-to-gypsum ratio, most mixtures fall between 0.8 and 1.0. Wheat straw content displays four main groups, with most of quantity fall between 25 to 75 g. The distributions for $CaCl_2$ and $Ca(OH)_2$ confirm that chemical additives were included with a fixed dosage of 5% for 45 samples and 0% dosage for 107 samples. Most mixtures exhibit compressive strengths below 8 MPa, while only a small subset reaches values above 20 MPa. Similarly, flexural strength is primarily concentrated below 2 MPa, with fewer samples showing values above 4 MPa, as shown in Figure 3. Overall, the figure highlights the range of mixture compositions and mechanical responses, demonstrating that the dataset spans a possible large parameter space for ML analysis.

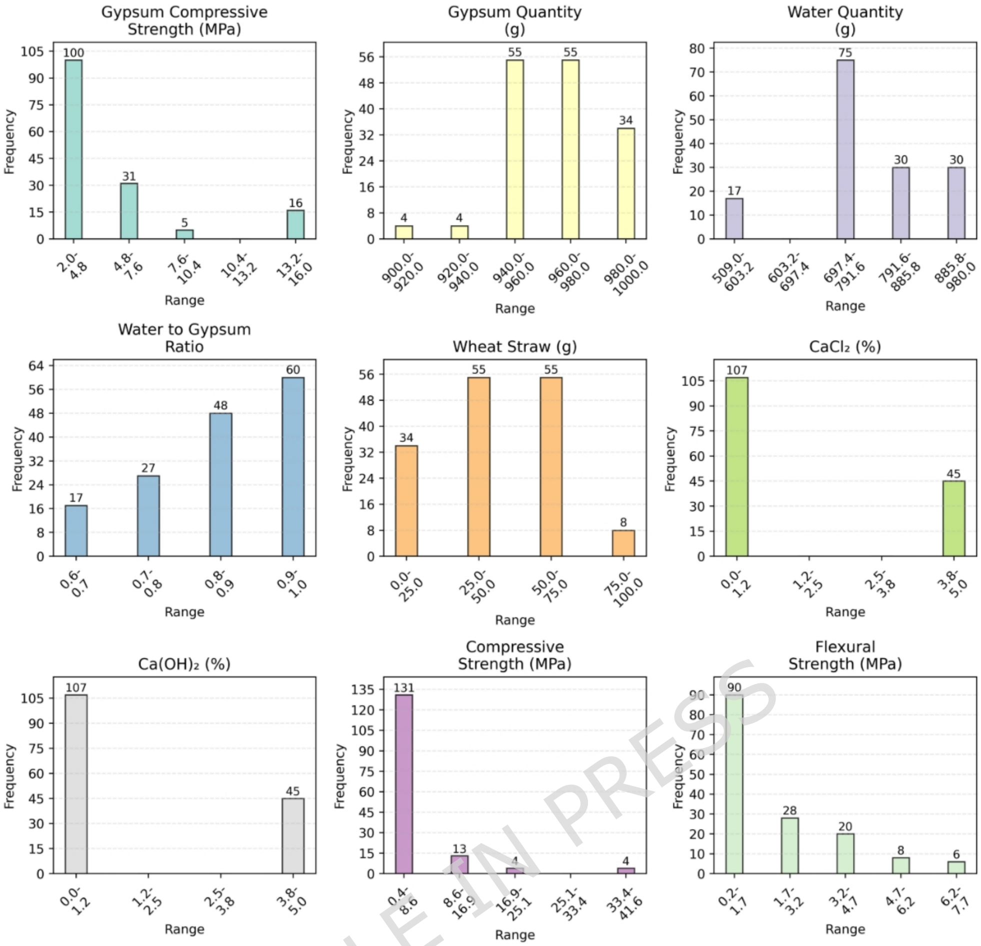


Figure 3*.* The distribution range of parameters.

2.2.1. Correlation between input parameters

A combined pairwise scatterplot matrix and correlation heatmap for all input parameters, enabling visual assessment of relationships within the dataset, is illustrated in Figure 4. The diagonal panels show the distribution of each variable as discussed above, while the lower triangle scatterplots illustrate how the parameters vary with respect to one another. The corresponding Pearson correlation coefficients are displayed in the upper triangular heatmap.

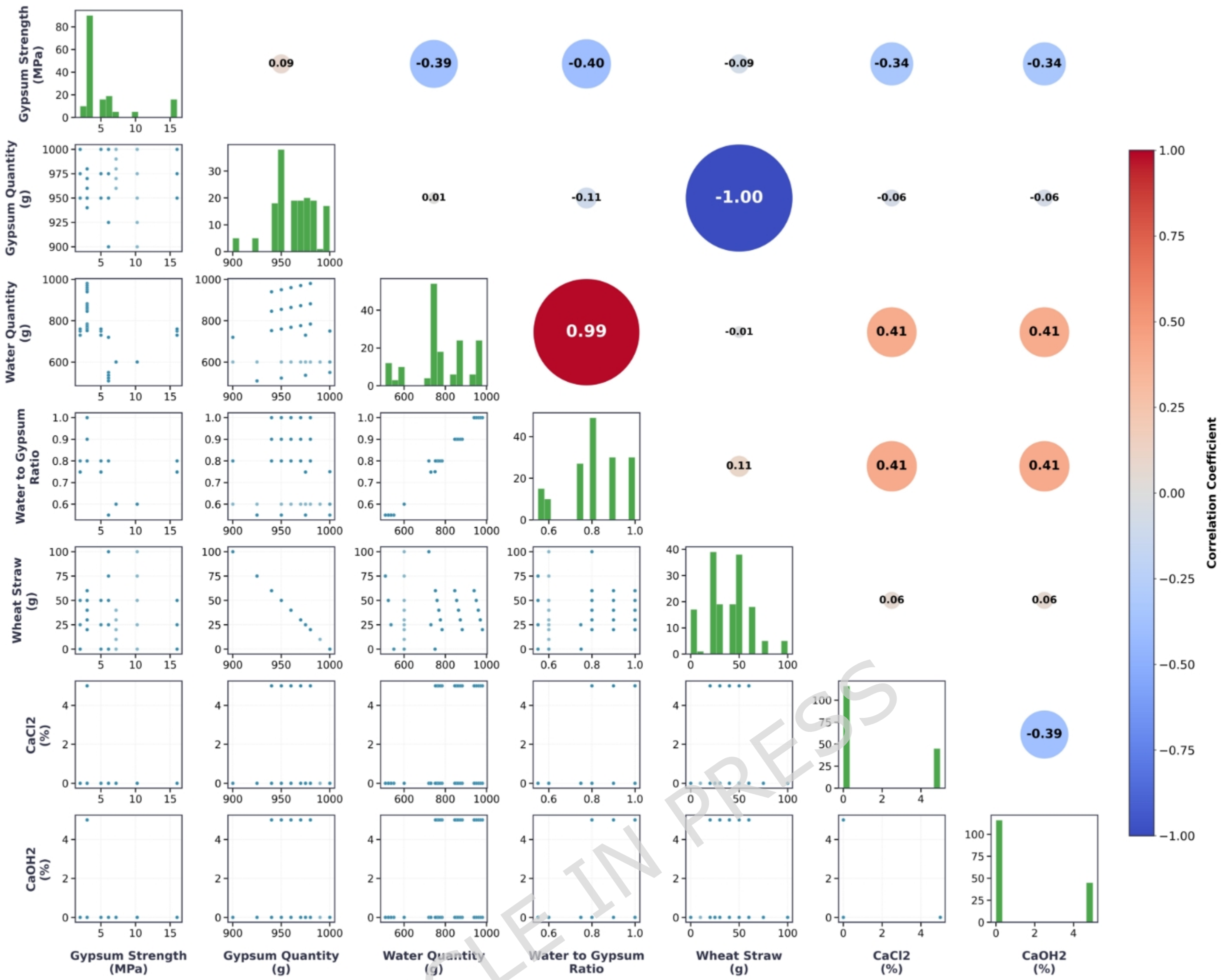


Figure 4. Summary of the correlation between input parameters.

The strongest relationship observed is the nearly perfect positive correlation between water quantity and the water-to-gypsum ratio (correlation coefficient = 0.99), reflecting the direct dependency of the water-to-gypsum ratio on water content across the mixtures. A similarly strong negative correlation is observed between water quantity and gypsum quantity (correlation coefficient = –1.00), due to the fixed mixture mass in several formulations. $CaCl_2$ and $Ca(OH)_2$ show a moderate positive correlation with both water quantity and the water-to-gypsum ratio (correlation coefficient = 0.41). Water quantity and water-to-gypsum ratio show a moderate negative correlation of around –0.40 with gypsum strength, indicating that increasing the water-to-gypsum ratio and water content introduces higher porosity and weaker bonding, thereby reducing gypsum strength. $CaCl_2$ and $Ca(OH)_2$ exhibit weak correlations with most parameters because the additives were included only in selected batches at a constant level of 5%. $CaCl_2$ and $Ca(OH)_2$ show a moderate negative correlation with gypsum strength (correlation coefficient = –0.34), suggesting moderate dependence.

Overall, the correlation matrix indicates that water-related parameters (water content and water-to-gypsum ratio) are the dominant drivers of variation within the dataset, while wheat straw content contributes secondarily. The additive variables display minimal statistical coupling due to their selective inclusion. These results provide a quantitative overview of parameter interactions relevant to modeling gypsum composite behavior.

2.2.2. Correlation of input parameters with compressive strength

The individual relationships between compressive strength and key mixture parameters, including gypsum strength, gypsum quantity, water quantity, water-to-gypsum ratio, wheat straw content, $CaCl_2$ content, and $Ca(OH)_2$ content, are shown in Figure 5. Scatter plots are accompanied by linear regression trends, along with shaded regions representing the 95% confidence interval and 95% prediction interval. A moderate positive correlation is observed between compressive strength and gypsum strength (correlation coefficient = 0.64), indicating that gypsum strength significantly enhances the mechanical performance of the composites. Gypsum quantity also exhibits a moderate positive correlation (0.48), suggesting its contribution to gypsum strength. In contrast, water quantity, water-to-gypsum ratio, wheat straw content, $CaCl_2$ content, and $Ca(OH)_2$ content show negative correlations ranging from –0.24 to –0.48, as shown in Figure 5, implying that with the increase in water quantity, water-to-gypsum ratio, and chemical additives negatively affect the compressive strength, with the possible reason of the increasing porosity or disrupting the gypsum matrix [52, 53].

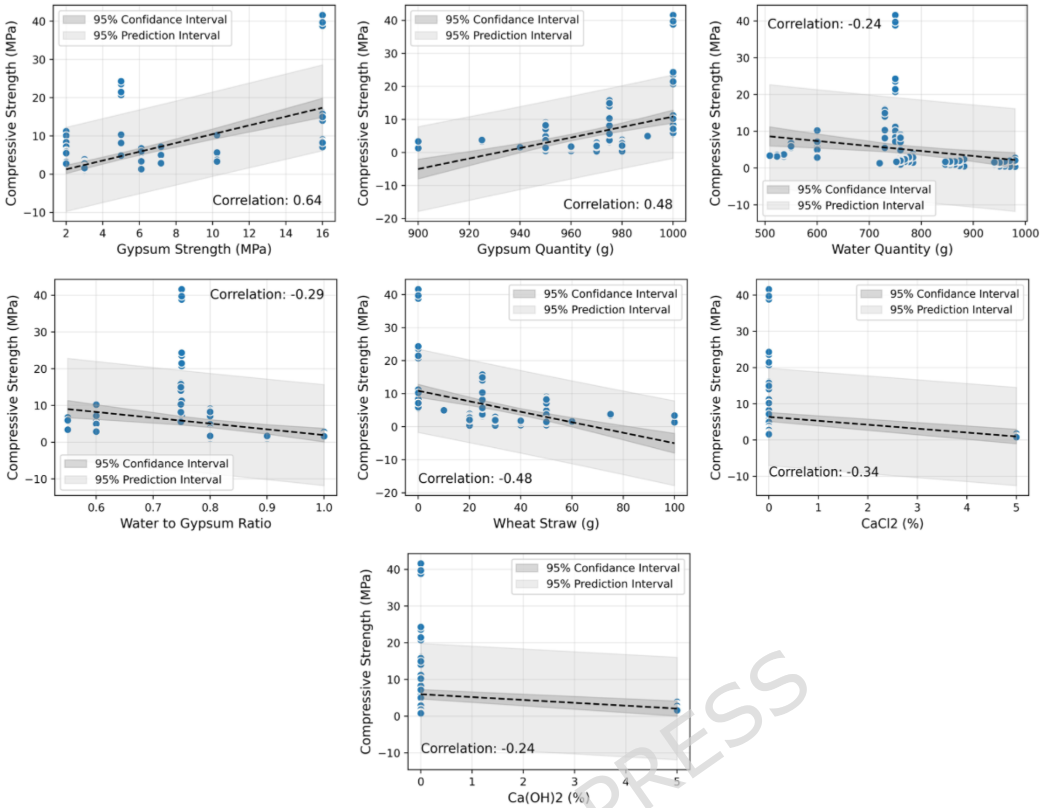


Figure 5. Correlation of input parameters with compressive strength.

2.2.3. Correlation of input parameters with flexural strength

The dependence of flexural strength on the mixture constituents of gypsum composites is shown in Figure 6. Similar to the above discussion, each subplot shows experimental data points along with linear regression lines, complemented by shaded regions representing the 95% confidence and prediction intervals. A strong positive correlation is observed between flexural strength and gypsum strength (correlation coefficient = 0.72), indicating that material strength markedly enhances flexural performance. Gypsum quantity also exhibits a moderate positive correlation (0.49), suggesting its contribution to improved flexural capacity. Conversely, water quantity, water-to-gypsum ratio, wheat straw content, $CaCl_2$ content, and $Ca(OH)_2$ content demonstrate moderate to strong negative correlations ranging from –0.36 to –0.53, as illustrated in Figure 6. These trends indicate that increased water content and wheat straw or additive dosages tend to reduce flexural strength, likely due to increased porosity, weakened interfacial bonding [52-54]. Overall, it highlights the sensitivity of flexural strength to mixture proportions and underscores the complex, multivariate interactions governing composite behavior.

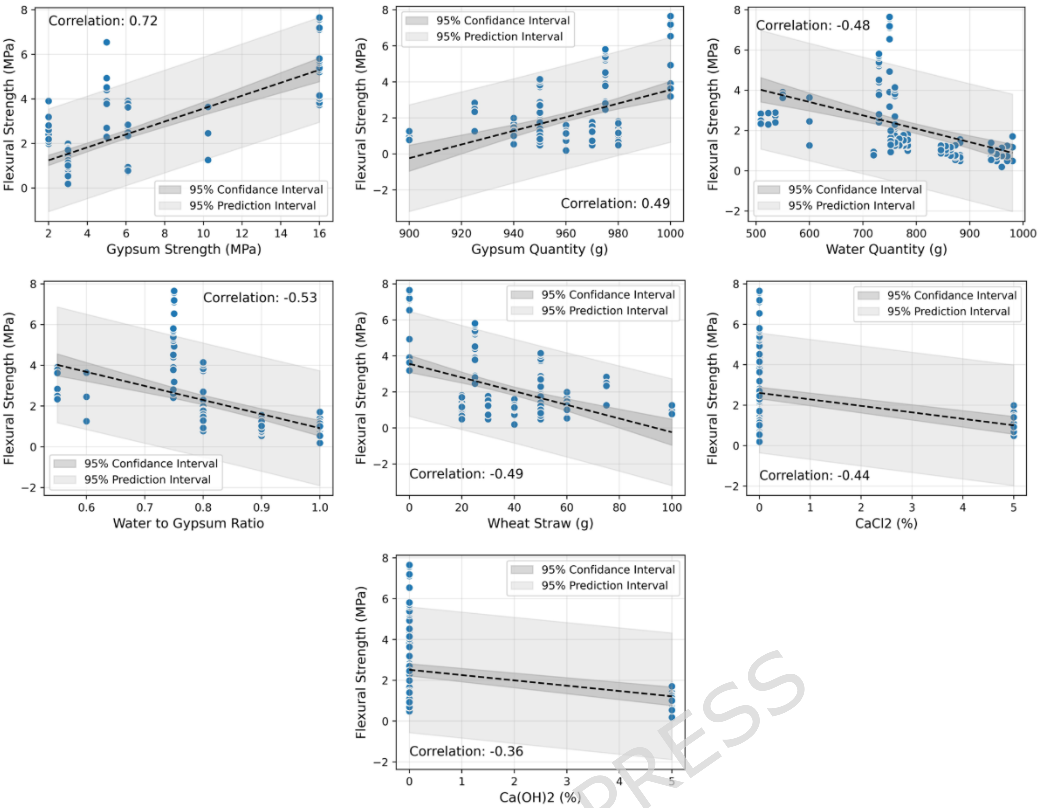


Figure 6. Correlation of input parameters with flexural strength.

2.2.4. Training and testing data

To assess the prediction performance of ML models, the dataset was randomly divided into training and testing subsets using a 70/30 ratio. This resulted in 113 and 106 samples for model training and 48 and 46 samples for independent testing in the compressive and flexural strength datasets, respectively. In addition, a 10-fold cross-validation procedure was applied to the training dataset, as illustrated in Figure 7, to optimize model hyperparameters. In this approach, the training data were partitioned into ten equal folds, where nine folds were used for training and one fold for validation, with the process repeated until each fold served as the validation set once. This combined strategy ensures reliable model development, provides evaluation of predictive performance.

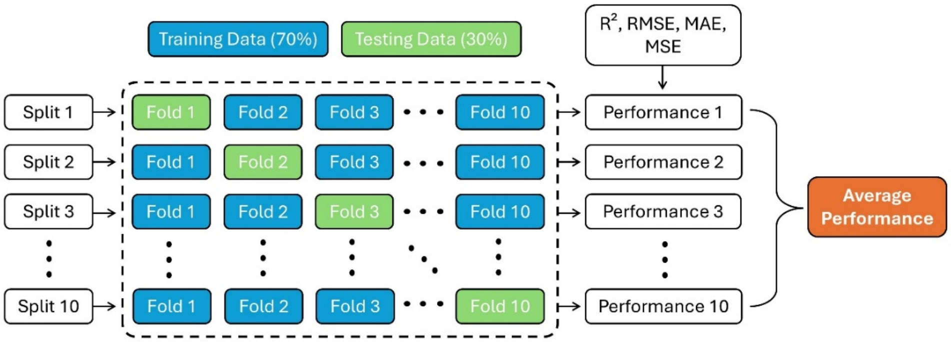


Figure 7. Flowchart of 10-fold cross-validation of data.

Figure 7 describes a 10-fold cross-validation approach that was employed with the five selected ML techniques to assess prediction performance with the aim of providing effective generalizability.

2.3. Theoretical background of adopted machine learning techniques

2.3.1. Artificial neural network

ANN [55] provides a powerful data-driven modeling framework capable of capturing complex behaviors between variables, which is widely used in the field of civil engineering [56-58]. The relationship between mixture components and compressive and flexural strength is multivariate and often difficult to express analytically. ANN developed in this study is a feedforward multilayer perceptron (MLP) consisting of an input layer, three hidden layers, and an output layer that predicts compressive and flexural strength, as shown in Figure 8.

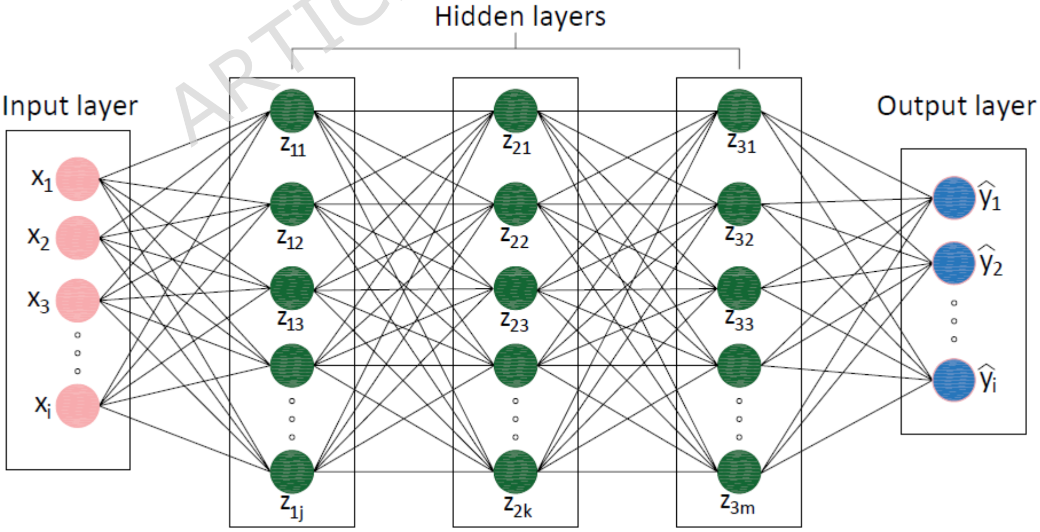


Figure 8. An example of an ANN structure.

Let the input vector be defined as:

$$x = [x_1, x_2, \ldots, x_i] \tag{1}$$

where x_i represents the □-th input variable (e.g., gypsum strength, water quantity, wheat straw content) comprising the input vector **x**. Each neuron performs a weighted summation of its inputs, followed by a nonlinear activation function. For the first hidden layer, the net input to the *j*-th neuron z_{1j} is given by:

$$z_{1j} = \sum_{i=1}^{n} w_{ij}^{(1)} x_i + b_j^{(1)} \tag{2}$$

$$h_{1j} = f(z_{1j}) \tag{3}$$

where, $w_{ij}^{(1)}$ and $b_j^{(1)}$ are the trainable weights and biases, $f(\cdot)$ is the activation function. For the second hidden layer, the neuron $z_{2k}$is expressed as:

$$z_{2k} = \sum_{j=1}^{n_2} w_{jk}^{(2)} h_{1j} + b_k^{(2)} \tag{4}$$

$$h_{2k} = f(z_{2k}) \tag{5}$$

Similarly, for the third hidden layer, the neuron z_{3m} is expressed as:

$$z_{3m} = \sum_{k=1}^{n_3} w_{km}^{(3)} h_{2k} + b_m^{(3)} \tag{6}$$

$$h_{3m} = f(z_{3m}) \tag{7}$$

Where n_1 and n_2 denote the number of neurons in the first and second hidden layers, respectively. The Rectified Linear Unit (ReLU) activation function was selected due to its computational efficiency and ability to model nonlinear relationships:

$$f(z) = \max(0, z) \tag{8}$$

The final output layer $\hat{y}_i$, representing the predicted mechanical property, is calculated as:

$$\hat{y}_i = \sum_{m=1}^{n_3} w_m^{(4)} h_{3m} + b^{(3)} \tag{9}$$

The loss function used for backpropagation is the Mean Squared Error (MSE):

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (10)$$

where y_i and $\hat{y}_i$ are the experimental and predicted outputs, respectively.

2.3.2. Gaussian process regression

GPR [59] is a non-parametric, probabilistic ML method that provides a framework for modeling complex nonlinear relationships and predicting the output with associated uncertainties, and is broadly used in civil engineering [60, 61]. The simplified architectural example of GPR is shown in Figure 9. In this study, GPR was employed to predict the mechanical properties of gypsum composites incorporating wheat straw fibers, based on mixture parameters where the underlying relationships are difficult to describe using conventional analytical models.

A Gaussian process (GP) is defined as a collection of random variables, any finite number of which follow a joint Gaussian distribution. A GP is fully specified by its mean function $m(x)$ and covariance (kernel) function $k(x,x')$ containing the input vectors x and x', and is expressed as:

$$f(x) \sim GP\ (m(x),\ k(x,x')) \qquad (11)$$

where, $f(x)$ represents the latent function that captures the relationship between input (mixture) parameters and output response (compressive or flexural strength).

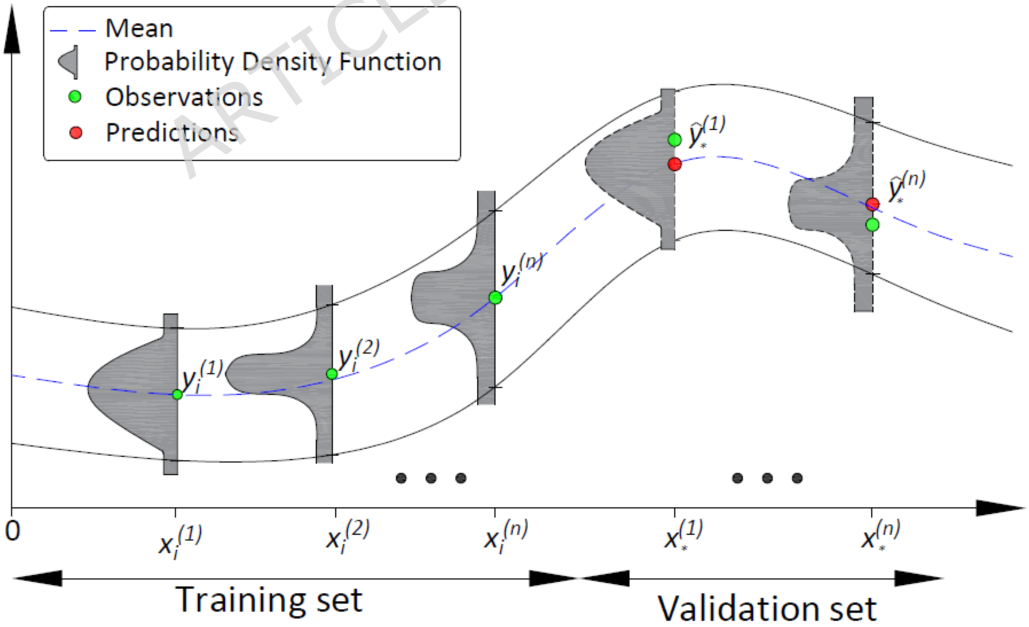


Figure 9. An example of a GPR architectural structure.

The mean function is commonly assumed to be zero without loss of generality:

$$m(x) = E[f(x)] \tag{12}$$

Given a training dataset $D = \{(x_i,y_i)\}_{i=1}^{N}$, the observed outputs are modeled as:

$$y_i = f(x) + \epsilon \tag{13}$$

$$\epsilon \sim N\left(0,\ \sigma_n^2\right) \tag{14}$$

where, ϵ represents independent Gaussian noise with variance σ_n^2. The covariance between input points is governed by the kernel function, which includes assumptions about smoothness and correlation in the data. In this study, the Radial Basis Function (RBF) kernel with a white noise kernel was adopted:

$$k(x,x') = \sigma_f^2 .\ \exp\left(-\frac{d(x,x')^2}{2l^2}\right) \tag{15}$$

where, σ_f^2 is the signal variance, $d(.,.)$ is the Euclidean distance and l is the length scale of the kernel. The white kernel was used to represent the variation in compressive or flexural strengths, and is represented as:

$$k(x,x') = noise\_level\ \mathit{if}\ x == x'\ \mathit{else\ 0} \tag{16}$$

For a new input point x_*, the joint distribution of training outputs y and predicted output $\hat{y}_*$ is given by:

$$\begin{pmatrix} y \\ \hat{y}_* \end{pmatrix} \sim N\left(0, \begin{bmatrix} K + \sigma_n^2 I & k_* \\ k_*^T & k(x_*,x_*) \end{bmatrix}\right) \tag{17}$$

where K is the covariance matrix of the training inputs and k_* is the covariance vector between the test input and training inputs, and I is an identity matrix. The predictive mean and variance are computed as:

$$\hat{y}_* = k_*^T\left[K + \sigma_n^2 I\right]^{-1} y \tag{18}$$

$$\sigma_*^2 = k(x_*,x_*) - k_*^T\left[K + \sigma_n^2 I\right]^{-1} k_* \tag{19}$$

Equation (18) presents the predictive mean, and Equation (19) provides the associated uncertainty as covariance.

2.3.3. Random forest

RF [62] is an ensemble learning method based on decision trees that enables the prediction by combining multiple weak learners into a single strong model, as illustrated in Figure 10. Owing to its ability to capture complex relationships and interactions among variables, RF has been commonly applied in civil engineering [63, 64] for prediction, such as the mechanical properties of composite materials. RF operates by constructing a large number of decision trees during training and aggregating their predictions through averaging (for regression problems), as shown in Figure 10. This ensemble strategy reduces model variance and mitigates overfitting compared to a single decision tree.

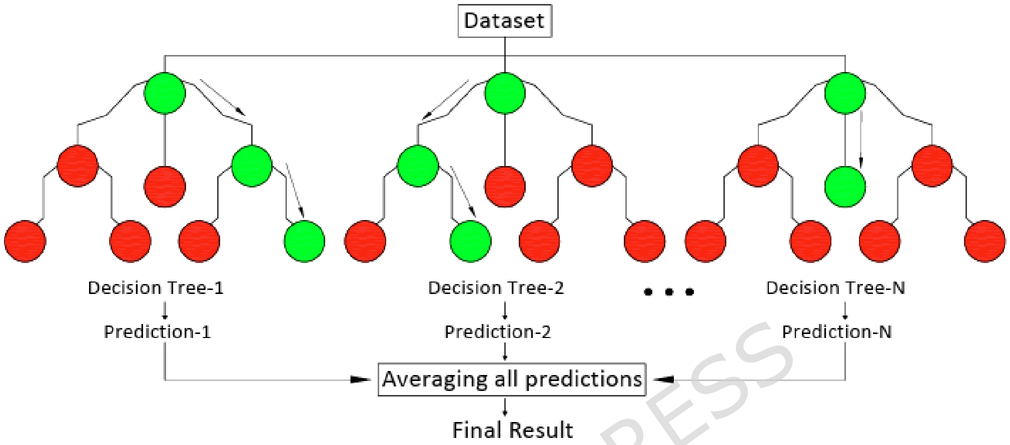


Figure 10. An example of an RF architectural structure.

Each decision tree divides the input data step by step into smaller groups so that samples with similar output values fall into the same group. For regression problems, this division is guided by minimizing the MSE, which measures how close the predicted values are to the actual values.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (20)$$

where y_i represents the target value and $\hat{y}_i$ is the mean response in the node. At each split, the feature and threshold that yield the maximum reduction in MSE are selected. The prediction of an individual tree $T_t(x)$ is the average of the target values in the terminal node corresponding to the input vector x. Given a training dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$, RF constructs T decision trees using bootstrap sampling. Each tree is trained on a randomly drawn subset of data, while a random subset of input features is considered at each split. RF regression prediction is obtained by averaging the predictions of all individual trees:

$$\hat{y}(x) = \frac{1}{T}\sum_{t=1}^{T} T_t(x) \qquad (21)$$

where $T_t(x)$ denotes the prediction from the t-th decision tree.

2.3.4. Extreme gradient boosting

XGBoost [65] is an advanced ensemble learning algorithm based on the gradient boosting framework. It builds a predictive model by sequentially combining multiple weak learners, typically decision trees, where each new tree is trained to correct the errors made by the previous ones, as shown in Figure 11. Due to its high predictive accuracy and computational efficiency, XGBoost has also been extensively used in civil engineering applications [66, 67] for modeling complex relationships.

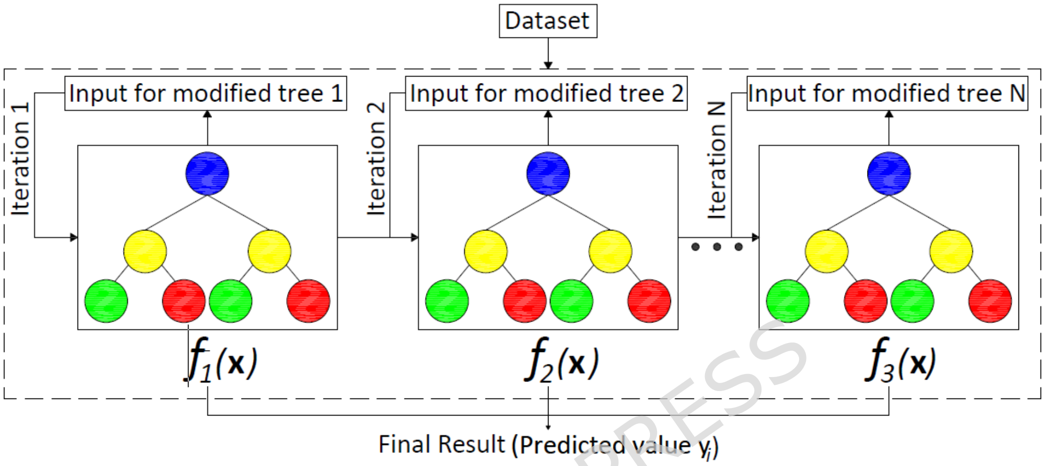


Figure 11. An example of an XGBoost architectural structure.

Generally, XGBoost is employed to predict the output parameters, where interactions among parameters are challenging to model using traditional regression techniques. Unlike RF, where trees are built independently, XGBoost constructs decision trees sequentially. At each iteration, a new tree is added to minimize the prediction error of the existing ensemble. The overall prediction after k trees are given by:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x) \tag{22}$$

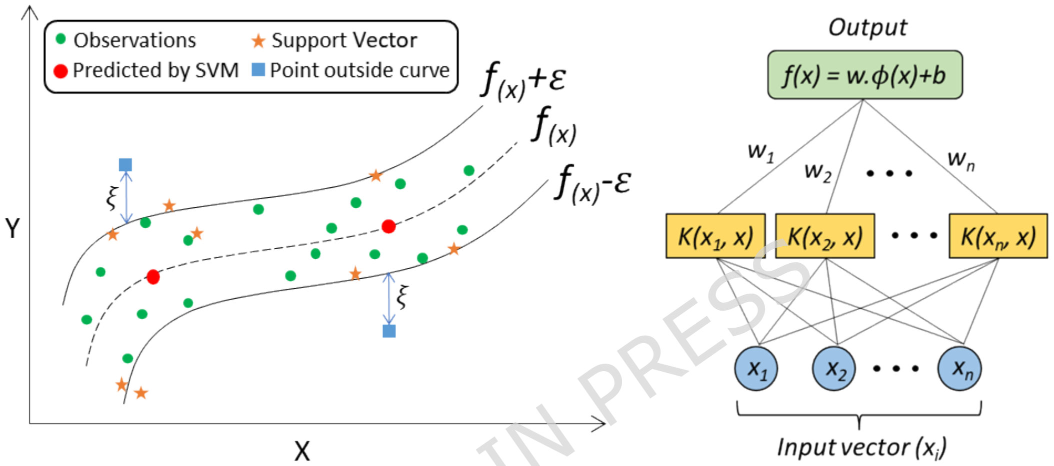
where f_k represents an individual decision tree and x is the input feature vector. XGBoost optimizes a regularized objective function that balances prediction accuracy and model complexity:

$$L = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{23}$$

where $l(\cdot)$ is typically the MSE for regression problems, and $\Omega(f_k)$ penalizes complex trees to reduce overfitting. Tree construction is accelerated using first and second-order derivatives of the loss function, enabling efficient identification of optimal splits. Due to its high accuracy, built-in regularization, and computational efficiency, XGBoost is well-suited for predicting the mechanical properties of gypsum-based composites.

2.3.5. Support vector machine

SVM [68] is a supervised learning algorithm widely used for regression and classification problems due to its strong generalization capability, particularly for nonlinear and high-dimensional datasets [69, 70]. Figure 12 presents the schematic representation of SVM. The blue curve represents the predicted regression function, while the dashed lines indicate the □-insensitive tube, as shown in Figure 12(a). Data points on the tube boundaries act as support vectors and govern the model, whereas points outside the tube are through slack variables, allowing controlled prediction errors.



(a) (b)

Figure 12. The schematic representation of SVM presenting the regression function with an insensitive tube (a) and mapping of input parameters (b).

SVM aims to determine a function that approximates the relationship between input parameters and the target response while maintaining model simplicity, as presented in Figure 12(b). The regression function is defined as:

$$f(x) = w^{T}\phi(x) + b \qquad (24)$$

where $\phi(x)$ maps the input data into a higher-dimensional feature space, w_n is the weight vector, and b is the bias term. The objective of SVM is to minimize the model complexity while ensuring that prediction errors lie within a predefined tolerance ε. This is achieved by solving the following optimization problem:

$$\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \qquad (25)$$

subject to:

$$y_i - f(x_i) \le \varepsilon + \xi_i$$
$$f(x_i) - y_i \le \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0$$

where □ is the penalty parameter controlling the trade-off between model complexity and training error, and ξ_i, ξ_i^* are slack variables. To handle complex relationships, SVM employs kernel functions, with the RBF kernel commonly used:

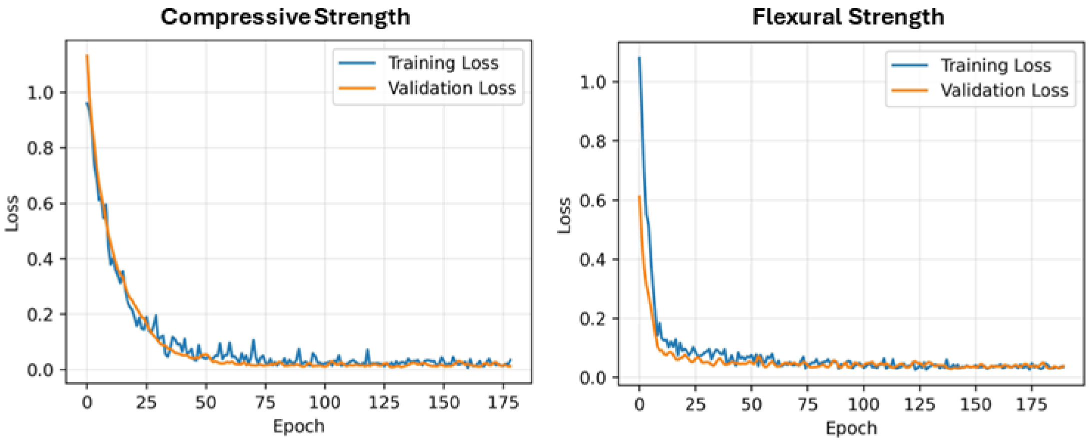
$$K(x,x') = \exp(- \gamma\|x,x'\|^2) \tag{26}$$

By using appropriate kernel functions, SVM can effectively model interactions between mixture parameters and mechanical properties.

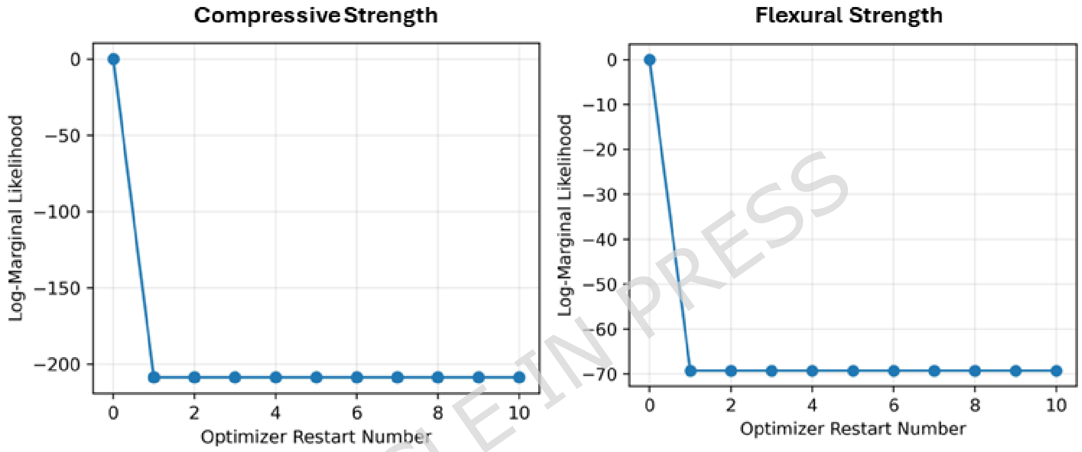
2.4. Optimization and learned hyperparameters

Although the predictive performance of the selected ML models was evaluated using multiple statistical indicators, a detailed discussion is provided in Section 3.1.1, including the coefficient of determination (R^2), root mean square error (RMSE), mean squared error (MSE), and mean absolute error (MAE), the convergence behavior of each model was illustrated using a single representative evaluation metric. Specifically, loss functions (MSE) were used for ANN, log-marginal likelihood for GPR, and R^2 score for RF, XGBoost, and SVM models, as these metrics are directly optimized during the training process of the respective algorithms. The summary of the convergence behavior of ANN, GPR, RF, XGBoost, and SVM models for predicting compressive and flexural strength is shown in Figure 13.

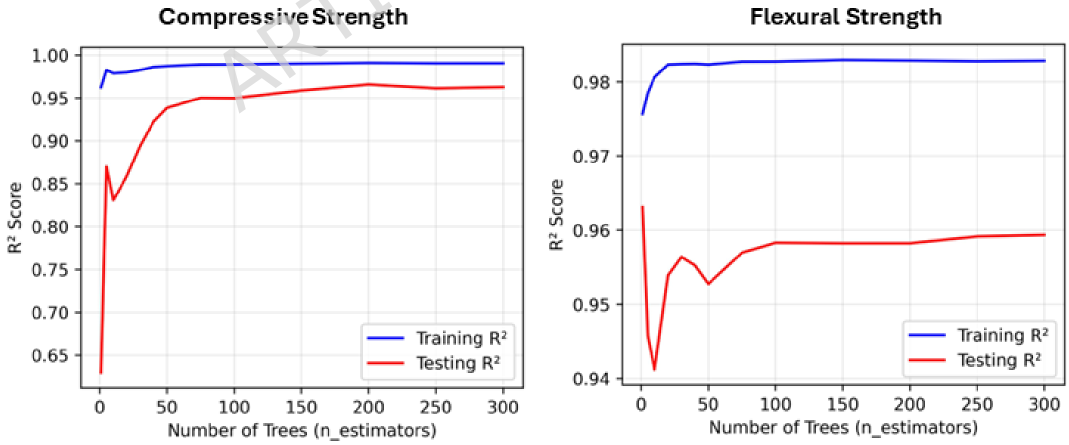
For ANN model, the training and validation loss curves exhibit a rapid decrease during initial epochs, followed by stabilization at low loss values, as shown in Figure 13(a), indicating efficient learning and good convergence without evident overfitting for both output responses. In the case of GPR, the log-marginal likelihood reaches a stable optimum within a small number of optimizer restarts, demonstrating kernel hyperparameter optimization and consistent convergence behavior, as shown in Figure 13(b). For RF model, the training and testing R^2 scores improve progressively with an increasing number of trees and eventually plateau, indicating that model performance stabilizes once a sufficient ensemble size is reached, as illustrated in Figure 13(c). For XGBoost model, the testing and training R^2 scores increase rapidly with the number of boosting rounds and reach stable, near-optimal values after a limited number of iterations, indicating fast convergence and strong generalization performance, as presented in Figure 13(d). In contrast, SVM model exhibits a gradual improvement in prediction accuracy with increasing training set size, as shown in Figure 13(e) by the learning curves. Overall, the convergence trends confirm the stability, robustness, and reliable learning behavior of the employed ML models for both compressive and flexural strength prediction.



(a) ANN

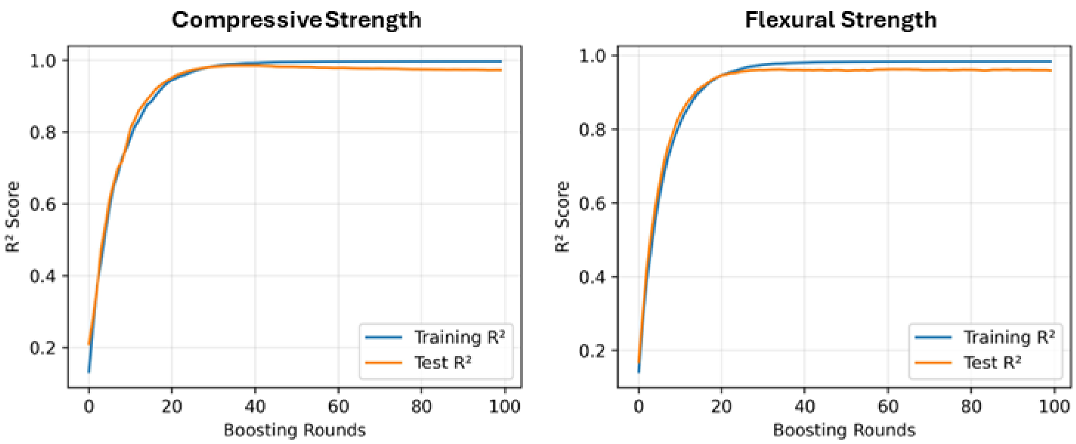


(b) GPR

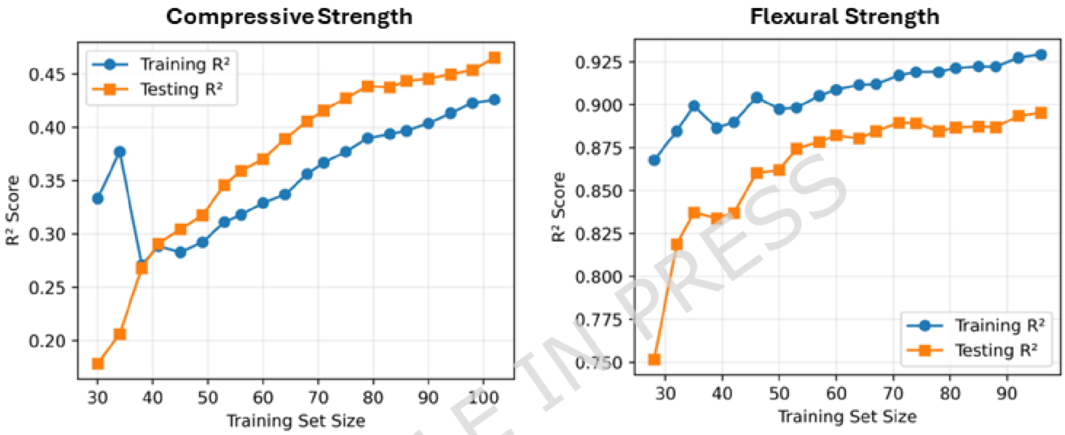


(c) RF

Figure 13. *(Continue)*



(d) XGBoost



(e) SVM

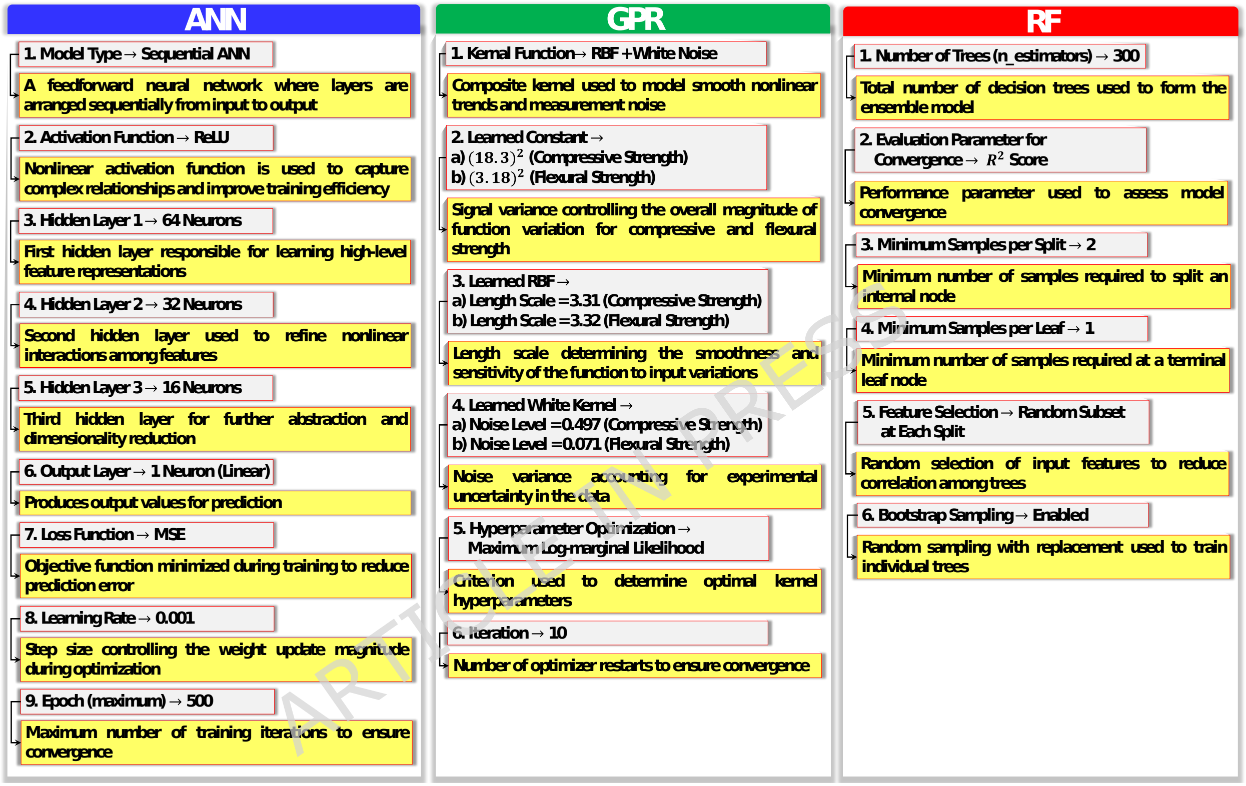
Figure 13. Hyperparameter convergence plots of models.

Figure 14 summarizes the key learning hyperparameters and architectural configurations of ANN, GPR, RF, XGBoost, and SVM models. ANN was implemented as a sequential feedforward network comprising three hidden layers with 64, 32, and 16 neurons, respectively, using the ReLU activation function, while the output layer adopted a linear activation, and the model was trained and achieved convergence using the MSE loss function, as shown in Figure 13(a).

GPR model utilized a composite kernel consisting of an RBF and a white noise kernel, with hyperparameters optimized automatically by maximizing the log-marginal likelihood over multiple iterations, as shown in Figure 13(b). The learned constant represents the overall signal variance and controls the global magnitude of variation in target response, reflecting the variability of compressive and flexural strength across the dataset. RBF kernel, characterized by its length scale parameter, captures the smoothness and correlation of the response with respect to changes in input parameters; larger length scale values indicate smoother variations and stronger correlations among neighboring data points, as illustrated in Figure 14.

Additionally, white noise kernel accounts for measuring experimental uncertainty present in the data.

RF model was configured with 300 decision trees, bootstrap sampling, and default split and leaf criteria (2, 1). Model convergence was assessed using R^2, which exhibited progressive improvement with an increasing number of trees before stabilizing, as illustrated in Figure 13(c). This indicated that the ensemble size was sufficient to capture the underlying relationships without overfitting. The use of bootstrap aggregation and random feature selection at each split further enhanced model performance and reduced variance.



ANN

1. Model Type → Sequential ANN
   - A feedforward neural network where layers are arranged sequentially from input to output
2. Activation Function → ReLU
   - Nonlinear activation function is used to capture complex relationships and improve training efficiency
3. Hidden Layer 1 → 64 Neurons
   - First hidden layer responsible for learning high-level feature representations
4. Hidden Layer 2 → 32 Neurons
   - Second hidden layer used to refine nonlinear interactions among features
5. Hidden Layer 3 → 16 Neurons
   - Third hidden layer for further abstraction and dimensionality reduction
6. Output Layer → 1 Neuron (Linear)
   - Produces output values for prediction
7. Loss Function → MSE
   - Objective function minimized during training to reduce prediction error
8. Learning Rate → 0.001
   - Step size controlling the weight update magnitude during optimization
9. Epoch (maximum) → 500
   - Maximum number of training iterations to ensure convergence

GPR

1. Kernal Function → RBF + White Noise
   - Composite kernel used to model smooth nonlinear trends and measurement noise
2. Learned Constant →
   a) $(18.3)^2$ (Compressive Strength)
   b) $(3.18)^2$ (Flexural Strength)
   - Signal variance controlling the overall magnitude of function variation for compressive and flexural strength
3. Learned RBF →
   a) Length Scale = 3.31 (Compressive Strength)
   b) Length Scale = 3.32 (Flexural Strength)
   - Length scale determining the smoothness and sensitivity of the function to input variations
4. Learned White Kernel →
   a) Noise Level = 0.497 (Compressive Strength)
   b) Noise Level = 0.071 (Flexural Strength)
   - Noise variance accounting for experimental uncertainty in the data
5. Hyperparameter Optimization → Maximum Log-marginal Likelihood
   - Criterion used to determine optimal kernel hyperparameters
6. Iteration → 10
   - Number of optimizer restarts to ensure convergence

RF

1. Number of Trees (n_estimators) → 300
   - Total number of decision trees used to form the ensemble model
2. Evaluation Parameter for Convergence → R^2 Score
   - Performance parameter used to assess model convergence
3. Minimum Samples per Split → 2
   - Minimum number of samples required to split an internal node
4. Minimum Samples per Leaf → 1
   - Minimum number of samples required at a terminal leaf node
5. Feature Selection → Random Subset at Each Split
   - Random selection of input features to reduce correlation among trees
6. Bootstrap Sampling → Enabled
   - Random sampling with replacement used to train individual trees

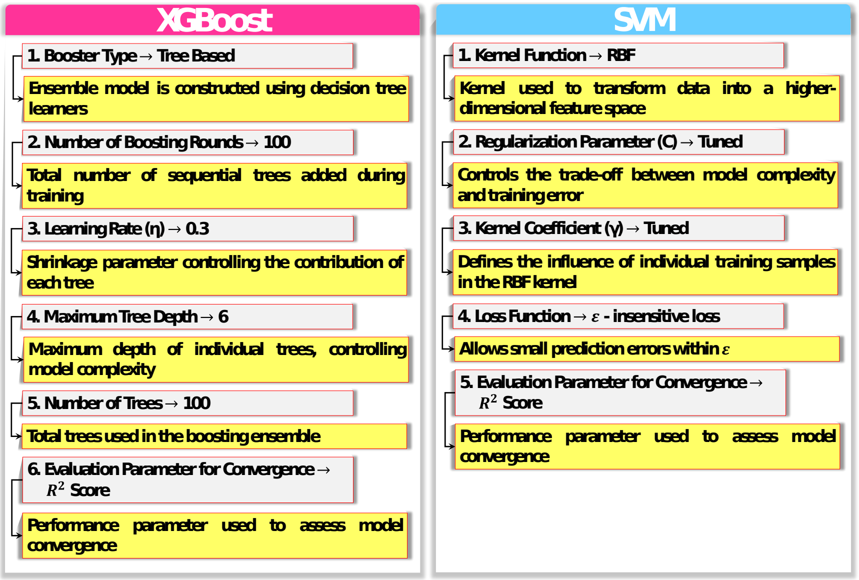


Figure 14. Learned hyperparameters of ML techniques for optimization.

XGBoost model was implemented using a tree-based booster, which constructs a sequence of decision trees in a sequential manner to progressively improve prediction accuracy. A learning rate of 0.3 was adopted to control the contribution of each newly added tree, enabling the model to learn complex patterns while maintaining stable convergence, as shown in Figure 13(d). The maximum tree depth was set to six, allowing the model to capture relationships and higher-order interactions among the mixture parameters without excessive model complexity. A total of 100 boosting rounds were employed, ensuring learning capacity while avoiding overfitting. Model convergence and performance stabilization were monitored using R^2.

For SVM model, an RBF kernel was adopted to capture relationships between input parameters and target mechanical properties. The regularization parameter (C) and kernel coefficient (□) were systematically tuned to obtain an optimal balance between prediction accuracy and model generalization. The parameter “□” controls how strongly the model penalizes prediction errors, whereas □ defines how far the influence of each data point extends in the feature space, as illustrated in Figure 14. Model convergence was assessed using R^2, ensuring that the selected hyperparameters yielded stable and reliable predictive performance on unseen data, as shown in Figure 13(e). Overall, the figure highlights the distinct learning mechanisms and parameter settings of each algorithm, which can provide sufficient understanding for the comparative analysis conducted in this study.

3. Results and discussion

3.1. Performance evaluation of models

3.1.1. Comparison of performance metrics using 10-fold cross-validation

To quantitatively assess the predictive performance of the selected ML models, four widely used statistical indicators were employed: the coefficient of determination (R^2), root mean squared error (RMSE), mean absolute error (MAE), and mean squared error (MSE). These metrics collectively provide a comprehensive evaluation of model accuracy. The R^2 measures the proportion of variance in the experimental data that is explained by the model predictions. A higher R^2 value indicates better agreement between predicted and observed values. The RMSE quantifies the standard deviation of prediction errors and is particularly sensitive to large deviations. Lower RMSE values reflect improved predictive precision and reduced dispersion of errors. The MAE represents the average size of the prediction errors, regardless of whether the model overestimates or underestimates the actual values. MAE provides a direct measure of model accuracy and is less sensitive to extreme errors compared to RMSE. The MSE evaluates the average squared difference between predicted and experimental values. These performance indicators are expressed as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y}_i)^2} \tag{27}$$

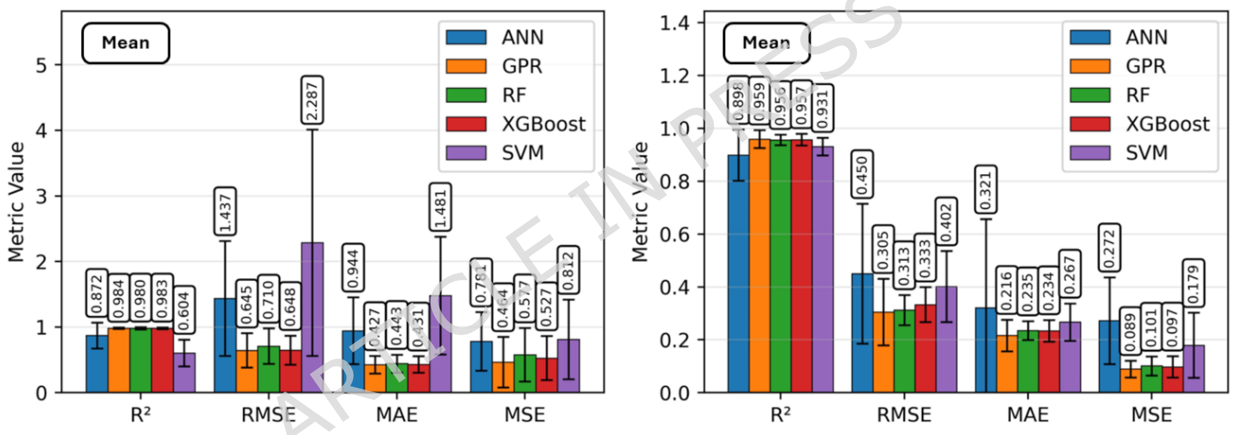
$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \tag{28}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{29}$$

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{30}$$

where y_i and $\hat{y}_i$are the experimental and predicted values, respectively, $\bar{y}_i$ is the mean of experimental values, and N is the number of samples.

The comparative predictive accuracy of models for compressive and flexural strength estimation using 10-fold cross-validation is shown in Figure 15. The bar plots represent the mean values of performance metrics across the ten folds, while error bars indicate the corresponding standard deviations, reflecting model stability. Among all models, for compressive strength predictions GPR model demonstrates the highest mean R^2 value along with the lowest RMSE, MAE, and MSE, indicating superior predictive accuracy. RF and XGBoost models exhibit performance metrics very close to those of GPR, ranking as the second most accurate model. In contrast, ANN shows moderate prediction accuracy, while SVM presents comparatively higher error values and variability across folds, as shown in Figure 15(a). For flexural strength prediction, the GPR model demonstrated the best overall performance, achieving the highest R^2 value and lowest RMSE, MAE, and MSE, which indicates relatively excellent stability and predictive accuracy, as illustrated in Figure 15(b). RF model followed closely as the second-best performer, exhibiting comparable accuracy with marginally higher error values than GPR. In contrast, ANN and XGBoost models provided moderate predictive capability, while SVM model showed relatively lower accuracy and greater variability across the cross-validation folds, as shown in Figure 15(b). These results highlight the robustness of GPR approach for compressive and flexural strength predictions, with XGBoost and RF, respectively, emerging as a reliable alternative under cross-validated conditions. In addition to 10-fold cross-validation, a 5-fold cross-validation was also conducted to examine model robustness. Although both approaches yielded comparable results, the 10-fold scheme demonstrated slightly superior predictive performance; therefore, it was adopted for subsequent analysis.



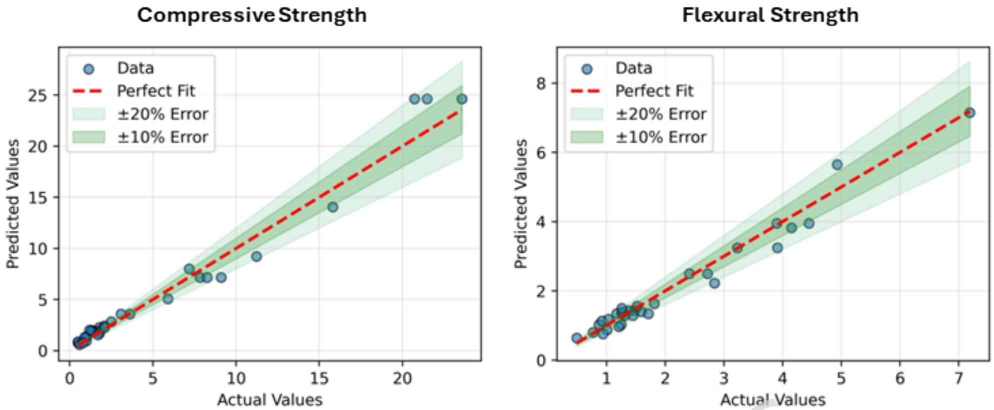
(a) Compressive Strength

(b) Flexural Strength

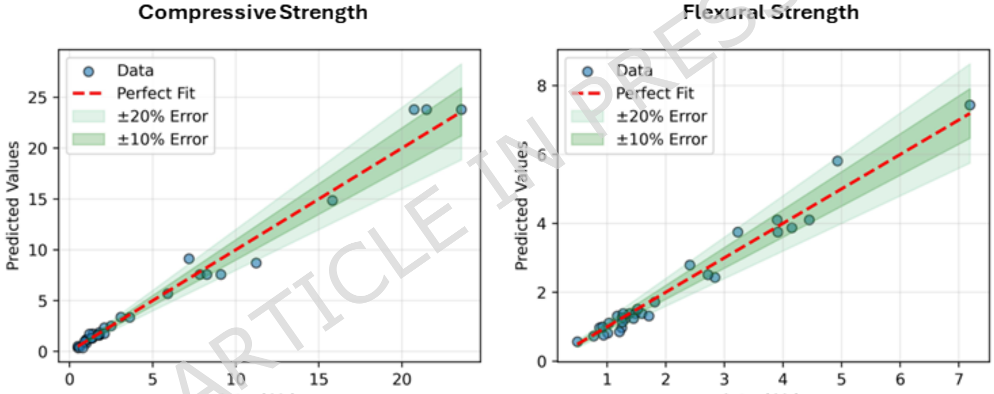
Figure 15. Comparison of the performance metrics of models obtained using 10-fold cross-validation.

3.1.2. Comparison between experimental and predicted values.

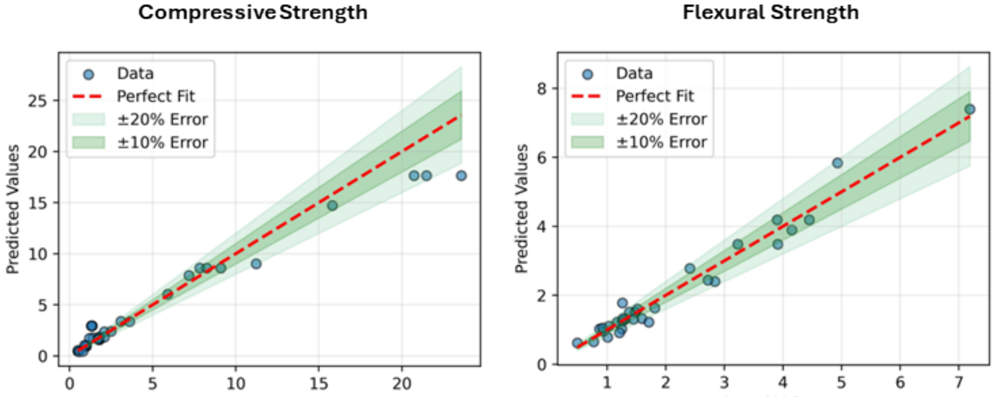
The comparison between actual values obtained from experiments and predicted values for compressive and flexural strength using ANN, GPR, RF, XGBoost, and SVM models is shown in Figure 16. In all subplots, the dashed red line represents the ideal line (perfect fit), while the shaded regions denote the ±10% and ±20% error bands around the perfect fit line, indicating tolerance ranges of prediction accuracy, respectively. A strong clustering of data points for low compressive and flexural strength values with relatively small errors indicates high predictive accuracy and low bias for low compressive and flexural strength. However, as the compressive and flexural strength values increase, a proportion of data points fall outside the ±10% band and approach or exceed the ±20% error limits, indicating reduced predictive accuracy at higher compressive and flexural strength values. This behavior suggests that the models are more reliable in predicting lower ranges of the output parameters, while prediction errors increase at higher compressive and flexural strength values, likely due to the limited number of high-strength samples and greater variability in material behavior in this region.



(a) ANN

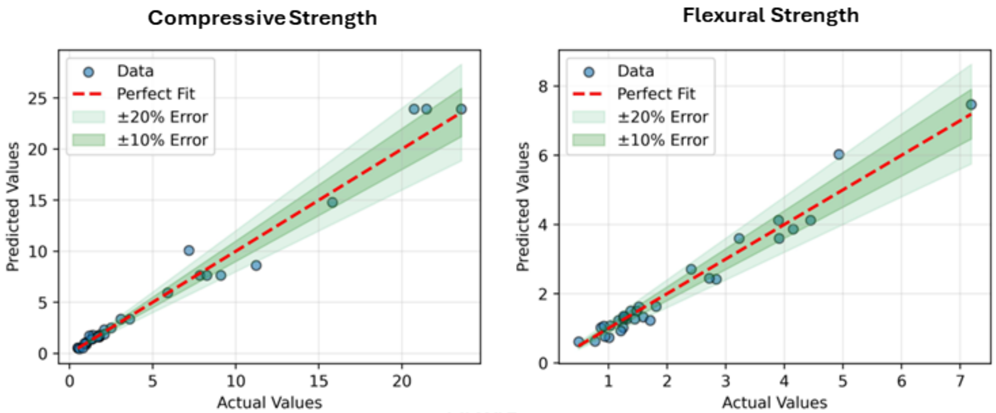


(b) GPR

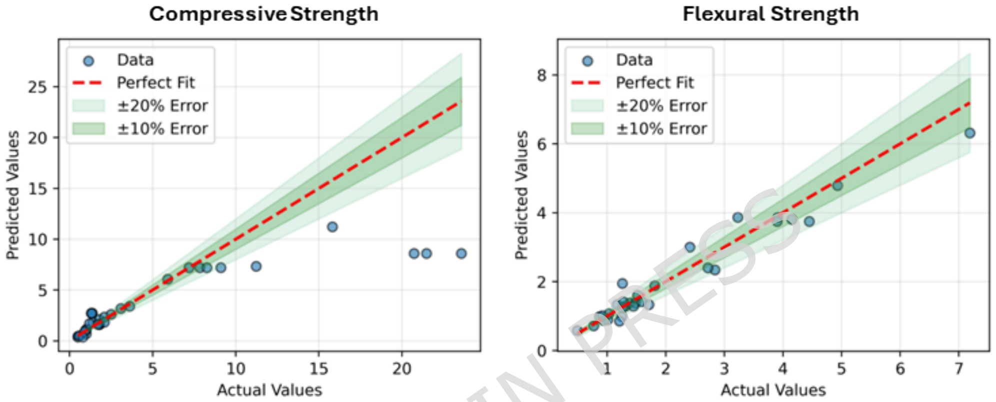


(c) RF

Figure 16. *(Continue)*



(d) XGBoost



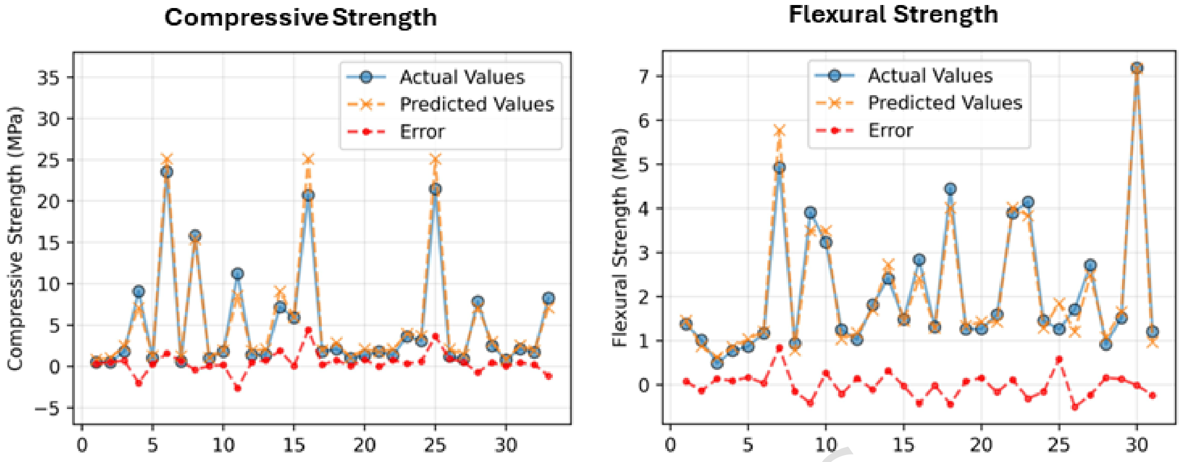
(e) SVM

Figure 16. Comparison between experimental and predicted values of compressive and flexural strength obtained using models.

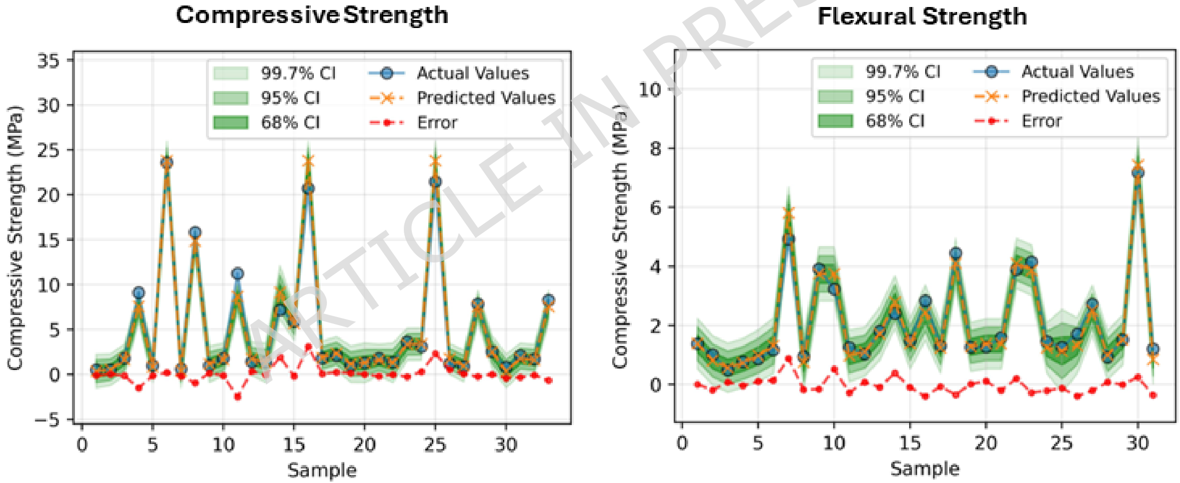
3.1.3. Prediction error analysis

The experimental and predicted values of compressive and flexural strength obtained using ANN, GPR, RF, XGBoost, and SVM models, along with the corresponding prediction errors for individual test samples, are shown in Figure 17. For all models, the predicted trends generally follow experimental values, indicating that the developed models can effectively capture the underlying relationship between input parameters and mechanical properties. However, noticeable differences in prediction accuracy and error dispersion can be observed among the models. ANN, RF, and XGBoost models show moderate fluctuations around the experimental data, with occasional larger deviations, particularly at higher strength values, as shown in Figure 17(a), Figure 17(c), and Figure 17(d). SVM model exhibits comparatively larger prediction errors and higher variability, reflecting its relatively lower predictive capability, as illustrated in Figure 17(e).

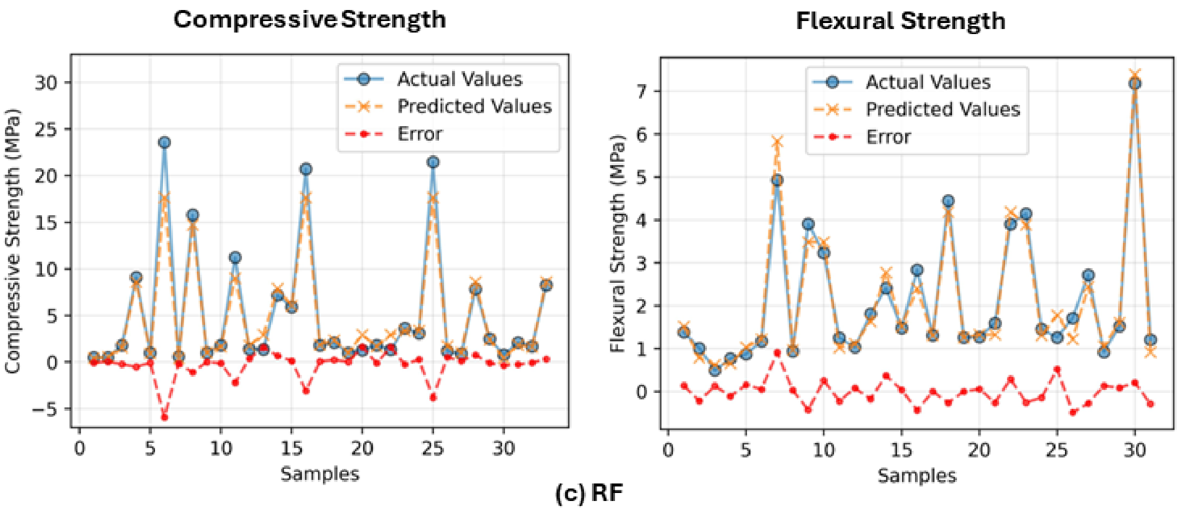
GPR model demonstrates superior performance, characterized by close alignment between experimental and predicted values and consistently smaller error magnitudes across samples, as shown in Figure 17(b). A key advantage of GPR is its ability to quantify predictive uncertainty through the confidence interval (CI). The CI was constructed from the predictive mean (μ) and standard deviation (σ) obtained from the GPR posterior, such that μ ±1 σ, μ ± 2σ, and μ ± 3σ, representing approximately 68%, 95%, and 99.7% CI, respectively, assuming a Gaussian predictive distribution.



(a) ANN

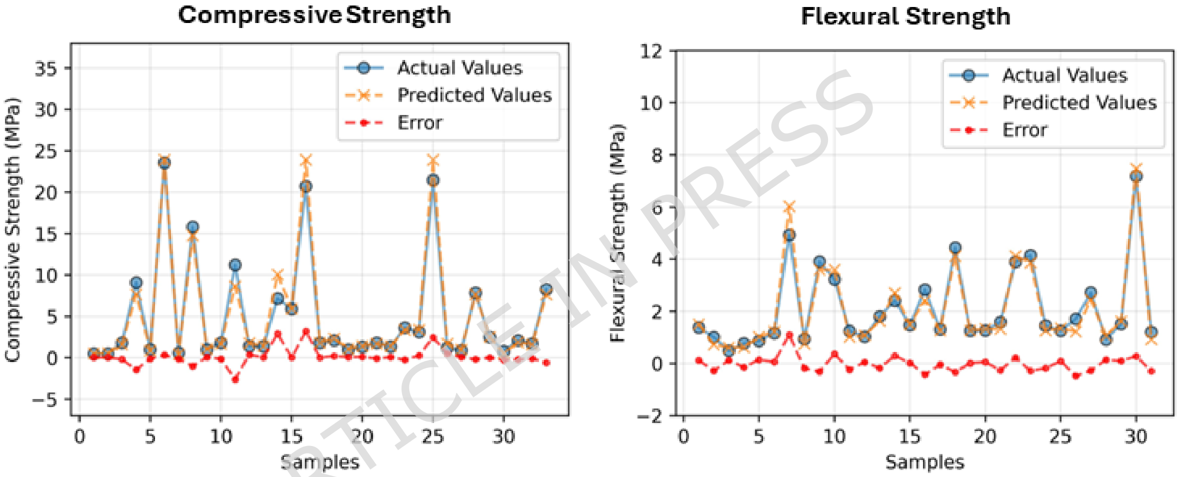


(b) GPR

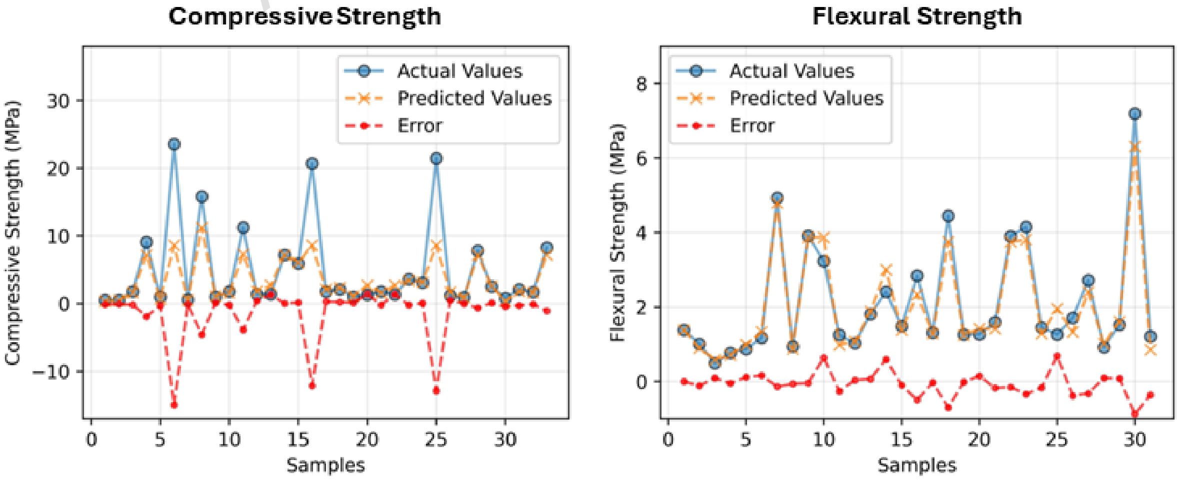


(c) RF

Figure 17. *(Continue)*



(d) XGBoost



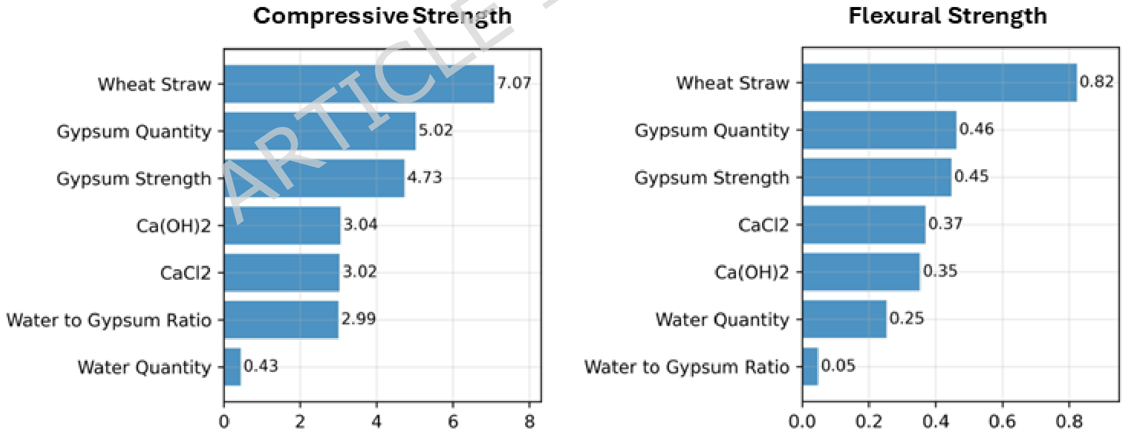
(e) SVM

Figure 17. Comparison of experimental and prediction values with error quantification.

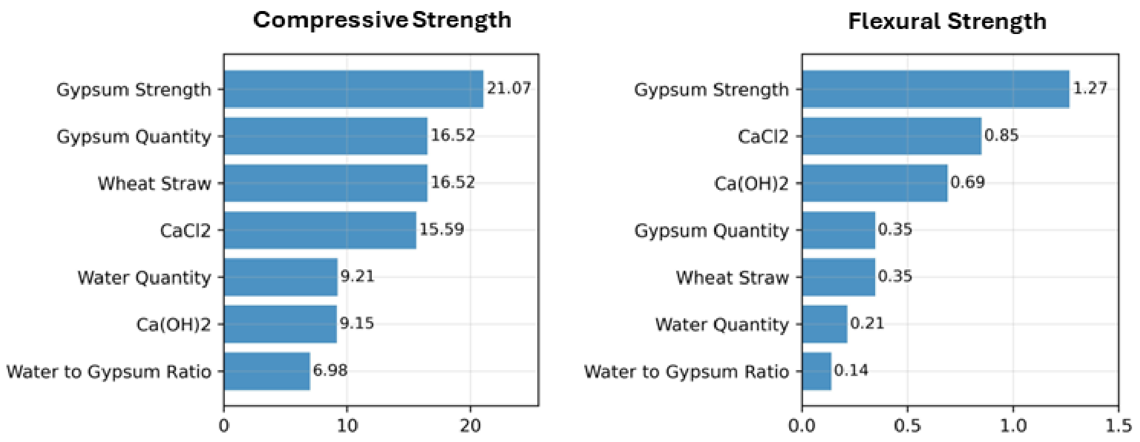
Most experimental data points fall within this interval, indicating reliable uncertainty estimation. This feature allows GPR not only to provide point predictions but also to assess associated uncertainties in experimental data. This makes it unique from other ML techniques and is particularly valuable for engineering applications involving material design and performance assessment [71]. Overall, the results highlight the robustness and reliability of GPR, while also confirming the effectiveness of other methods, such as RF and XGBoost, and comparatively limited performance of SVM for the present dataset.

3.2. Importance feature analysis and physical interpretation

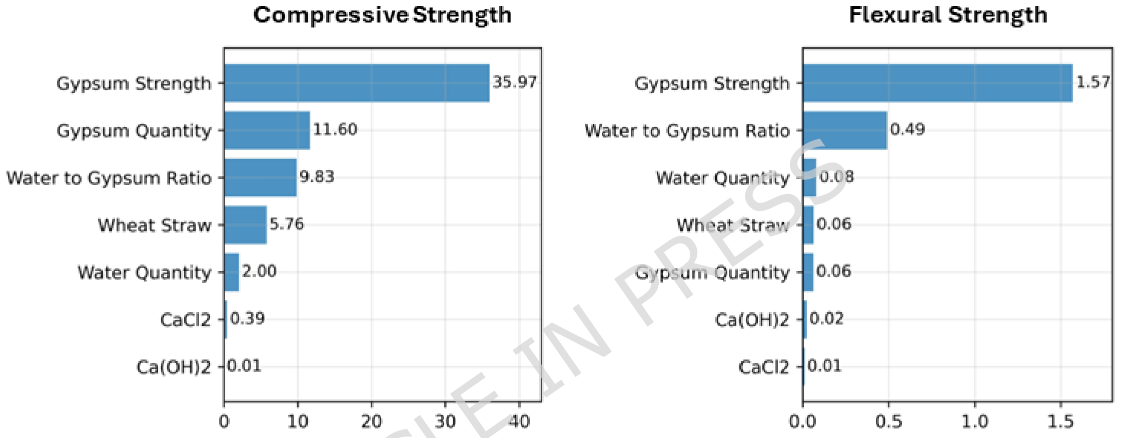
The permutation-based feature importance results for compressive and flexural strength prediction were obtained using ANN, GPR, RF, XGBoost, and SVM models are presented in Figure 18. Across GPR, RF, and XGBoost for both mechanical properties, gypsum strength consistently emerged as the most influential parameter, exhibiting the highest importance scores. However, in the ANN model, wheat straw emerged as the most influential parameter for both mechanical properties, while in the SVM model, $CaCl_2$ was the most influential parameter for compressive strength and gypsum strength for flexural strength. This highlights the dominant role of the different parameters in models in governing both compressive and flexural performance.



(a) ANN

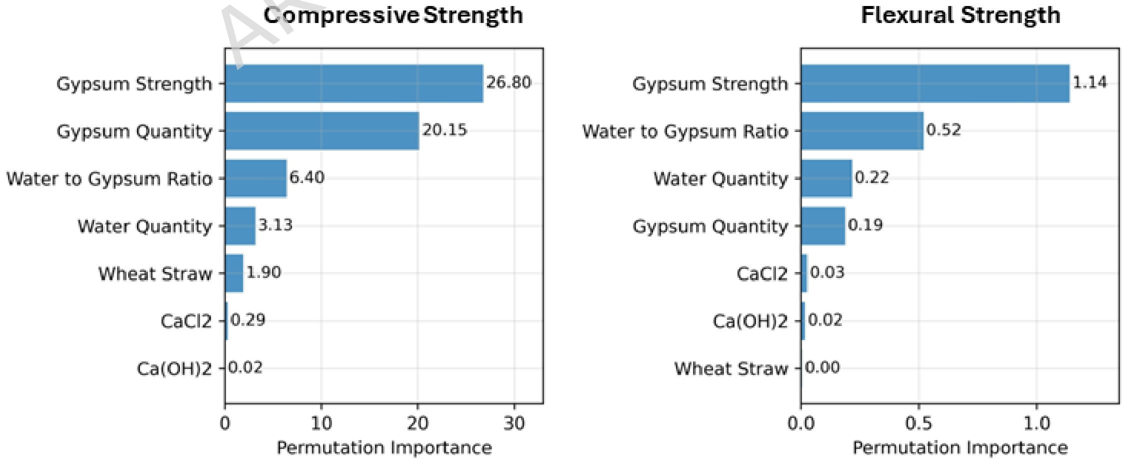


(b) GPR



(c) RF

Figure 18. *(Continue)*



(d) XGBoost

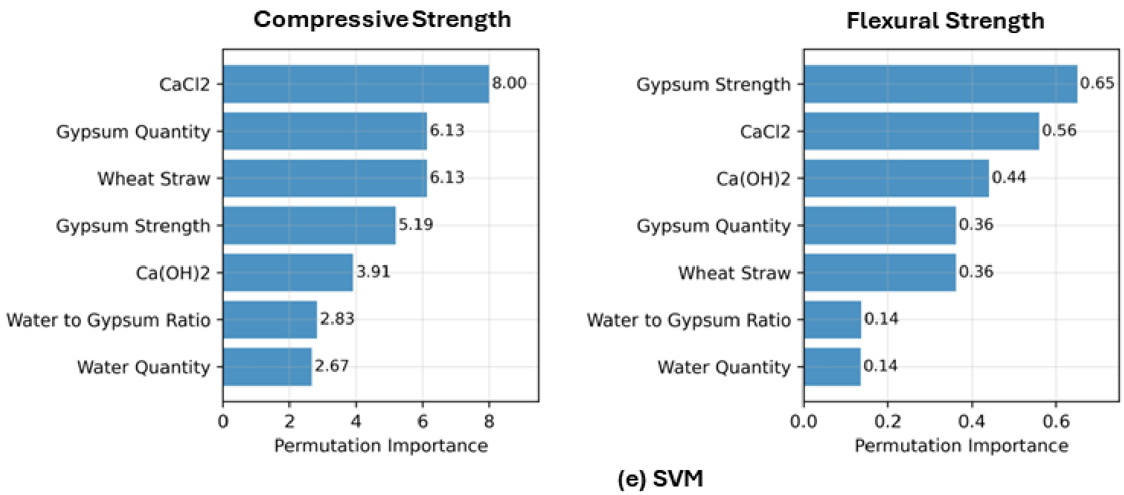


Figure 18. Comparison of feature importance of models.

For compressive strength, gypsum quantity ranked as the next most influential parameter in all models, indicating that binder content can affect load-bearing capacity, as concluded in [47-51]. In contrast, wheat straw and water-related parameters (water quantity and water-to-gypsum ratio) showed moderate importance, reflecting their possible indirect influence of fiber addition and porosity, respectively. Chemical additives such as $CaCl_2$ and $Ca(OH)_2$ generally exhibited lower importance values, suggesting a minor role in strength development under the studied dosage ranges, as shown in Figure 18.

For flexural strength, each model behaves differently from the others in its predictions. ANN model identifies gypsum quantity as a major contributor after wheat straw, as shown in Figure 18(a), indicating the possible effect of binding provided by gypsum quantity in enhancing flexural performance. Gypsum strength shows a moderate influence, whereas chemical additives and water-related parameters contribute marginally. For GPR model, $CaCl_2$ and $Ca(OH)_2$ content show relatively high importance after gypsum strength, compared to other models, as illustrated in Figure 18(b). This suggests that GPR possibly captures the influence of chemical activation on microstructural development. Additionally, wheat straw quantity, gypsum quantity, and water-to-gypsum ratio also show a relatively less contribution compared to other parameters. In RF model, the water-to-gypsum ratio ranks second after gypsum strength, as shown in Figure 18(c). This indicates that RF likely emphasizes mixture proportioning effects corresponding to water and gypsum, particularly the balance between workability, which directly affects flexural strength. Whereas the other parameters show the minor rule in contribution to predictions. XGBoost model ranks water-to-gypsum ratio as a major contributor after gypsum strength, as shown in Figure 18(d), suggesting a similar interpretation as RF model. The water and gypsum quantity shows a moderate contribution after the water-to-gypsum ratio. Whereas the other parameters show the minor rule in contribution to predictions. For SVM model, although the model shows relatively less accuracy compared to other models, feature importance analysis shows a meaningful interpretation of input parameters for predictions. $CaCl_2$, $Ca(OH)_2$, gypsum quantity, and wheat straw exhibit comparable secondary importance after gypsum strength, as presented in Figure 18(e), reflecting the possible influence of chemical activation and fiber reinforcement effects on microstructural development as discussed above.

Overall, despite differences in relative rankings, all models consistently confirm that gypsum strength is the governing factor. These findings provide possible physical interpretability and validate the reliability of ML models in capturing the underlying material behavior of gypsum composites. It is important to note that certain input parameters exhibit strong multicollinearity, particularly between water content and the water-to-gypsum ratio, as well as between water and gypsum quantities. While the employed ML models are capable of handling correlated inputs in terms of predictive accuracy, multicollinearity may influence feature importance attribution. In permutation-based analyses and SHAP-based interpretations (section 3.4.1 and section 3.4.2), predictive contributions may be redistributed among correlated variables, potentially affecting their relative rankings. Therefore, feature importance results should be interpreted with consideration of underlying correlations.

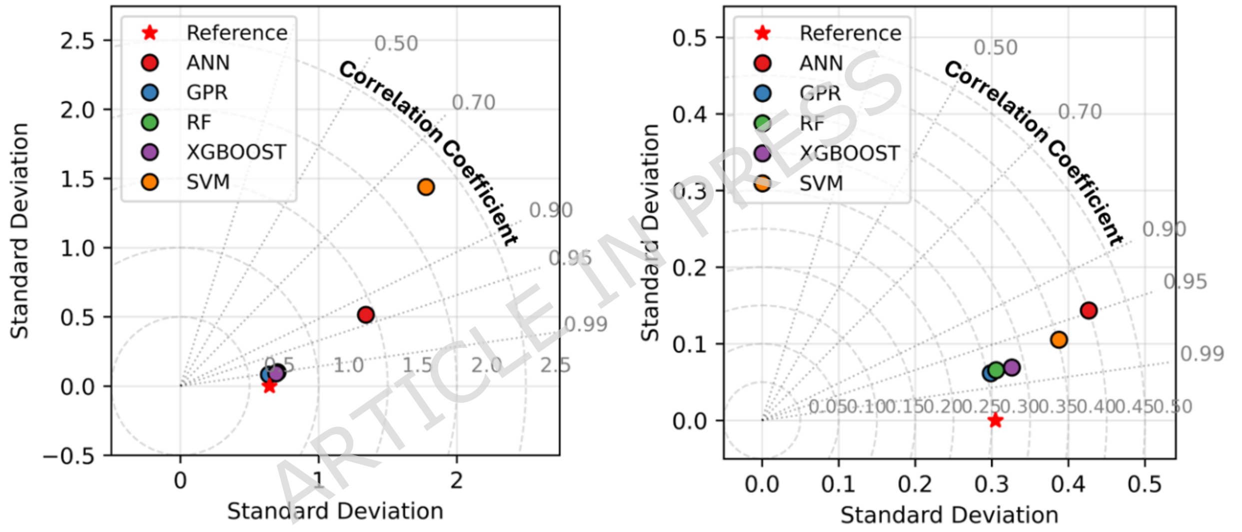
3.3. Interpretation of Taylor diagrams for model performance

Taylor diagrams illustrating the comparative performance of ANN, GPR, RF, XGBoost, and SVM models in predicting compressive strength and flexural strength are presented in Figure 19. The Taylor diagram provides a compact statistical summary by simultaneously displaying the correlation coefficient, standard deviation of RMSE difference between model predictions and experimental reference data. In these diagrams, the reference point represents experimental measurements, while model performance improves as points approach the reference with higher correlation and closer standard deviation to the reference point.

For compressive strength prediction, GPR and XGBoost models are located closest to the reference point, indicating high correlation coefficients (approaching 0.95-0.99) and standard deviations comparable to the experimental data, as shown in Figure 19(a). This confirms their superior ability to reproduce both the variability and trend of the compressive strength measurements. RF and ANN models exhibit slightly lower correlation and moderate deviations in standard deviation, suggesting reasonable but comparatively less accurate predictions. In contrast, SVM model lies farther from the reference, characterized by lower correlation and substantially higher standard deviation, reflecting larger prediction variability and reduced reliability for compressive strength estimation.

For flexural strength prediction, all models show improved clustering near the reference point, indicating generally higher predictive consistency compared to compressive strength. Among them, GPR model again demonstrates the closest agreement with the experimental data, exhibiting high correlation and minimal deviation in standard deviation, as illustrated in Figure 19(b). RF and XGBoost models also perform well, with strong correlations and slightly higher dispersion. ANN model shows moderate agreement, while SVM model displays relatively lower correlation and larger deviation, although its performance improves compared to the compressive strength case.

Overall, Taylor diagrams confirm that GPR provides a relatively balanced and accurate representation of both compressive and flexural strength, capturing not only the trend but also the statistical variability of experimental data. XGBoost and RF follow as possible alternatives, whereas ANN and particularly SVM exhibit comparatively lower predictive consistency. These observations are consistent with the quantitative performance metrics and further validate the performance of ML models for predicting the mechanical properties of gypsum composites.



(a) Compressive Strength **(b) Flexural Strength**

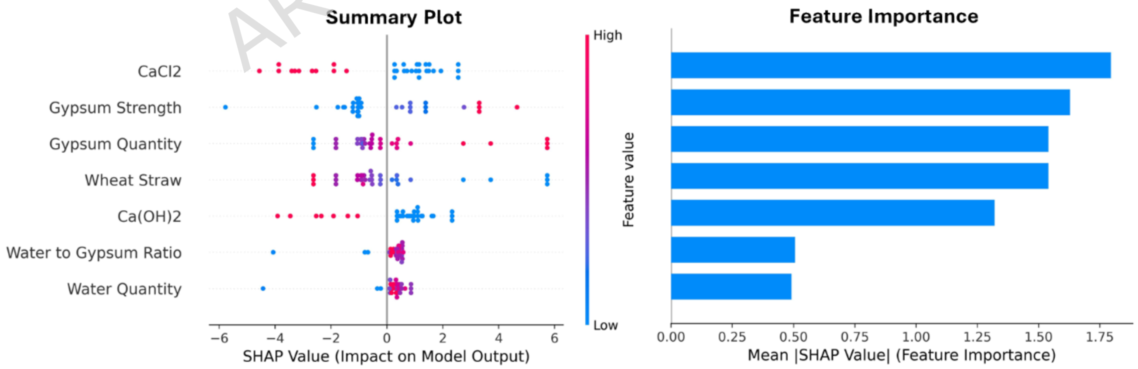
Figure 19. Taylor diagrams comparing the performance of models.

3.4. SHAP interpretability of GPR model predictions

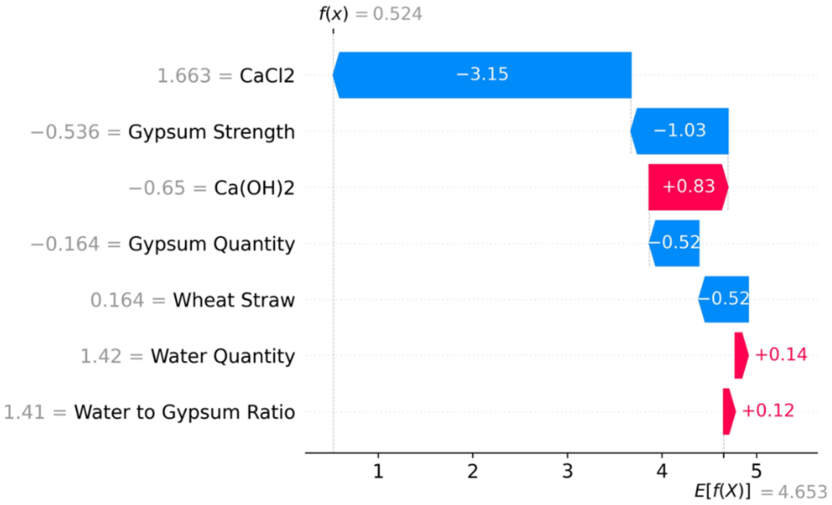
To further enhance the interpretability of GPR model as the best-performing approach, SHapley Additive exPlanations (SHAP) [72] were employed to quantify the contribution of individual input parameters to the prediction of compressive and flexural strength. SHAP analysis provides both global explanations, through summary plot and mean SHAP-based feature importance, and local explanations, through waterfall and force plots, by assigning each feature a contribution value. In SHAP summary plots, red-colored points represent higher feature values, while blue-colored points correspond to lower feature values. The horizontal position of each point indicates the magnitude and direction of its impact on the predicted response, where positive SHAP values increase the predicted output value and negative values decrease it. The feature importance plots present the mean absolute SHAP values, offering a global ranking of parameters according to their overall influence on model predictions across the dataset. In contrast, the waterfall and force plots provide instance-level explanations, illustrating how individual features cumulatively shift the prediction from the model's baseline value to the final output for a specific sample. Here, the baseline value $E[f(x)]$ represents the average model prediction over all training samples when there is no specific information about a sample, while the final output $f(x)$ corresponds to the predicted value after accounting for the combined contributions of all features. Together, these SHAP visualizations enable a comprehensive understanding of both the dominant predictors that govern the global behavior and the sample-specific predictions of the model, thereby strengthening the explainability of the GPR model.

3.4.1. SHAP values for compressive strength prediction

Figure 20 presents the SHAP summary plots, mean SHAP-based features importance, waterfall, and force plots for compressive strength. $CaCl_2$ exhibits the highest mean SHAP value, confirming its dominant negative influence on the global predicted response. Gypsum strength, gypsum quantity, and wheat straw content show substantial SHAP contributions, as shown in Figure 20(a), suggesting possible roles, such as binder strength and content and fiber addition, to affect the load-carrying capacity through reinforcement. $Ca(OH)_2$ exhibits a moderate negative effect with SHAP values, suggesting that higher amounts of these additives tend to reduce the predicted compressive strength. Water-related parameters show a minor positive SHAP influence on compressive strength within the investigated range.



(a) Summary Plot and Feature Importance



(b) Waterfall Plot

(c) Force Plot

Figure 20. SHAP analysis and interpretation for compressive strength of GPR prediction.

The local interpretability of the model for compressive strength prediction using SHAP waterfall and force plots in Figure 20(b) and Figure 20(c). The waterfall plot explains a single prediction by showing how each input parameter incrementally shifts the model output from the baseline value $E[f(x)] =$ 4.653 MPa to the final predicted value $f(x) =$ 0.524 MPa. Features shown in blue contribute negatively to the prediction, reducing compressive strength, whereas red bars indicate positive contributions. For this specific sample, $CaCl_2$ (−3.15 MPa), gypsum strength (−1.03 MPa), gypsum quantity (−0.52 MPa), and wheat straw (−0.52 MPa), with corresponding SHAP value in grey, exert a negative influence, indicating that higher values of these parameters reduce the predicted compressive strength for this sample. In contrast, $Ca(OH)_2$ (+0.83 MPa), water quantity (+0.14 MPa), and the water-to-gypsum ratio (+0.12 MPa), with corresponding SHAP value, provide positive contributions.

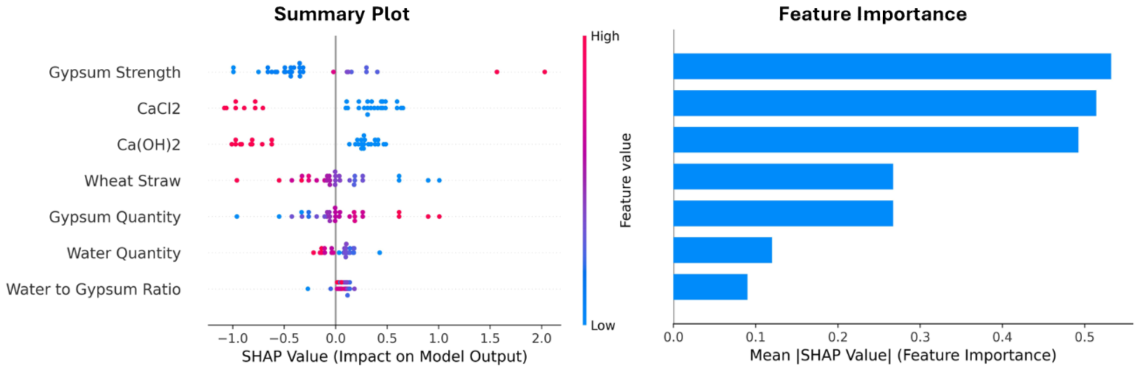
The corresponding force plot provides a complementary visualization by showing how features push the prediction away from the baseline value. Red arrows represent parameters that increase compressive strength, while blue arrows indicate parameters that decrease it. The length of each arrow reflects the magnitude of influence. Consistent with the waterfall plot, $CaCl_2$ and gypsum strength dominate the downward shift in the predicted value, while $Ca(OH)_2$ and water quantity contribute positively. Together, these plots demonstrate the capability of the GPR model to quantify both the direction (positive or negative) and magnitude of individual feature effects on compressive strength at the sample level.

3.4.2. SHAP values for flexural strength prediction

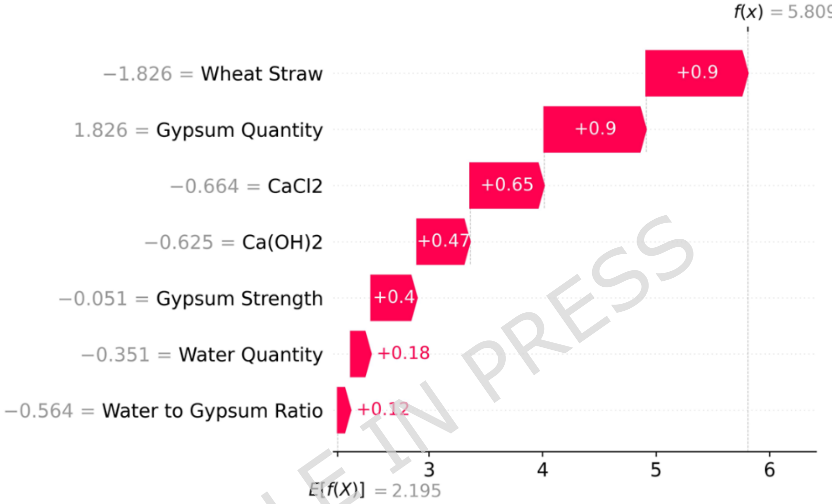
For flexural strength prediction, the global and local interpretability of the model, as shown in Figure 21, reveals a different contribution pattern. Although gypsum strength remains the most influential parameter on global prediction response, showing the higher SHAP values with positive influence, with a possible reason of directly affecting the mechanical properties of gypsum composites. $CaCl_2$, $Ca(OH)_2$, and wheat straw content demonstrate a negative influence with higher SHAP contributions, as shown in Figure 21(a), indicating that the increasing chemical additives and fiber addition can cause a decrease in the flexural strength globally in the studied ranges. The gypsum quantity and water-to-gypsum ratio exhibit a relatively higher SHAP value with a positive influence, indicating that, within the studied range, binder content and water availability can promote bonding, thereby enhancing flexural strength. The negative SHAP contribution of water quantity suggests that excessive quantity may introduce weak bonding, thereby reducing flexural strength.

The local interpretability for flexural strength prediction is illustrated in Figure 21(b) and Figure 21(c). For the selected sample, the prediction shifts from the baseline value $E[f(x)] = 2.195$ MPa to the final output $f(x) = 5.809$ MPa. The waterfall plot shows that wheat straw (+0.9 MPa) with the corresponding SHAP value contributes to the largest positive shift. Similarly, gypsum quantity (+0.9 MPa), $CaCl_2$ (+0.65 MPa), $Ca(OH)_2$ (+0.47 MPa), gypsum strength (+0.4 MPa), water quantity (+0.18 MPa), and wheat straw (+0.12 MPa) exert additional positive influences.

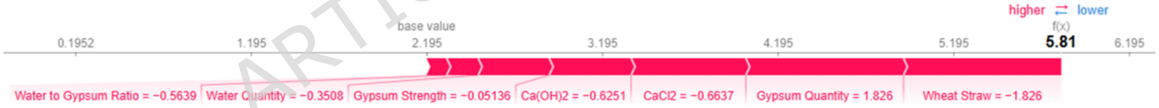
The force plot, as shown in Figure 21(c), complements the waterfall plot by visually summarizing these effects along a single axis, clearly highlighting how negatively contributing variables (blue) dominate the prediction shift, while positively contributing variables (red) provide partial recovery toward higher flexural strength as discussed in Section 3.4.1. SHAP explanations highlight the complex, nonlinear interactions captured by the GPR model and provide insight into the decision-making process underlying individual predictions, which cannot be captured by permutation importance alone.



(a) Summary Plot and Feature Importance



(b) Waterfall Plot



(c) Force Plot

Figure 21. SHAP analysis and interpretation for flexural strength of GPR prediction.

4. Conclusions

This study developed and evaluated a comprehensive ML framework for predicting the compressive and flexural strength of wheat straw reinforced gypsum composites using a collected dataset of 161 experimental mixtures. Five ML models, ANN, GPR, RF, XGBoost, and SVM, were systematically trained, validated, and compared under 10-fold cross-validation to ensure reliable and unbiased performance assessment.

- Based on quantitative evaluation metrics, GPR model consistently demonstrated the highest predictive accuracy for both compressive and flexural strength, achieving the highest R^2 values and lowest RMSE, MAE, and MSE. XGBoost ranked as the second-best model for compressive strength prediction, exhibiting performance close to GPR, while RF emerged as the second-best model for flexural strength estimation, whereas SVM exhibited comparatively lower predictive performance and higher variability. The performance comparison is further confirmed by the Taylor diagram analysis, where GPR exhibits the closest agreement with the reference point, characterized by a high correlation coefficient and minimal standard deviation error.

- Experimental versus predicted comparisons further highlighted the relatively strong agreement of GPR predictions with measured values, accompanied by predicted uncertainty. A key advantage of GPR is its inherent capability to quantify prediction uncertainty, which is particularly valuable for engineering decision-making and material design under uncertainty.
- Features importance analysis revealed that gypsum strength is the dominant parameter governing both compressive and flexural performance across most of models. Water-related parameters, wheat straw content, and gypsum quantity exhibited secondary but meaningful influence, while chemical additives showed variable contributions depending on the mechanical response.
- SHAP analysis provided valuable insights into both global and local prediction behavior of the best-performing GPR model. Globally, $CaCl_2$ was identified as the most influential parameter governing compressive and gypsum strength for flexural strength, while water-related parameters and chemical additives exhibited varying degrees of positive and negative influence depending on the response type. Locally, SHAP waterfall and force plots revealed how individual input parameters interact to shift predictions from the baseline to the final output, highlighting sample-specific mechanisms. SHAP provides deeper insight by linking data-driven predictions with physically meaningful material behavior.

Overall, this study demonstrates that ML, particularly GPR, can serve as a reliable, interpretable, and efficient tool for predicting the mechanical properties of sustainable gypsum composites. The proposed framework reduces experimental research efforts and enhances the understanding of material behavior. Accordingly, the findings of this study support the development of low-cost and sustainable gypsum-based construction products incorporating agricultural waste for practical use in both newly constructed and existing buildings. Although external validation using a completely independent dataset would further strengthen the generalization assessment, such data were not available in this study.

Future research should focus on expanding the experimental database to include additional parameters, such as water absorption and drying shrinkage, to further improve model generalization. Moreover, integrating advanced hybrid and physics-informed machine learning models could support the sustainable design of gypsum composites. Notably, the datasets were compiled from previously published experimental investigations, the differences in specimen preparation procedures, mixing protocol, and curing regimes may introduce inherent variability that could influence model predictions. Establishing a standardized and fully controlled experimental database is recommended for future investigations to further improve model accuracy and generalization. Finally, to support practical implementation, the trained models and prediction scripts have been made publicly accessible as mentioned in the availability of data and material section. The repository includes data preprocessing, trained model parameters, hyperparameter optimization and prediction scripts.

Acknowledgments

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Haseeb Ahmad: Formal analysis, Investigation, Software, Writing – original draft, Visualization. **Muhammad Fahad Ejaz:** Conceptualization, Data curation, Methodology, Formal analysis, Software, Writing – original draft, Writing – review and editing. **Muhammad Rizwan Riaz:** Conceptualization, Supervision, Visualization, Validation. **Shaban Shahzad:** Data curation, Investigation, Resources. **Sarah El Kadri:** Resources, Supervision, Validation. **Maria G. Kmeid:** Methodology, Project Administration.

Funding

This study did not receive any funding.

Availability of data and material

The dataset generated during and/or analyzed during the current study are available in this article. The trained models and prediction scripts have been made publicly accessible at repository (https://doi.org/10.5281/zenodo.18877579).

Conflict of interest

The authors declare no conflict of interests.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Appendix. Test database for gypsum wheat-straw reinforced composite specimens experiencing flexure and compression tests.

Table A1: Database of gypsum composites incorporating wheat straw.

Source	Sample ID	Gypsum Strength (MPa)	Gypsum (g)	Water (g)	Water-to-gypsum Ratio	Wheat Straw (g)	CaCl$_2$ (%)	Ca(OH)$_2$ (%)	Flexure Strength (MPa)	Compressive Strength (MPa)
[47]	1	6.1	1000	550	0.55	0	0	0	3.913	5.567
	2	6.1	1000	550	0.55	0	0	0	3.786	6.787
	3	6.1	1000	550	0.55	0	0	0	3.623	5.953
	4	6.1	975	536	0.55	25	0	0	2.738	3.164
	5	6.1	975	536	0.55	25	0	0	2.896	3.587
	6	6.1	975	536	0.55	25	0	0	2.404	3.719
	7	6.1	975	536	0.55	25	0	0	-	3.740
	8	6.1	950	523	0.55	50	0	0	2.867	3.397
	9	6.1	950	523	0.55	50	0	0	2.279	3.587
	10	6.1	950	523	0.55	50	0	0	2.301	3.090
	11	6.1	950	523	0.55	50	0	0	-	3.164
	12	6.1	925	509	0.55	75	0	0	2.839	3.397
	13	6.1	925	509	0.55	75	0	0	2.494	3.624
	14	6.1	925	509	0.55	75	0	0	2.344	3.317
	15	6.1	925	509	0.55	75	0	0	-	3.397
	16	6.1	900	720	0.8	100	0	0	0.940	1.373
	17	6.1	900	720	0.8	100	0	0	0.722	1.300
	18	6.1	900	720	0.8	100	0	0	0.777	1.300
	19	6.1	900	720	0.8	100	0	0	-	1.337
[48]	20	2	1000	750	0.75	0	0	0	3.23	11.23
	21	2	1000	750	0.75	0	0	0	3.87	10.10
	22	2	1000	750	0.75	0	0	0	3.19	8.36
	23	2	1000	750	0.75	0	0	0	3.91	7.23
	24	2	975	730	0.75	25	0	0	2.41	6.17
	25	2	975	730	0.75	25	0	0	2.61	5.91
	26	2	975	730	0.75	25	0	0	2.81	5.53
	27	2	950	760	0.8	50	0	0	1.99	3.29
	28	2	950	760	0.8	50	0	0	2.09	3.23
	29	2	950	760	0.8	50	0	0	2.19	2.75
	30	5	1000	750	0.75	0	0	0	4.94	23.57
	31	5	1000	750	0.75	0	0	0	4.93	20.72
	32	5	1000	750	0.75	0	0	0	6.55	21.47
	33	5	1000	750	0.75	0	0	0	6.54	24.32
	34	5	975	730	0.75	25	0	0	3.85	8.25
	35	5	975	730	0.75	25	0	0	4.45	10.25
	36	5	975	730	0.75	25	0	0	3.90	8.45
	37	5	975	730	0.75	25	0	0	4.40	10.05
	38	5	975	730	0.75	25	0	0	3.78	8.17
	39	5	975	730	0.75	25	0	0	4.52	10.33
	40	5	950	760	0.8	50	0	0	2.28	4.47
	41	5	950	760	0.8	50	0	0	2.72	5.07
	42	5	950	760	0.8	50	0	0	2.29	4.57

	43	5	950	760	0.8	50	0	0	2.71	4.97
	44	5	950	760	0.8	50	0	0	2.30	4.65
	45	5	950	760	0.8	50	0	0	2.70	4.89
	46	16	1000	750	0.75	0	0	0	7.11	41.62
	47	16	1000	750	0.75	0	0	0	7.19	39.72
	48	16	1000	750	0.75	0	0	0	7.57	38.82
	49	16	1000	750	0.75	0	0	0	7.65	39.74
	50	16	975	730	0.75	25	0	0	5.21	16.15
	51	16	975	730	0.75	25	0	0	5.31	14.93
	52	16	975	730	0.75	25	0	0	5.39	15.81
	53	16	975	730	0.75	25	0	0	5.63	14.05
	54	16	975	730	0.75	25	0	0	5.71	14.05
	55	16	975	730	0.75	25	0	0	5.81	14.97
	56	16	950	760	0.8	50	0	0	3.71	9.10
	57	16	950	760	0.8	50	0	0	3.73	7.84
	58	16	950	760	0.8	50	0	0	3.83	8.60
	59	16	950	760	0.8	50	0	0	4.03	7.10
	60	16	950	760	0.8	50	0	0	4.13	7.10
	61	16	950	760	0.8	50	0	0	4.15	8.26
[49]	62	10.24	1000	600	0.6	0	0	0	3.64	10.24
	63	10.24	975	600	0.6	25	0	0	2.46	5.693
	64	10.24	950	600	0.6	50	0	0	1.26	3.765
	65	10.24	925	600	0.6	75	0	0	1.26	3.811
	66	10.24	900	600	0.6	100	0	0	1.26	3.352
[50]	67	7.18	1000	600	0.6	0	0	0	-	7.18
	68	7.18	990	600	0.6	10	0	0	-	5
	69	7.18	980	600	0.6	20	0	0	-	3.18
	70	7.18	970	600	0.6	30	0	0	-	3.03
	71	7.18	960	600	0.6	40	0	0	-	2.89
[51]	72	3	980	784	0.8	20	5	0	1	1.4
	73	3	980	784	0.8	20	5	0	1.21	1.5
	74	3	980	784	0.8	20	5	0	1.26	1.5
	75	3	970	776	0.8	30	5	0	1.31	1.7
	76	3	970	776	0.8	30	5	0	1.25	1.8
	77	3	970	776	0.8	30	5	0	1.45	1.7
	78	3	960	768	0.8	40	5	0	1.38	1.6
	79	3	960	768	0.8	40	5	0	1.48	1.7
	80	3	960	768	0.8	40	5	0	1.6	1.7
	81	3	950	760	0.8	50	5	0	1.27	1.4
	82	3	950	760	0.8	50	5	0	1.39	1
	83	3	950	760	0.8	50	5	0	1.75	1.8
	84	3	940	752	0.8	60	5	0	1.65	1.9
	85	3	940	752	0.8	60	5	0	1.48	1.4
	86	3	940	752	0.8	60	5	0	1.99	1.8
	87	3	980	882	0.9	20	5	0	0.49	0.6
	88	3	980	882	0.9	20	5	0	0.78	0.8
	89	3	980	882	0.9	20	5	0	0.67	0.5
	90	3	970	873	0.9	30	5	0	0.54	0.5
	91	3	970	873	0.9	30	5	0	0.7	0.6
	92	3	970	873	0.9	30	5	0	0.77	0.6
	93	3	960	864	0.9	40	5	0	0.76	1
	94	3	960	864	0.9	40	5	0	1.05	0.8
	95	3	960	864	0.9	40	5	0	1.21	1
	96	3	950	855	0.9	50	5	0	0.92	1
	97	3	950	855	0.9	50	5	0	0.98	1.1

98	3	950	855	0.9	50	5	0	1.17	0.9
99	3	940	846	0.9	60	5	0	0.92	1
100	3	940	846	0.9	60	5	0	1.09	1
101	3	940	846	0.9	60	5	0	1.09	1
102	3	980	980	1.0	20	5	0	0.49	0.6
103	3	980	980	1.0	20	5	0	0.49	0.4
104	3	980	980	1.0	20	5	0	0.49	0.4
105	3	970	970	1.0	30	5	0	0.49	0.8
106	3	970	970	1.0	30	5	0	0.49	0.5
107	3	970	970	1.0	30	5	0	0.73	0.4
108	3	960	960	1.0	40	5	0	0.75	0.6
109	3	960	960	1.0	40	5	0	0.78	0.5
110	3	960	960	1.0	40	5	0	0.74	0.5
111	3	950	950	1.0	50	5	0	0.68	0.6
112	3	950	950	1.0	50	5	0	0.49	0.6
113	3	950	950	1.0	50	5	0	0.91	0.5
114	3	940	940	1.0	60	5	0	0.69	0.5
115	3	940	940	1.0	60	5	0	0.93	0.8
116	3	940	940	1.0	60	5	0	1.4	0.8
117	3	980	784	0.8	20	0	5	1.78	3.9
118	3	980	784	0.8	20	0	5	1.81	3.3
119	3	980	784	0.8	20	0	5	1.54	2.9
120	3	970	776	0.8	30	0	5	1.51	2.7
121	3	970	776	0.8	30	0	5	1.77	2.9
122	3	970	776	0.8	30	0	5	1.52	3
123	3	960	768	0.8	40	0	5	1.59	2.5
124	3	960	768	0.8	40	0	5	1.22	2.7
125	3	960	768	0.8	40	0	5	1.27	2.4
126	3	950	760	0.8	50	0	5	1.25	2.2
127	3	950	760	0.8	50	0	5	1.4	2.3
128	3	950	760	0.8	50	0	5	1.41	2
129	3	940	752	0.8	60	0	5	0.93	1.8
130	3	940	752	0.8	60	0	5	1.47	2
131	3	940	752	0.8	60	0	5	1.29	1.7
132	3	980	882	0.9	20	0	5	1.51	2.5
133	3	980	882	0.9	20	0	5	1.56	2.5
134	3	980	882	0.9	20	0	5	1.4	2.4
135	3	970	873	0.9	30	0	5	1.23	1.7
136	3	970	873	0.9	30	0	5	1.26	2.3
137	3	970	873	0.9	30	0	5	1.27	2
138	3	960	864	0.9	40	0	5	1.03	2.1
139	3	960	864	0.9	40	0	5	1.04	1.9
140	3	960	864	0.9	40	0	5	1.17	1.8
141	3	950	855	0.9	50	0	5	0.86	1.4
142	3	950	855	0.9	50	0	5	1.03	1.7
143	3	950	855	0.9	50	0	5	1.23	1.8
144	3	940	846	0.9	60	0	5	1.25	1.8
145	3	940	846	0.9	60	0	5	0.87	1.7
146	3	940	846	0.9	60	0	5	1.03	1.7
147	3	980	980	1.0	20	0	5	1.23	1.9
148	3	980	980	1.0	20	0	5	1.71	2.8
149	3	980	980	1.0	20	0	5	1.17	2.1
150	3	970	970	1.0	30	0	5	1.19	2
151	3	970	970	1.0	30	0	5	1.23	2.1
152	3	970	970	1.0	30	0	5	1.24	2

153	3	960	960	1.0	40	0	5	1.08	1.2
154	3	960	960	1.0	40	0	5	1.14	1.6
155	3	960	960	1.0	40	0	5	0.19	1.8
156	3	950	950	1.0	50	0	5	0.69	1.2
157	3	950	950	1.0	50	0	5	1.01	1.7
158	3	950	950	1.0	50	0	5	0.83	1.5
159	3	940	940	1.0	60	0	5	0.54	1.1
160	3	940	940	1.0	60	0	5	0.94	1.4
161	3	940	940	1.0	60	0	5	1.01	1.6

References

1. Cement and concrete around the world. https://gccassociation.org/concretefuture/cement-concrete-around-the-world/ (accessed March 03, 2026).
2. Sharma, R.D. and N. Singh. Evaluation of strength and absorption behaviour of iron slag and recycled aggregates concrete and its comparative environmental estimation by life cycle assessment. *Cleaner Materials*. **10**, 100210 (2023).
3. Ma, J., H. Yuan, J. Zhang, and P. Zhang. Enhancing concrete performance: A comprehensive review of hybrid fiber reinforced concrete. *Structures*. **64**, 106560 (2024).
4. Kaptan, K., S. Cunha, and J. Aguiar. A review: construction and demolition waste as a novel source for CO2 reduction in Portland cement production for concrete. *Sustainability*. **16**(2), 585 (2024).
5. Rajendran, N. et al. Economic and environmental impact analysis of cellulose nanofiber-reinforced concrete mixture production. *Resources, Conservation and Recycling*. **212**, 107917 (2025).
6. Zhaurova, M., R. Soukka, and M. Horttanainen. Multi-criteria evaluation of CO2 utilization options for cement plants using the example of Finland. *International Journal of Greenhouse Gas Control*. **112**, 103481 (2021).
7. Argalis, P.P., M. Sinka, and D. Bajare. Recycling of cement–wood board production waste into a low-strength cementitious binder. *Recycling*. **7**(5), 76 (2022).
8. Shah, S.A.R. et al. Optimization of fresh and mechanical properties of sustainable concrete composite containing ARGF and fly ash: An application of response surface methodology. *Construction and Building Materials*. **362**, 129722 (2023).
9. Benhelal, E., G. Zahedi, E. Shamsaei, and A. Bahadori. Global strategies and potentials to curb CO2 emissions in cement industry. *Journal of cleaner production*. **51**, 142-161 (2013).
10. Jia, B., L. Peng, and Y. Zhao. Recycling waste glass powder in lightweight aggregate concrete: Towards lightweight, sustainable and durable marine engineering structures. *Construction and Building Materials*. **472**, 140690 (2025).
11. Riaz, M.R., M.M. Azhar, M.F. Ejaz, R. Azam, and M.F. Javaid. Mechanical characterization and impact resistance of concrete

incorporating recycled thin-walled hollow tubular GFRP sections. *Innovative Infrastructure Solutions*. **10**(12), 585 (2025).

12. Murali, G., L.S. Wong, I. Abdulkadir, H.A. Algaifi, and M. Abdellatief. Sustainable transformation of waste phosphogypsum into geopolymer concrete: comprehensive review on strength, durability, and microstructural characteristics. *Journal of Building Engineering*. **111**, 113597 (2025).
13. Abdullah, M., A.U. Qazi, Q.S. Khan, S.M.S. Kazmi, and M.J. Munir. Experimental and analytical investigations on shear performance of Ambient-Cured reinforced geopolymer concrete beams. *Buildings*. **14**(1), 204 (2024).
14. Aziz, W. et al. Mechanical properties, drying shrinkage and structural performance of coconut shell lightweight concrete. *Structures*. **35**, 26-35 (2022).
15. Ejaz, M.F. et al. Coconut shell waste as an alternative lightweight aggregate in concrete-A review. *Advances in Materials Research*. **11**(4), 299-330 (2022).
16. Domanskaya, I., V. Petropavlovskaya, T. Novichenkova, K. Petropavlovskii, and R. Fediuk. Potential of Gypsum-Based Matrices for Sustainable Composite Materials: A Comprehensive Review: Domanskaya, Petropavlovskaya, Novichenkova, Petropavlovskii, and Fediuk. *JOM*. **77**(4), 2367-2382 (2025).
17. del Río-Merino, M. et al. A review of the research about gypsum mortars with waste aggregates. *Journal of Building Engineering*. **45**, 103338 (2022).
18. Sonnier, R. et al. Fire behaviour of hemp, clay and gypsum-based light biobased concretes and renders. *Construction and Building Materials*. **331**, 127230 (2022).
19. Khrissi, Y., A. Tilioua, N. Laaroussi, and A. Bybi. Recycling date palm waste in a gypsum-based composite: experimental study of thermal, acoustic, mechanical and hydric performance. *Innovative Infrastructure Solutions*. **10**(7), 312 (2025).
20. Ramzi, S. and H. Hajiloo. Evaluation of the Fall-Off of Gypsum Board in Lightweight and Mass Timber Constructions and Implications in Fire Resistance. *Fire Safety Journal*. 104455 (2025).
21. Zhang, L. et al. Valorization of industrial by-product gypsum for sustainable lightweight brick: Performance, microstructure and environmental safety assessments. *Sustainable Chemistry and Pharmacy*. **48**, 102179 (2025).
22. Zhang, F. et al. Innovative valorization of basic oxygen furnace slag in gypsum-based sustainable building blocks. *Journal of Building Engineering*. **105**, 112489 (2025).
23. Togun, H. et al. Innovative PCM-Gypsum plasterboard technologies for smart building envelopes: Recent advances and applications. *Sustainable Energy Technologies and Assessments*. **82**, 104470 (2025).

24. Kou, Y.-f., G.-l. Xu, Q.-q. Ren, J.-p. Hao, and L.-m. Tian. Experimental investigations into the cyclic performance of innovative original bamboo frame and gypsum-based mortar composite walls. *Structures*. **80**, 109890 (2025).
25. Javed, M.H., M.R. Riaz, R. Azam, M. Kashif, and A. Fatima. Physico-mechanical characterization of eco-friendly gypsum composites incorporating shredded surgical face masks. *Innovative Infrastructure Solutions*. **9**(6), 186 (2024).
26. Doleželová, M., J. Krejsová, L. Scheinherrová, M. Keppert, and A. Vimmrová. Investigation of environmentally friendly gypsum based composites with improved water resistance. *Journal of Cleaner Production*. **370**, 133278 (2022).
27. Sharifi Teshnizi, E., M. Mirzababaei, J. Karimiazar, R. Arjmandzadeh, and K. Mahmoudpardabad. Gypsum and rice husk ash for sustainable stabilization of forest road subgrade. *Quarterly Journal of Engineering Geology and Hydrogeology*. **57**(1), qjegh2023-008 (2024).
28. Songkhla, W.N., C. Chaikaew, B. Chatveera, and G. Sua-iam. Experimental investigation of recycled HDPE waste fiber-reinforced gypsum with emphasis on durability and mechanical performance. *Results in Engineering*. 105330 (2025).
29. Seid, A.M. et al. Recent Progress on the Physical, Thermal, and Mechanical Properties of Expanded Polystyrene Waste–Based Composites. *International Journal of Polymer Science*. **2025**(1), 9285040 (2025).
30. Beldjilali, S. et al. Waste Brick as Partial Replacement of Gypsum in Mortars: Mechanical Performance and Environmental Benefits for Sustainable Construction. *Sustainability*. **17**(16), 7452 (2025).
31. Rouf, M.A., M.R. Alam, S.A. Belal, Y. Ali, and M.Z. Rahman. Mechanical and thermal performances of banana fiber–reinforced gypsum composites. *International Journal of Polymer Science*. **2025**(1), 8120082 (2025).
32. Pedreño-Rojas, M.A., M.J. Morales-Conde, and I. Flores-Colen. Incorporation of wood waste as aggregate in the production of recycled gypsum products: mechanical and thermal analysis. *Journal of Building Engineering*. 114567 (2025).
33. Chen, C., L. Fang, S. Jiu, Y. Chen, and Y. Liu. Preparation of novel straw-based shape-stabilized phase change materials and their effect on the properties of gypsum-based composites. *Journal of Energy Storage*. **116**, 116087 (2025).
34. Li, S. and G. Chen. Agricultural waste-derived superabsorbent hydrogels: Preparation, performance, and socioeconomic impacts. *Journal of Cleaner Production*. **251**, 119669 (2020).
35. Raza, W. et al. A review on the deteriorating situation of smog and its preventive measures in Pakistan. *Journal of Cleaner Production*. **279**, 123676 (2021).
36. SAARC, *Energy Center Report. In Possible Uses of Crop Residue for Energy Generation Instead of Open Burning; SAARC Energy Centre*. 2019: Islamabad, Pakistan.

37. Habib, A., S. Nasim, and A. Shahab. Charting Pakistan’s air quality policy landscape. *International Growth Centre: London, UK*. (2021).
38. Saeed, M. et al. Agricultural waste biomass energy potential in Pakistan. in *Proceedings of the International Conference held in Shanghai, PR China*. 2015. Leeds.
39. Kazemi, F., A.Ö. Çiftçioğlu, T. Shafighfard, N. Asgarkhani, and R. Jankowski. RAGN-R: A multi-subject ensemble machine-learning method for estimating mechanical properties of advanced structural materials. *Computers & Structures*. **308**, 107657 (2025).
40. Wang, R. et al. Application and prospect of machine learning techniques in cost estimation of building projects. *Engineering, Construction and Architectural Management*. **32**(12), 8445-8471 (2025).
41. Yuan, Z., W. Zheng, and H. Qiao. Machine learning based optimization for mix design of manufactured sand concrete. *Construction and Building Materials*. **467**, 140256 (2025).
42. Shaaban, M., M. Amin, S. Selim, and I.M. Riad. Machine learning approaches for forecasting compressive strength of high-strength concrete. *Scientific Reports*. **15**(1), 25567 (2025).
43. Sinkhonde, D., T. Bezabih, D. Mirindi, D. Mashava, and F. Mirindi. Ensemble machine learning algorithms for efficient prediction of compressive strength of concrete containing tyre rubber and brick powder. *Cleaner Waste Systems*. **10**, 100236 (2025).
44. Gharagoz, M.M., M. Noureldin, and J. Kim. Explainable machine learning (XML) framework for seismic assessment of structures using Extreme Gradient Boosting (XGBoost). *Engineering Structures*. **327**, 119621 (2025).
45. Zhan, H., W. He, H. Xu, J. Yang, and C. Shi. Machine learning-enhanced prediction of traffic-induced ground vibrations with vehicle–road surface–soil coupling and experimental validation. *Soil Dynamics and Earthquake Engineering*. **202**, 109995 (2026).
46. Aziz, M.T. et al. State-of-the-art artificial intelligence techniques in structural engineering: A review of applications and prospects. *Results in Engineering*. **28**, 107882 (2025).
47. Ejaz, M.F. et al. Physico-mechanical characterization of gypsum-agricultural waste composites for developing eco-friendly false ceiling tiles. *Sustainability*. **14**(16), 9797 (2022).
48. Vavřínová, N., K. Stejskalová, J. Teslík, K. Kubenková, and J. Majer. Research of mechanical and thermal properties of composite material based on gypsum and straw. *Journal of Renewable Materials*. **10**(7), 1859 (2022).
49. Gomes, D.A.C. et al. Analysis of the influence of wheat residues on gypsum composites. *Innovative Infrastructure Solutions*. **8**(1), 31 (2023).
50. El hammouti, A. et al. Application analysis and environmental impact of straw reinforced gypsum plaster for improving the energy efficiency in buildings in the six climate zones of Morocco. *Journal of Building Engineering*. **74**, 106829 (2023).

51. Regulska, K. and A. Repelewicz. Properties of gypsum composites with straw fillers. in *IOP Conference Series: Materials Science and Engineering*. 2019. IOP Publishing.
52. Durgun, M.Y. Experimental research on gypsum-based mixtures containing recycled roofing tile powder at ambient and high temperatures. *Construction and Building Materials*. **285**, 122956 (2021).
53. Petrella, A. et al. Experimental investigation on environmentally sustainable cement composites based on wheat straw and perlite. *Materials*. **15**(2), 453 (2022).
54. Morales-Conde, M.J., C. Rodríguez-Liñán, and M.A. Pedreño-Rojas. Physical and mechanical properties of wood-gypsum composites from demolition material in rehabilitation works. *Construction and Building Materials*. **114**, 6-14 (2016).
55. Bishop, C.M., *Neural networks for pattern recognition*. Oxford university press.(1995)
56. Khan, A.Q., H.A. Awan, M. Rasul, Z.A. Siddiqi, and A. Pimanmas. Optimized artificial neural network model for accurate prediction of compressive strength of normal and high strength concrete. *Cleaner Materials*. **10**, 100211 (2023).
57. Morsy, A.M., A.E. Abd Elmoaty, and A.B. Harraz. Predicting mechanical properties of engineering cementitious composite reinforced with PVA using artificial neural network. *Case Studies in Construction Materials*. **16** (2022).
58. Congro, M. et al. Prediction of the residual flexural strength of fiber reinforced concrete using artificial neural networks. *Construction and Building Materials*. **303**, 124502 (2021).
59. Williams, C. and C. Rasmussen. Gaussian processes for regression. *Advances in neural information processing systems*. **8** (1995).
60. Haruna, S. et al. Gaussian process regression model for the prediction of the compressive strength of polyurethane-based polymer concrete for runway repair: A comparative approach. in *IOP Conference Series: Earth and Environmental Science*. 2022. IOP Publishing.
61. Al-Kahtani, M., H. Zhu, Y.E. Ibrahim, S. Haruna, and S. Al-Qahtani. Study on the mechanical properties of polyurethane-cement mortar containing nanosilica: RSM and machine learning approach. *Applied Sciences*. **13**(24), 13348 (2023).
62. Breiman, L. Random forests. *Machine learning*. **45**(1), 5-32 (2001).
63. Farooq, F. et al. A comparative study of random forest and genetic engineering programming for the prediction of compressive strength of high strength concrete (HSC). *Applied Sciences*. **10**(20), 7330 (2020).
64. Qing, S. and C. Li. Data-driven prediction on critical mechanical properties of engineered cementitious composites based on machine learning. *Scientific Reports*. **14**(1), 15322 (2024).
65. Chen, T. XGBoost: A Scalable Tree Boosting System. *Cornell University*. (2016).

66. Alshboul, O., G. Almasabha, A. Shehadeh, and K. Al-Shboul. A comparative study of LightGBM, XGBoost, and GEP models in shear strength management of SFRC-SBWS. in *Structures*. 2024. Elsevier.
67. Nguyen, H., T. Vu, T.P. Vo, and H.-T. Thai. Efficient machine learning models for prediction of concrete strengths. *Construction and Building Materials*. **266**, 120950 (2021).
68. Cortes, C. and V. Vapnik. Support-vector networks. *Machine learning*. **20**(3), 273-297 (1995).
69. Jueyendah, S., M. Lezgy-Nazargah, H. Eskandari-Naddaf, and S. Emamian. Predicting the mechanical properties of cement mortar using the support vector machine approach. *Construction and Building Materials*. **291**, 123396 (2021).
70. Cemiloglu, A. et al. Support vector machine (SVM) application for uniaxial compression strength (UCS) prediction: a case study for Maragheh limestone. *Applied Sciences*. **13**(4), 2217 (2023).
71. Ahmad, H. and Y. Matsumoto. An Investigation of the Effect of Age Degradation on Natural Frequency of Single Span Prestressed Concrete Bridge Under Environmental Variation Using Gaussian Process Regression. in *International Conference on Experimental Vibration Analysis for Civil Engineering Structures*. 2025. Springer.
72. Lundberg, S.M. and S.-I. Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*. **30** (2017).