Proceedings of the 8th International Conference on Civil Engineering for Sustainable Development (ICCESD 2026), 5~7 February 2026, KUET, Khulna, Bangladesh

COMPARATIVE ANALYSIS OF MACHINE LEARNING AND EMPIRICAL MODELS FOR PREDICTING CONCRETE COMPRESSIVE STRENGTH

Tuli Rani Das*[1] and Joy Prokash Das[2]

[1] *Graduate Student, Leading University, Bangladesh, e-mail: ce_2012060021@lus.ac.bd*
[2] *Graduate Student, Khulna University of Engineering & Technology, Bangladesh, e-mail: das1905078@stud.kuet.ac.bd*

***Corresponding Author**

ABSTRACT

The accurate forecast of concrete compressive strength is critical for assuring structural safety, durability, and quality control in civil engineering projects. Conventional empirical methods, such as those recommended by the American Concrete Institute (ACI), rely on linear correlations, which frequently fail to capture the complex, nonlinear interactions between the many constituents of concrete mixtures. This paper presents a machine learning (ML)-based prediction system that estimates compressive strength directly from mix design parameters, avoiding the need for time-consuming and destructive laboratory testing. Eight regression models are trained and analyzed using the publicly accessible UCI Concrete Compressive Strength dataset. These models include Linear Regression, Ridge, Lasso, Support Vector Regression (SVR), Decision Tree, Random Forest, XGBoost, and Artificial Neural Network (ANN). Model effectiveness is assessed by metrics like R^2, Mean Absolute Error (MAE), RMSE, and 10-fold cross-validation. The study found that XGBoost, Random Forest, and ANN outperform standard empirical approaches and other machine learning models, with R^2 scores exceeding 0.91. Furthermore, SHAP (SHapley Additive ExPlanations) values and sensitivity analysis are used to better understand the model predictions, with curing age, cement content, and water being the most important variables. This proposed data-driven strategy provides a speedy, accurate, and cost-effective alternative to traditional approaches, especially in places with limited testing facilities.

Keywords: *Concrete Compressive Strength, Machine Learning, Mix Design Parameters, Regression Models, SHAP Analysis*

1. INTRODUCTION

Concrete is a universal construction material because of its versatility, economy, and structural performance. Its compressive strength is the crucial mechanical characteristic of a structure for bearing loads safely without collapsing (Xue & Zhou, 2018). It represents the capacity of concrete to support load without failure and maintain long-term safety, reliability, and durability during its service life. Compressive Strength is usually measured by standard laboratory tests in which cylindrical or cuboid specimens are formed, cured, and compressed with a hydraulic testing machine, usually after 28 days of curing. The 28-day testing is a common yardstick for checking any definition requirement of the concrete (Alsanusi & Bentaher, 2015). But the price one pays for predicting wrongly can be severe: Underprediction means delays and cost blowout; overprediction may cause fatal structural failure (Prayogo, 2018).

The practical testing of compressive strength is done experimentally but is subject to certain practical restrictions. These include being time, labour, and conductive intensive, destructive test methods and laboratory facilities, tight curing conditions required, and a skilled operator to prepare the testing set-up (Haryanto et al., 2023). As those tests take 7-28 days for a definite outcome (Alone et al., 2020), this delays construction on-site decision as to when to strike the formwork or apply load. Another drawback of the multiple specimen requirement is further waste of material and a higher cost for the project. The quality control of the prepared samples, as well as curing and testing conditions, are additional sources of variation, introducing a further source of measurement uncertainty (Pujitha et al., 2017). Lack of conventional techniques to meet these requirements has therefore precluded real-time on-site Quality control on large or fast-track projects. The demand is therefore increasing for a quick, low-cost cost and non-destructive predictive method which will enable accurate real-time determination of strength in concrete, leading to smarter (and more sustainable) construction.

Recent advances in ML have offered powerful solutions as an alternative to the conventional compressive strength test and enabled rapid, non-destructive data-driven prediction of CSP based on mix design criteria, including cement content, water dosage, aggregate size distribution (graded or ungraded), admixture type, and curing age. Complex nonlinear relationships of concrete composition data that could not be modelled by conventional regression analysis can be learnt by ML models. Techniques like Random Forest, ANN, SVMs, and gradient boosting e. g., XGBoost and LightGBM, have been reported to achieve high prediction accuracy with efficiency (Song et al., 2021). These methods save time, money, and material for testing, and improve quality control and real-time decision-making during construction. Indeed, unnecessary trials can be minimized and the possibility of an optimised mix design supported when ML-based prediction is integrated in construction practice to line with smarter and more sustainable ways (Kashifi et al., 2024).

Throughout 2020-2025, many research efforts have been done and demonstrated high accuracy of ML for predicting the Compressive Strength of concrete. Zhang et al. (2024) obtained a strong $R^2 = 0.91$ with the Random Forest model, which lends support to the power of ensemble-based prediction methods. Liu et al. (2023) obtained $R^2 = 0.94$ through an exercise of ANN, confirming the forecasting precision of deep learning in nonlinear relation force. Singh and Kumar (2024) utilized XGBoost alongside the Harris Hawks Optimization algorithm, obtaining excellent $R^2 = 0.96$ with compaction error of RMSE. Liu (2022) reported that for the prediction of HPC compressive strength with the XGBoost algorithm $R^2 = 0.9993$. For the prediction of compressive strength of eco-bricks, SHAP and machine learning models like ANN and Random Forest were employed, with the former obtaining R^2 greater than 0.99 while also identifying critical predictors such as age and fly ash (Chandra et al., 2025). Kumar et al. (2025) use Bi-LSTM for ultra-high-performance concrete, reaching an R^2 of 0.9464 and showing the possibility to optimize material selection and lower the experimental effort. Taken together, these results confirm that advanced ML methodologies can reliably and rapidly forecast 24-h concrete strength, providing an effective alternative to conventional testing for Quality assurance and smart construction applications.

Although considerable advances have been achieved, there are still some research issues to be addressed in the ML prediction of concrete compressive strength. The majority of current studies are based on a small or homogeneous database, which tends to reduce the generalisation capacity of models for different concrete mixtures, climatic conditions and curing systems. There is also insufficient consistent cross comparison, including multiple ML algorithms and limited external validation on independent datasets-which is important when real-world robustness needs to be gauged. Furthermore, limited open-access datasets and reproducible frameworks for benchmarking and collaborative development are lacking. And last, many models are still blackbox systems that focus on performance rather than explainability, which creates obstacles to their use in safety-critical civil engineering applications where transparency, explainability, and trust are also crucial.

As such, the current research aims at developing and benchmarking machine learning algorithms on concrete compressive strength with mix composition information. To achieve this goal, we systematically compare eight algorithms-Linear Regression, Ridge, Lasso, SVR, Decision Tree, Random Forest, XGBoost, and ANN-based on standardized evaluation criteria. In addition, this work studies feature importance and model interpretability with respect to important parameters using SHAP analysis. The novelty of this work lies in formulating a robust, interpretable ML framework with high prediction accuracy for the data-driven strength.

2. METHODOLOGY

The approach used in this study consists of five major steps: data acquisition, preprocessing, development of predictive models, performance evaluation, and interpretation of results. The overall workflow is schematically represented in Figure 1.

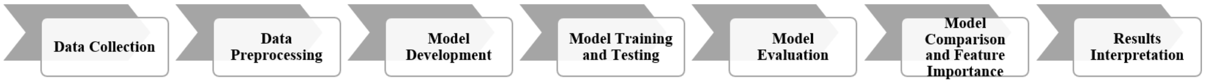

Figure 1: Workflow of the study

2.1 Data Collection

The Concrete Compressive Strength Dataset was utilized for the present study, which is publicly accessible from the UCI Machine Learning Repository and contains 1030 observations. There were 8 independent variables for concrete mix components in the dataset: Cement, Blast Furnace Slag, Fly Ash, Water, Superplasticizer, Coarse Aggregate, Fine Aggregate, and Age; and one dependent variable, which is concrete compressive strength. Since the dataset is well-established and widely referenced, no laboratory testing or field sampling was required, making it suitable for machine learning-based regression studies in civil engineering.

2.2 Data Preprocessing

The data set was cleaned to mitigate model-performance degradation and maintain data consistency. First, the dataset was explored to look for missing and inconsistent values, but these were not found. And then feature scaling is conducted using the StandardScaler method to make the models input parameters be unified. The data was split into training and testing (80:20) so that the models can be evaluated impartially.

All preprocessing was performed in Python 3.10 on Google Colab; libraries used included Pandas, NumPy, and Scikit-learn.

2.3 Model Development

Eight restricted-sense learning models were employed to predict the concrete compressive strength as follows: Linear Regression, Ridge Regression, Lasso Regression, Support Vector Regression, Decision Tree Regressor (DTR), Random Forest Regressor (RFR), Extreme Gradient Boosting (XGBoost), and Artificial Neural Network (ANN). All models were coded in Python and trained from the same pre-processed data to make the comparison of performance fair.

The plain linear regression served as our baseline model, and we also used Ridge and Lasso Regressions by imposing the L2, or the L1 regularization to enhance generalization. The SVR learned the nonlinear relationships by means of kernel functions. Decision Tree was a basic non-linear model, and Random Forest and XGBoost used the technology of ensemble learning and gradient boosting to improve prediction results. The ANN employed an eight-neuron input layer, a hidden layer with ReLU activation function, and an output neuron representing concrete strength.

2.4 Model Training and Evaluation

The hyperparameters of the models were empirically tuned to find the best performance. The training set was used for training the models, whereas the unseen test set (20%) ensured unbiased predictions. Hyperparameters are set to standard values and further adjusted after iterative testing. We measured the performance after training, and all implementations were performed in Python with standard machine learning packages.

2.5 Model Comparison and Validation

All models' results were compared to select the most effective approach of prediction. The predictive performance of each model was assessed using three standard statistical indicators:

RMSE - serves to measure the magnitude of prediction error.

Mean Absolute Error (MAE) - the average absolute difference between the predictions and observations.

Coefficient of Determination (R^2) - indicates the proportion of variance in compressive strength explained by the model.

Performance comparison also revealed the ability of this model to represent complex nonlinear relationships between blend components. These criteria are chosen due to their widespread use in regression problems and the fact that they can measure error and explanation capacity. In addition, the importance of features was analysed to obtain the key parameters affecting compressive strength prediction.

2.6 Feature Importance & Comparison

We ran all experiments at Google Colab open-source projects using Python 3.10. The main libraries employed in this paper are Scikit-learn, Pandas, NumPy, Matplotlib, XGBoost, and TensorFlow/Keras. This computational configuration allowed for rapid training and evaluation of the interactions.

Feature importance analysis was conducted using the models that are best suited to determine the most significant input variables on compressive strength. Then, a performance comparison of models based on evaluation metrics was carried out to identify which algorithm would perform best in terms of its predictive accuracy.

3. RESULTS AND DISCUSSION

This section focuses on presenting the prediction achievements of the eight machine-learning methods and the ACI empirical equation in predicting concrete compressive strength.

3.1 Descriptive Statistics of the Dataset

A preliminary description of the input parameters and statistical analysis was carried out to determine their distribution and variability. Table 1 shows the minimum, maximum, mean, and standard deviation of some key attributes in the dataset. The large variability in cement and water contents indicates that mix designs are very different across the test records; curing age is highly scattered and similarly presents significant contributions to model learning. The compressive strength region will include both low and high-strength concretes, which will allow the development of strong predictive models. Several SCM combinations (fly ash with slag) were also employed to increase the variety of data and to better learn nonlinearly related models using ML.

Table 1: Descriptive statistics of the dataset

Feature	Unit	Min	Max	Mean	Std
Cement	kg/m^3	102	540	281.17	104.51
Slag	kg/m^3	0	359.40	73.90	86.28
Water	kg/m^3	121.75	247	181.57	21.36
Fly Ash	kg/m^3	0	200.10	54.19	63.99
Fine Aggregate	kg/m^3	594	992.60	773.58	80.18
Coarse Aggregate	kg/m^3	801	1145	972.92	77.75
Superplasticizer	kg/m^3	0	32.20	6.20	5.97
Curing Age	days	1	365	45.66	63.17
Compressive Strength	MPa	2.33	82.60	35.82	16.71

3.2 Model Performance Comparison

Model Performance comparisons are shown in Table 2. The XGBoost model achieved the best performance in accuracy ($R^2 = 0.9059$), followed by Random Forest ($R^2 = 0.8819$), ANN ($R^2 = 0.8479$). While the performances of Linear Regression, Ridge, Lasso, and SVR were lower moderate level, and ACI formulas provided negative values (that is, no generalization to our dataset).

Table 2: Model performance comparison

Model	R^2	MAE	RMSE
XGBoost	0.9059	3.2247	4.924
Random Forest	0.8819	3.7581	5.5167
ANN	0.8479	4.771	6.2598
SVR	0.6547	7.5154	9.4328
Decision Tree	0.6439	7.3822	9.5792
Ridge	0.6276	7.7518	9.7964
Lasso	0.6275	7.7522	9.7967
Linear Regression	0.6275	7.7454	9.7967
ACI Empirical	-4.1029	32.8923	36.2619

These evidences demonstrate that the nonlinear ensemble models are superior to both linear forms and typical empirical formulas, which in turn verifies that there exist nonlinear interactions among the individual cementitious materials, aggregates, water, and age.

3.3 R² Score Comparison

The highest performance was achieved with XGBoost (accuracy: 0.91), followed by Random Forest and ANN performances, both with accuracies of 0.88 and 0.85, respectively, suggesting high RSA in mix satisfaction prediction (Figure 2). Finally, SVR and Decision Tree (DT) were moderately acceptable with an accuracy of 0.65; yet Linear Regression (LR), Ridge, and Lasso had the weakest linkages with an accuracy of 0.63.

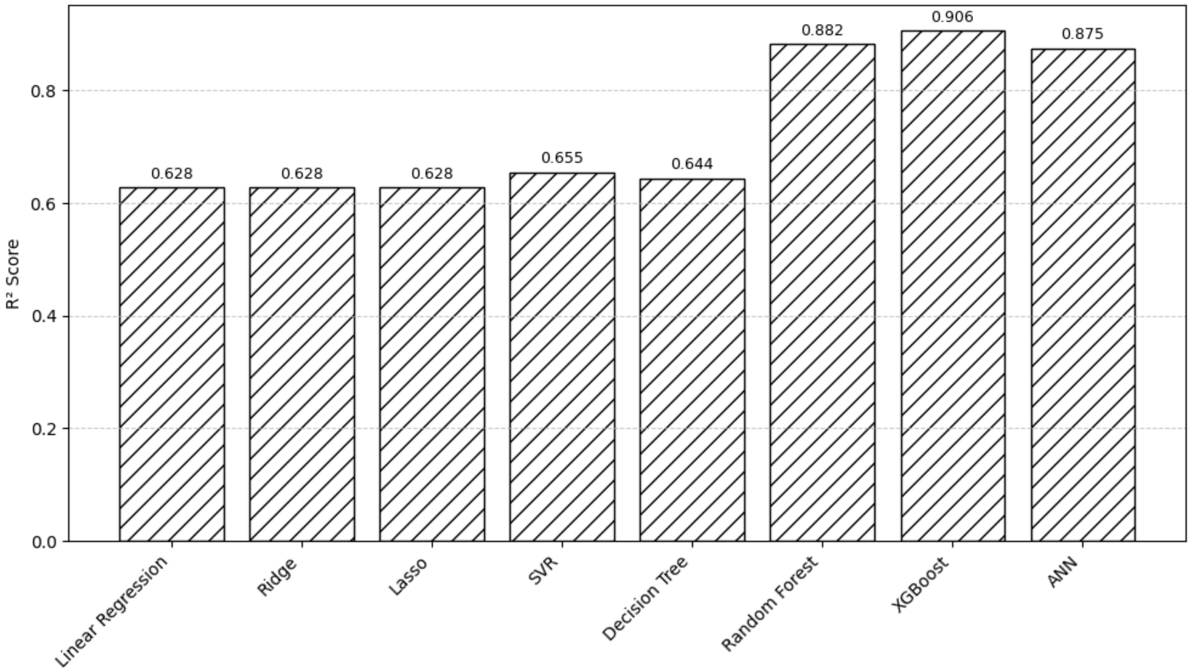

Figure 2: R² comparison of all predictive models

This section details the comparison of all the proposed machine learning models based on 10-fold cross-validation. Regarding the performance validation, R² is adopted to measure how well created models obtain an appropriate estimation of concrete compressive strength in this work. The mean R² values with their standard deviations are presented in Figure 3.

XGBoost was shown to be the best predictor of performance when considering model R² (mean = 0.91±0.02), followed by Random Forest with a mean R² on test sets of 0.88±0.03 [3, 5]. The ensemble methods (XGBoost, Random Forest) are clearly better in comparison to the single causal estimators (Decision Tree, SVR) and linear models (Ridge, Linear Regression, Lasso). These results verify that ensemble learning works well to retain the nonlinear relationships of input mix parameters and compressive strength.

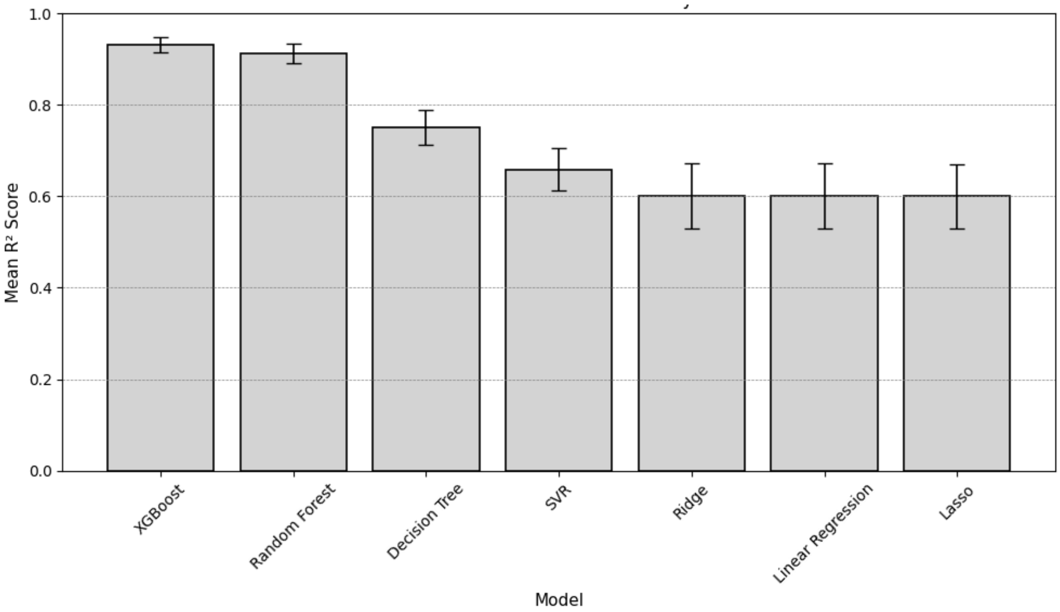

Figure 3: 10-fold cross-validation: mean R^2 by model

3.4 Prediction Error Histogram

Histograms for the prediction errors (y_{pred} - y_{true}) obtained by all models are shown in Figure 4, where the horizontal axis denotes prediction error in MPa.

Both XGBoost and Random Forest exhibit the narrowest and the most centered error distributions, which means predictions tend to be accurate and unbiased. The ANN is moderately good with a slightly wider spread. Linear models, SVR, and Decision Tree have wider error distributions, suggesting a high degree of variance. In contrast, deviations of the ACI empirical model demonstrate a pronounced negative bias, with errors predominantly shifted toward negative values.

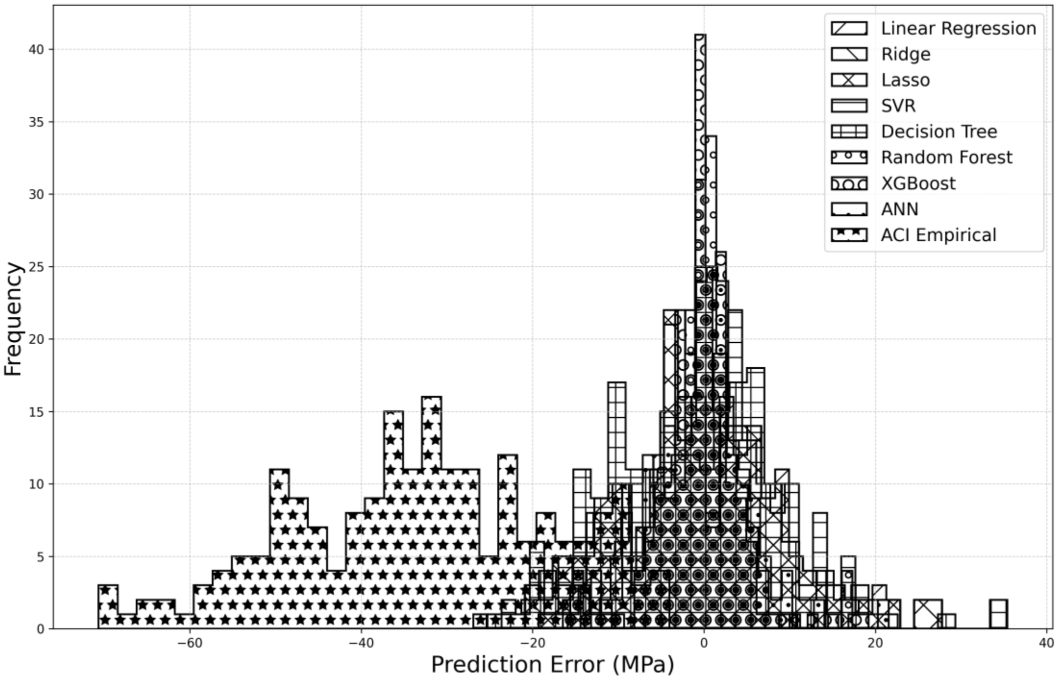

Figure 4: Histogram for prediction errors (y_{pred} - y_{true}) in MPa for all models

3.5 Actual vs Predicted - XGBoost

The scatter plot of predicted vs. actual compressive strength is shown in Figure 5 for the XGBoost model, where the points appear to be tightly packed around the 45° reference line, indicating an excellent accuracy with low prediction error overall strength ranges.

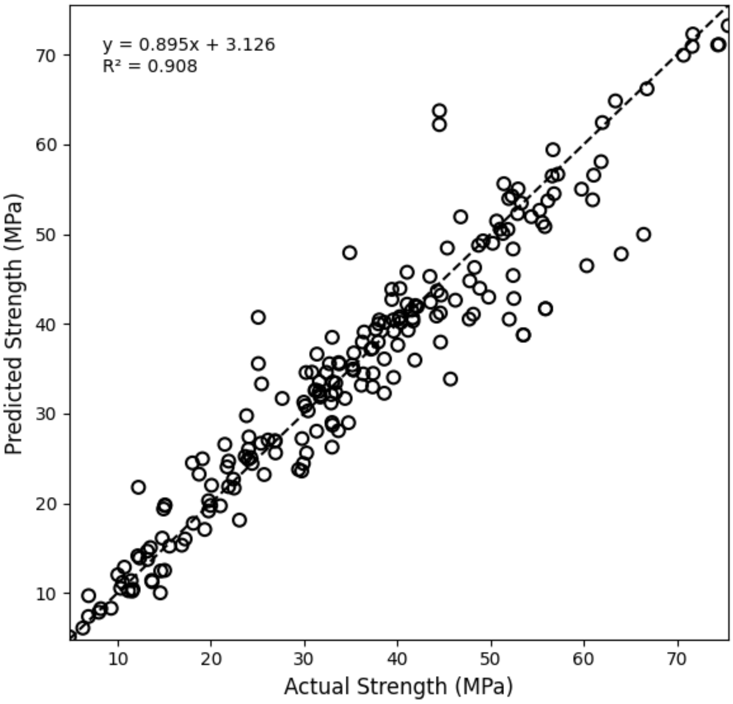

Figure 5: Actual vs Predicted strength (XGBoost)

At the higher values of strength, there is little scatter, which may be attributed to fewer samples at high strength and more variability. Based on these indicators, XGBoost has a good ability to describe nonlinear relationships between the mix design values and compressive strength.

3.6 Feature Importance (SHAP Analysis)

SHAP values were obtained to measure the effect of each input factor on concrete compressive strength prediction. The most useful information provided is that the age of testing contributes the highest to the variability in strength, followed by cement content, water, slag, and superplasticiser dose. At the same time, aggregate and fly ash are less significant. This indicates that the model primarily captures strength evolution governed by binder hydration mechanisms rather than aggregate-controlled effects.

Figure 6 indicates the ranking of feature importance by SHAP analysis, and age, as well as binder-related components, are ranked as significant predictors.

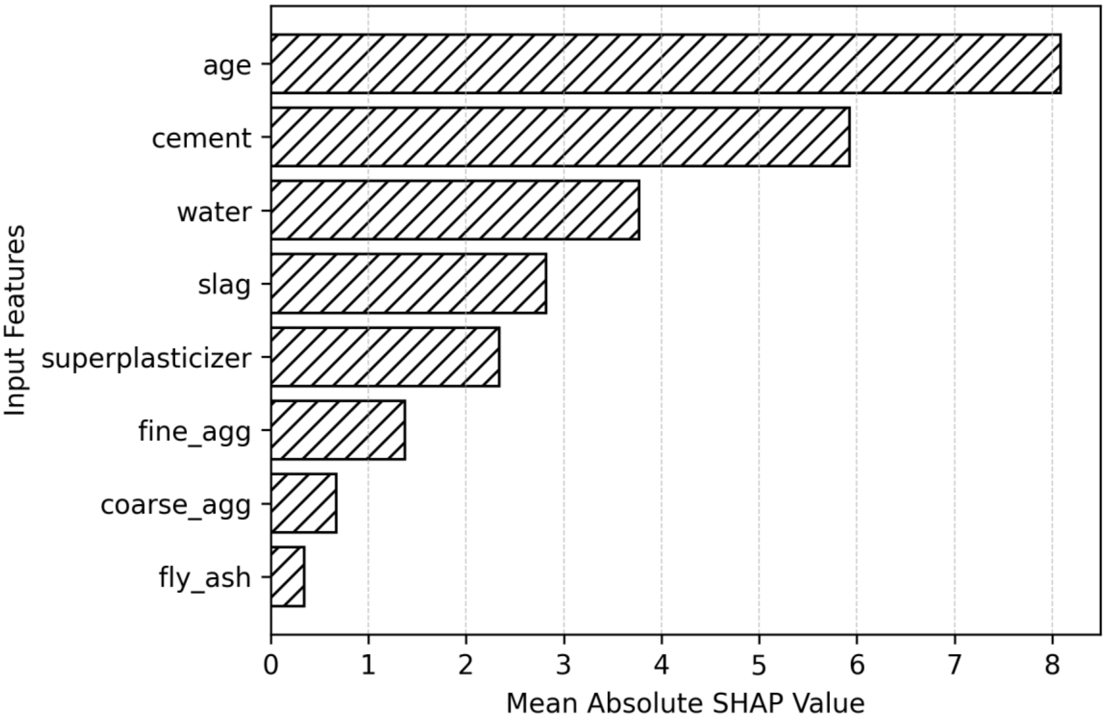

Figure 6: Mean SHAP importance of input features

3.7 Sensitivity Analysis

A sensitivity analysis was performed to investigate how the prediction of compressive strengths changes with each of such input parameters. Within the investigated range, it can be seen from Figure 7 that with the aging and increase of cement content, strength increases primarily due to improved hydration progress and paste bonding. However, excessive cement can actually lead to higher porosity and tiny cracks, which can hinder any additional gains in strength. It has been observed that when water content increases, strength decreases because of a higher water/cement ratio. The incremental addition of slag, fly ash, and superplasticizer at a low concentration level improves the long-term strength (≥28 days) performance of mixtures, but it is likely that an excessive replacement with SCM reduces their fresh state behavior. In this study, the influence of fine and coarse aggregates is less significant on strength development, which means they have a secondary effect on strength generation relative to binder-related ingredients.

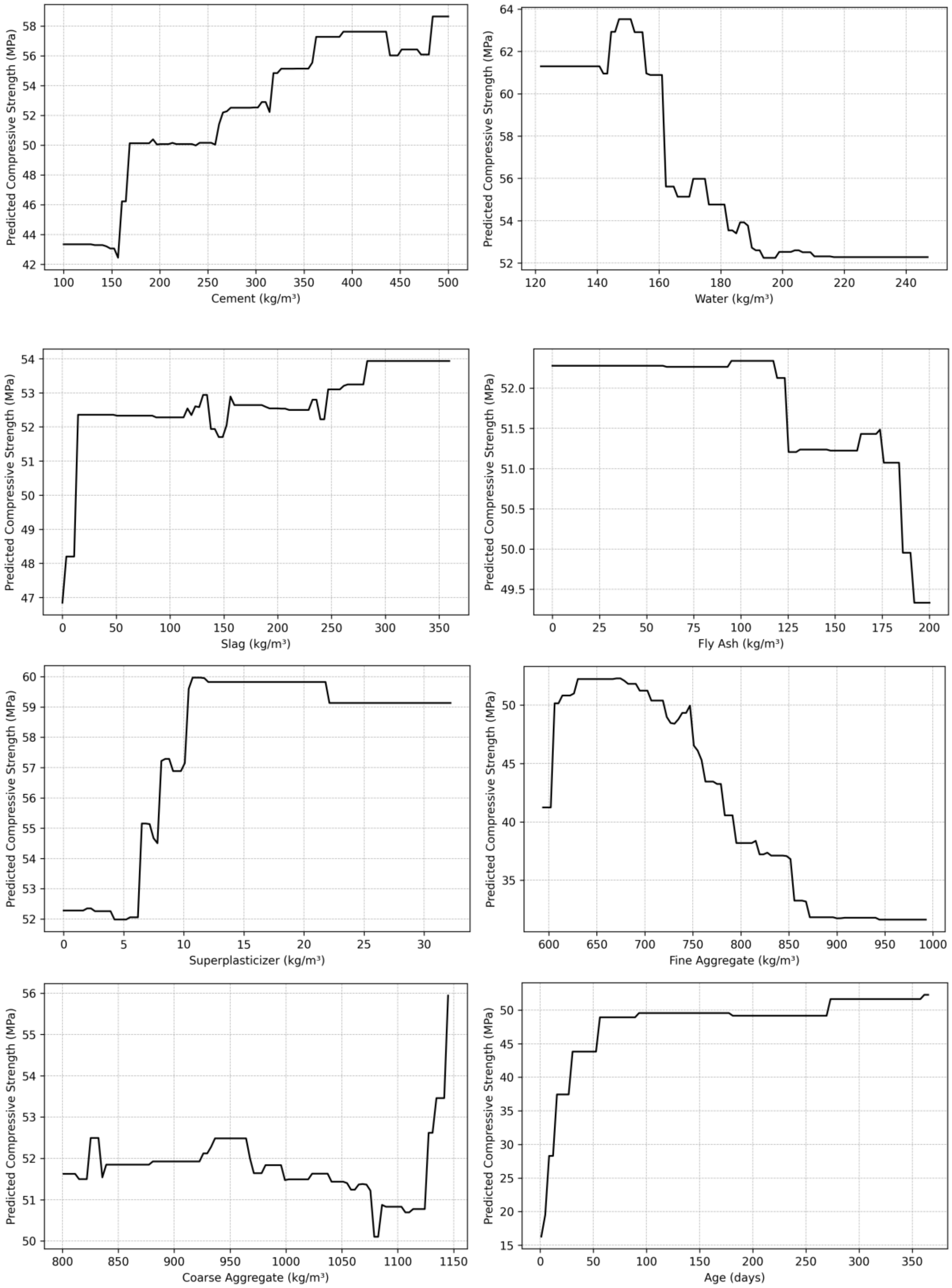

Figure 7: Sensitivity analysis of all input features

3.8 Overall Discussion

The comparative study indicates that the most accurate model to predict the compressive strength of concrete is XGBoost, having the highest R² and lower prediction error. Random Forest and ANN also predict well the nonlinear nature of materials. On the other hand, linear models are not competitive due to their incapacity to account for the presence of interactions, whilst the ACI empirical formula also can't be generalized across a broad spectrum of mix compositions.

Results obtained from SHAP and sensitivity analysis prove the importance of age and binder content as the two most influential factors, based on well-established physical reasons, which will improve the reliability and interpretability of machine learning results. On a practical level, such models can simplify the design of mixes, reduce trial-and-error methodology, and assist with decision making when limited laboratory testing facilities are available. Nonetheless, the study has some constraints as it was based on available systems and not all environmental parameters were taken into account, while a better performance of the ANN through further calibration is possible.

4. CONCLUSIONS

In this study, the performance of eight algorithms from machine learning (ML) that could predict the Compressive Strength of Concrete was investigated using the UCI dataset. In these models, XGBoost yielded the best predictions ($R^2 = 0.9059$), followed by Random Forest and ANN, essentially demonstrating the power of ensembled models in modeling non-linear relationships between mix parameters. All of the linear and empirical methods, as well as the ACI equation, failed to make accurate predictions, indicating that these models have low generality. Feature-importance and SHAP studies revealed age and the cement content as the most influential factors, both of which are in accordance with known material behavior. In general, the results demonstrate that high-level ensemble learning approaches are a reliable data-driven substitute for classical hand-calculation-based ones to accurately estimate strength in an efficient and interpretable manner, which is indispensable for smart mix design and QC practice as well as sustainable development of concrete.

DECLARATION OF USE OF AI

AI tools (such as ChatGPT and Quilbot) were utilized exclusively for language correction and clarification. All ideas, conclusions, and analyses are solely the authors' own, and any AI-assisted text has been examined by the authors.

REFERENCES

Alone, D., Meshram, G., & Ekare, P. (2020). An Experimental Investigation Of Compressive Strength On Concrete By Using Waste Cardboard - IRE Journals. 3(9), 54–58. https://www.irejournals.com/formatedpaper/1701992.pdf

Alsanusi, S., & Bentaher, L. (2015). Prediction of Compressive Strength of Concrete from Early Age Test Result Using Design of Experiments (RSM). Journal of Civil and Environmental Engineering, 9(12), 1567–1571. https://publications.waset.org/10003114/pdf

Chandra, S. S., Kumar, R., Arjunasamy, A., Galagali, S., Tantri, A., & Naganna, S. R. (2025). Predicting the compressive strength of polymer-infused bricks: A machine learning approach with SHAP interpretability. Scientific Reports, 15(1), 8090.

Haryanto, V. M., Cahayanti, S. R., Suroso, A., & Amin, M. (2023). Economic Analysis and Compressive Strength of Concrete, Cost and Utilization Time of Carbon-Sulfur and Dirty Sulfur Solid Waste. Syntax Idea, 5(6), 722-738.

Kashifi, M. T., Salami, B. A., Rahman, S. M., & Alimi, W. (2024). Using explainable machine learning to predict compressive strength of blended concrete: A data-driven metaheuristic approach. Asian Journal of Civil Engineering, 25(1), 219-236.

Kumar, R., Kumar, D. R., Wipulanusat, W., Thongchom, C., Samui, P., & Rai, B. (2025). Estimation of the compressive strength of ultrahigh performance concrete using machine learning models. Intelligent Systems with Applications, 25, 200471.

Liu, X., Qin, J., Zhao, K., Featherston, C. A., Kennedy, D., Jing, Y., & Yang, G. (2023). Design optimization of laminated composite structures using artificial neural network and genetic algorithm. Composite Structures, 305, 116500.

Liu, Y.-N. (2022). High-Performance Concrete Strength Prediction Based on Machine Learning. Computational Intelligence and Neuroscience, 2022, 1–7. https://doi.org/10.1155/2022/5802217

Prayogo, D. (2018). Metaheuristic-Based Machine Learning System for Prediction of Compressive Strength based on Concrete Mixture Properties and Early-Age Strength Test Results. Civil Engineering Dimension, 20(1), 21–29. https://doi.org/10.9744/CED.20.1.21-29

Pujitha, B., Swathi, N., & Saheb, Sk. J. (2017). Experimental Investigation of Compressive Strength Properties of Eco-Friendly Concrete. International Journal of Advance Research, Ideas and Innovations in Technology, 3(6), 733–743. https://www.ijariit.com/manuscripts/v3i6/V3I6-1357.pdf

Singh, R., & Kumar, V. (2024). Ensemble multi-view feature set partitioning method for effective multi-view learning. Knowledge and Information Systems, 66(8), 4957-5001.

Song, H., Ahmad, A., Farooq, F., Ostrowski, K., Maślak, M., Czarnecki, S., & Aslam, F. (2021). Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. Construction and Building Materials, 308, 125021. https://doi.org/10.1016/J.CONBUILDMAT.2021.125021

Xue, X., & Zhou, H. (2018). Neuro-fuzzy based approach for estimation of concrete compressive strength. Computers and Concrete, 21(6), 697–703. https://doi.org/10.12989/CAC.2018.21.6.697

Zhang, W., Guo, J., Ning, C., Cheng, R., & Liu, Z. (2024). Prediction of concrete compressive strength using a Deepforest-based model. Scientific Reports, 14(1), 18918.