AI-driven Modeling for the Optimization of Concrete Strength for Low-Cost Business Production in the USA Construction Industry

Md. Habibur Rahman Sobuz
Department of Building Engineering and Construction Management, Khulna University of Engineering & Technology, Khulna – 9203, Bangladesh | Faculty of Engineering & Quantity Surveying, INTI International University (INTI-IU), Persiaran Perdana BBN, Putra Nilai, Nilai 71800, Negeri Sembilan, Malaysia
habib@becm.kuet.ac.bd (corresponding author)

Mohammad Abu Saleh
Department of Business Administration, International American University, 3440 Wilshire Blvd STE 1000, Los Angeles, CA 90010, USA
saleh7576@yahoo.com

Md. Samiun
Department of Business Administration, International American University, 3440 Wilshire Blvd STE 1000, Los Angeles, CA 90010, USA
engineersamiun@gmail.com

Mohammad Hossain
Department of Business Administration, International American University, 3440 Wilshire Blvd STE 1000, Los Angeles, CA 90010, USA
mhossain.eee@gmail.com

Anupom Debnath
Department of Business Administration, International American University, 3440 Wilshire Blvd STE 1000, Los Angeles, CA 90010, USA
debnathanupom05@gmail.com

Mahafuj Hassan
Department of Business Administration, International American University, 3440 Wilshire Blvd STE 1000, Los Angeles, CA 90010, USA
anikmahafuj8@gmail.com

Sanchita Saha
Department of Business Administration, Westcliff University, 17877 Von Karman Ave 4th Floor, Irvine, CA 92614, USA
s.saha.552@westcliff.edu

Rakibul Hasan
Department of Business Administration, Westcliff University, 17877 Von Karman Ave 4th Floor, Irvine, CA 92614, USA
r.hasan.179@westcliff.edu

Md. Kawsarul Islam Kabbo
Department of Building Engineering and Construction Management, Khulna University of Engineering & Technology, Khulna – 9203, Bangladesh
kabbo@becm.kuet.ac

Md. Munir Hayet Khan
Faculty of Engineering & Quantity Surveying, INTI International University (INTI-IU), Persiaran Perdana BBN, Putra Nilai, Nilai 71800, Negeri Sembilan, Malaysia
munir.hayetkhan@newinti.edu.my

Received: 26 November 2024 | Revised: 24 December 2024 | Accepted: 8 January 2025

ABSTRACT

The need to develop ecologically friendly sustainable building materials is made apparent by the worldwide construction industry's substantial contribution to global greenhouse gas emissions. The use of supplemental materials in concrete is one potential solution to lessen the environmental footprint. Thus, the purpose of this work is to use Machine Learning (ML) algorithms to forecast and create an empirical formula for the Compressive Strength (CS) of concrete with supplemental materials. Six distinct ML models—XGBoost, Linear Regression, Decision Tree, k-Nearest Neighbors, Bagging, and Adaptive Boosting—were trained and tested using a dataset that included 359 experimental data of varying mix proportions. The most significant factors used as input parameters are cement, aggregates, water, superplasticizer, silica fume, ambient curing, and supplemental material. Several statistical measures, such as Mean Absolute Error (MAE), coefficient of determination (R^2), and Mean Square Error (MSE), were used to evaluate the models. XGBoost model outperformed the other models with R^2 values of 0.99 at the training stage. To ascertain how the input parameters affected the outcome, feature importance analysis using Shapely Additive exPlanations (SHAP) was conducted. It was demonstrated that curing age and cement type significantly affected the strength of concrete with high SHAP values. By eliminating experimental procedures, reducing the demand for labor and resources, increasing time efficiency, and offering insightful information for enhancing sustainable manufacturing of concrete, this research advances the low-cost production of concrete in the USA construction industry.

Keywords-AI; construction materials; ML; business production; strength prediction

I. INTRODUCTION

Concrete is the most prevalent construction material due to its remarkable versatility, availability of raw resources, and minimal maintenance expenses [1]. Nonetheless, the worldwide annual output of over 25 billion tons of concrete manufacture has led to considerable environmental strain due to its contribution to CO_2 emissions [2]. Seven percent of annual CO_2 emissions originate from clinker production for Portland cement [3]. The concrete production industry is improving its compliance with regulations pertaining to sustainable development and energy saving. Minimizing cement use by partly replacing it with mineral admixtures or Supplementary Cementitious Materials (SCMs) might substantially aid in attaining this objective [4]. SCMs are used to mitigate environmental impact and improve the workability, mechanical attributes, and durability of concrete [5-7]. Ground granulated blast furnace slag, fly ash, bottom ash, glass powder, marble powder, granite powder, coral waste powder, palm oil clinker, and limestone powder have demonstrated efficacy in partially substituting cement in concrete to mitigate its adverse environmental and economic impacts [8, 9]. Consequently, it is logical to substitute a part of the cement with cost-effective, inert, and eco-friendly materials for sustainable development [10]. The intricate and evolving characteristics of cement hydration, along with our limited comprehension of pozzolanic reactivity, render it very challenging to simulate the mechanical properties of concrete including SMCs by empirical models [11]. The characteristics of the SMCs must be determined by comprehensive testing and laboratory studies, which may be costly, time-intensive, and arduous [12, 13]. Moreover, doing comprehensive laboratory work might be difficult due to specific constraints or variables, such as authorized zones for the storage and curing of concrete mixtures [14]. The Compressive Strength (CS) of concrete is affected by several factors, including fine and coarse aggregate composition, curing duration, and concrete mixture proportions, all of which need investigation, consuming time [15-17].

The data-driven Machine Learning (ML) technique in civil engineering has recently attracted significant interest. However, since ML algorithms just need data input and do not demand extensive theoretical examination, it is expected that they would exhibit comparable performance for concrete, owing to their robust predictive capabilities [18, 19]. The ML-based approach has the capability to provide a robust and reliable alternative for elucidating the complex link between input parameters and the desired output parameters using bigger data sets [20, 21]. Several prominent ML methods, including Artificial Neural Networks (ANNs), Random Forests (RFs), and Decision Trees (DTs), have been successfully used to address complex regression issues in civil engineering [22].

The RF model was identified as the most accurate for forecasting the mechanical parameters of roller-compacted concrete pavement in comparison to ANN models. Authors in [23] employed ML models based on SVM to predict the mechanical properties of granite waste-based concrete, resulting in a major accuracy enhancement against conventional predicting methods. Ensemble learning methods like Bagging and Boosting have also gained popularity for their robustness and reliability and can further improve prediction accuracy [24]. The introduction of Bagging [25] showed that an averaging ensemble could compensate for the increase in variance, thus increasing the model stability, a critical aspect of complex material behavior in civil engineering applications.

Boosting methods, especially Ada-Boost and Gradient Boosting (GB) have been successfully applied for CS prediction as they train initially weak models sequentially to improve the accuracy of the predictive model [26]. In [27], the authors focused on how Boosting algorithms can iteratively adjust model weights in order to reduce misclassifications, making them extremely useful and relevant for more complicated prediction problems where the data may contain various features that interact with each other. Authors in [28] used a Genetic Programming (GP) model, adaptive boosting (AdaBoost), and GB to forecast concrete strength. Authors in [29] developed shear strength prediction models for beam-column junctions with RF models. To predict the CS of concrete, authors in [30] used boosting-based techniques, including GB, AdaBoost, and Extreme Gradient Boost (XGB). XGB demonstrated superior outcomes with the highest R^2 value over 0.90. XGB had the highest accuracy among the ML models used in [31] for anticipating the shear capability of RC beams, with a RMSE of 1.346 and a MAE of 0.704. Authors in [24] discovered that Light Gradient Boosting (LGB) has superior performance in assessing the strength of 3D-printed concrete compared to XGB, Support Vector Regression (SVR), and RF models. Their research underscores the advantages of ML compared to traditional analytical or empirical methods.

Two significant deficiencies exist in the current ML techniques employed to predict the CS of concrete with additives: firstly, there is an absence of thorough parametric analyses such as SHapley Additive exPlanations (SHAP), and secondly, there is a necessity for more advanced and precise ML models. The results may lack generalizability since most previous studies used datasets that were either insufficiently sized or lacked enough data points. This study addresses this gap in the literature by using and comparing six ML models, namely XGB, DT, AdaBoost, Linear Regression (LR), k-Nearest Neighbors (kNN), and Bagging algorithms. The study used 359 data points from prior research and seven critical input parameters, which significantly enhanced the precision of the CS prediction. Furthermore, the models underwent a thorough verification process by juxtaposing the anticipated results with actual data samples and using a comprehensive array of statistical measures to evaluate performance during training and testing. To comprehend the factorial influence of input materials on the CS prediction, as well as optimize the mix materials, we assessed the significance of the characteristics. The outcomes of this study could be an effective solution for accelerating concrete production with minimal cost, which will introduce a new era in the US construction industry.

II. RESEARCH METHODOLOGY

Six distinct models were assessed to assess the most accurate ML model for assessing the CS of supplement-based concrete. The database employs a random division for testing and training applicable to all six ML models. The division is upheld at 80% for the training set and 20% for the testing dataset. Figure 1 presents the link network of the ML algorithms for this study.

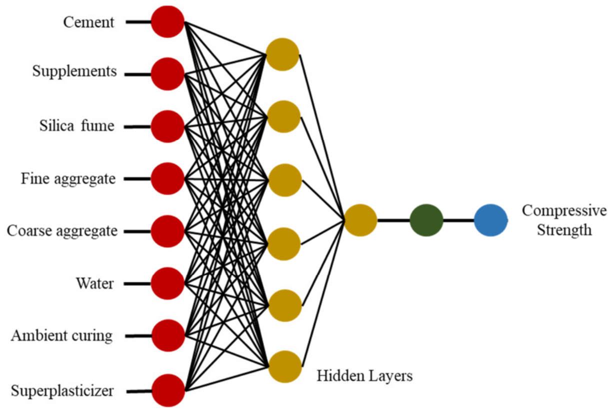

Fig. 1. ML algorithm working network.

A. *Machine Learning Algorithms*

AdaBoost is a boosting technique that uses DT regression as weak learners to assess data properties [32]. AdaBoost uses only a proportion of that data, making hypotheses on different subsets and adjusting them with respect to wrongly classified examples from previous classifiers, thus reducing the overfitting vulnerability.

The Bagging regressor is an ensemble technique that reduces overfitting by making predictions on multiple DTs. Though it increases the robustness of models, it adds sophistication and could be less This method reduces errors in initial predictions and gives estimates.

LR is one of the most fundamental and widely used statistical and ML methods for modeling the relationship between a dependent variable and one or more independent variables. The method assumes a linear relationship between the predictors and the response, expressed mathematically as:

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon \quad (1)$$

where y represents the dependent variable, x_1, $x_2 \ldots x_n$ are the independent variables, β_0 is the intercept, $\beta_1 \ldots\ldots \beta_n$ are the coefficients of the predictors, and ϵ is the error term accounting for unexplained variability. LR is widely appreciated for its simplicity, interpretability, and efficiency, particularly when the relationship between variables is approximately linear.

A DT is a non-parametric supervised learning method used for both classification and regression. The model utilizes recursively binary trees based on feature values and can handle both categorical and continuous data. In this framework, internal nodes represent the decision on an attribute, branches represent the outcome of the decisions, and leaf nodes represent the final prediction. Figure 2 illustrates the training and testing of the DT model, indicating that the minimum Root Mean Square Error (RMSE) for training occurred at 20 iterations. The minimum RMSE values for the training set were below 2.5 MPa.

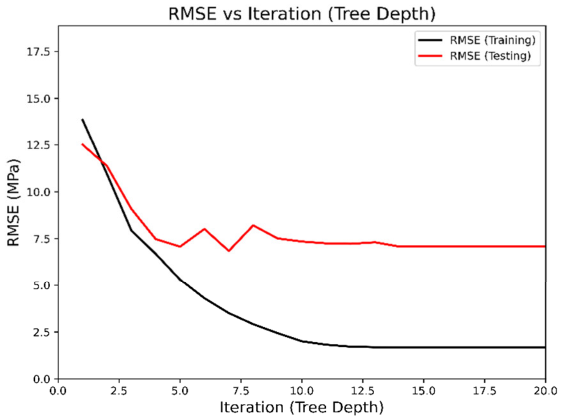

Fig. 2. RMSE vs iteration number for the DT model.

kNN is an easy, non-parametric ML classifier and repressor. The kNN algorithm classifies a data point according to how its neighbors are classified. Generally, the distance among points is commonly measured with Euclidean distance, but other metrics may also be used as appropriate [33]. kNN is noted for its conceptual simplicity and effectiveness, especially with problems where the decision boundary is not linear.

XGB is an optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable. With extra features such as tree pruning and weighted quantile sketch, XGB is good for heavy-duty modeling involving very large complex datasets. Due to these optimizations, XGB is well-regarded in ML competitions and is used in many practical use cases [34].

B. Performance Evaluation

The ML models' productivity and efficiency were assessed using MSE, R^2, and MAE. The relative contribution of each feature value to the model's prediction is ascertained by SHapley Additive exPlanations (SHAP). SHAP is a game theoretic approach that explains the output of any ML model. In order to do this, it uses some principles from game theory, particularly Shapley values, which allow us to attribute each feature's contribution to a model's predictions. SHAP is designed to provide insight into feature importance and interaction, which can be helpful in increasing model interpretability [35]. For the purpose of illustrating these feature contributions, SHAP offers several visual aids, including force, dependence, and summary plots.

III. RESULTS AND DISCUSSION

A. Dataset Analysis and Data Distribution

Table I presents the statistical measures of the database features, including parameters, units, standard deviation, mean, 25th percentile, median, and 75th percentile values. As can be seen, the cement content ranged from 139.6 to 540 kg/m^3, with an average of 311.94 kg/m^3. The coarse aggregate exhibited a maximum value of 1134, a minimum value of 801, and a mean value of 975.16. The supplements exhibited a mean density of 75.76 kg/m^3, with a range from 0 to 282.8 kg/m^3 and a standard deviation of 82.45 kg/m^3.

TABLE I. DATASET STATISTICS

Models	Mean	Deviation	25%	50%	75%
Cement (kg/m^3)	311.9490251	94.82850026	230	295.7	380
Supplements (kg/m^3)	75.76685237	82.45572056	0	53.8	132.4
Silica fume (kg/m^3)	47.86908078	57.74486475	0	0	100.4
Fine aggregate (kg/m^3)	776.5512535	94.56087066	755.8	780.6	852.2
Coarse aggregate (kg/m^3)	975.1671309	73.31961889	932	961.2	1030
Water (kg/m^3)	175.8891365	27.52027935	155.6	168.1	190.65
Superplasticizers (kg/m^3)	8.960445682	6.7239417	4.6	9.5	11.7
Ambient curing (days)	60.58774373	79.04126965	14	28	90
CS (MPa)	44.35523677	17.07372469	32.92	43.06	55.9

In this study, materials that have pozzolanic activity and the ability to enhance strength when used as a binder replacement have been considered and taken as supplements. For this particular study, two waste-based byproducts (waste glass powder and waste quartz powder) were considered as supplements. These materials similar compositions and when used as SCM enhance concrete strength [36, 37]. CS ranged from 7.4 MPa to 82.46 MPa. Other input variables exhibit moderate deviation, as indicated in Table I. The cement and fine aggregate exhibited the highest standard deviation values, exceeding 90 kg/m^3. Figure 3 illustrates the relationship between concrete strength values and various concrete mix components for the dataset considered in this study. As seen in Figure 3(a), cement content shows almost equal scattering between 200-500 kg/m^3. However, supplements visualize small clusters of data points between 100-200 kg/m^3 (Figure 3(b)). Silica fume showed a dense cluster between 100 and 125 kg/m^3, which indicates most of the mix proportions preferred silica fume dosage within 100-125 kg/m^3. Figures 3(d)-(h) also visualize the distribution of data points and their association with the strength of concrete. The illustration is important for understanding the data distribution and how every point of the input variables correlates with the output variables.

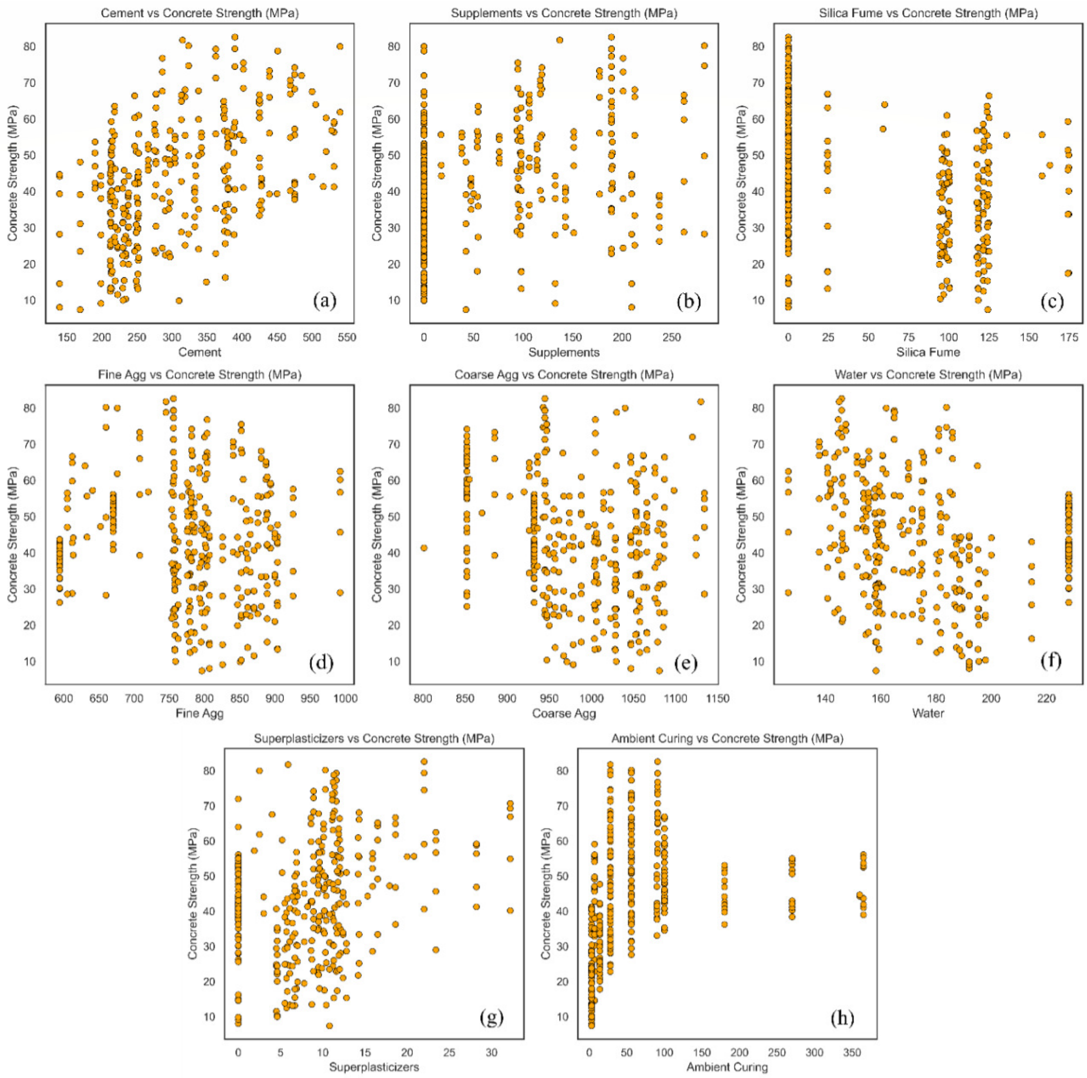

Fig. 3. Data distribution and the relationship between input and output variables.

B. Pearson Correlation among Features

The Pearson correlation coefficients between each variable were calculated, and the results are shown in Figure 4. High correlation coefficient values between the input variables, whether positive or negative, might result in approaches that are inefficient and make it more difficult to determine how the variables impact the answer. This demonstrates that the independent input variables do not display a substantial number of correlations with one another. On the basis of the findings, it is clear that there is a substantial association between the values of cement and concrete strength (0.48), which means that concrete strength mostly depends on the quantity of binders. Additionally, supplements and concrete strength also exhibited a high positive correlation value (0.34). Due to the pozzolanic activity of the supplement materials used in this study (glass powder and quartz powder), they also influence the strength and incorporating higher supplement content will also increase the strength of concrete [36, 38]. Superplasticizers (+0.31) and curing age (+0.25) also visualize positive association with strength. This aligns with prior research findings, which indicate that while superplasticizers contribute to reduced water content and improved particle packing, their influence is secondary to the core binder components [39, 40]. It should be noted that multicollinearity is a problem that occurs when predictors have a substantial connection with one another. This is an extremely important point to keep in mind.

C. Evaluating Machine Learning Models

The regression plots of the ML models are illustrated in Figure 5. The models' R^2 scores show different performance patterns. While both XGB and DT exhibit near-perfect R^2 scores in training (0.999), meaning they nearly perfectly fit the training data, their test R^2 scores (0.880 for XGB and 0.80 for DT) indicate a possible overfitting because they generalize to new data with a slightly lower effectiveness. The regression plots show better performance when the R^2 values are closer to 1 [39]. Bagging is very competitive with XGB because it finds an optimum balance between the test train dataset, with a test R^2 of 0.854 and a training R^2 of 0.981.

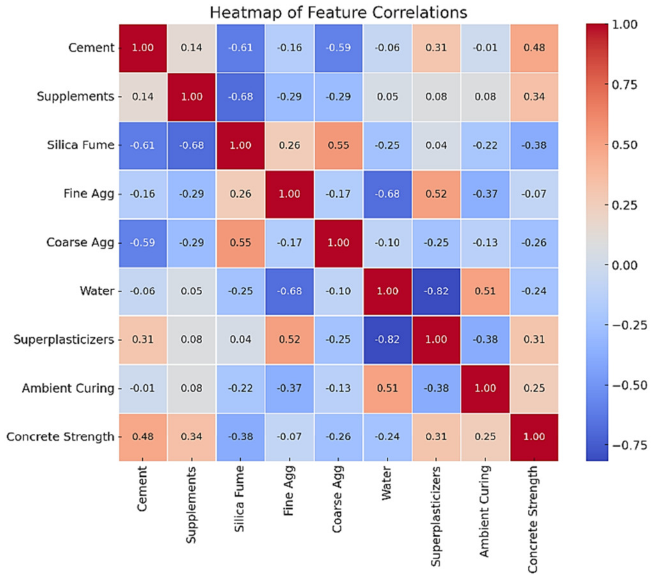

Fig. 4. Pearson correlation between features.

AdaBoost also shows satisfactory performance with an R^2 value of 0.890 (training) and 0.80 (testing). In contrast, LR shows poor performance with R^2 scores lower than 0.60, suggesting that it struggles to capture the underlying patterns. kNN has moderate R^2 scores, offering basic predictive power but falls behind models like XGB and BAG. Overall, while XGB was the superior model, the DT and Bagging models also showed very good prediction performance. In a past study on ML models' ability to predict the CS of palm oil fuel ash concrete [8], the XGB model had satisfactory R^2 values, which were very similar to the ones acquired in this study.

D. Performance Metrics

Figure 6 and Table II illustrate the error distribution of the models. The performance measures (MSE, MAE) for the ML models exhibit distinct variations. XGB has a minimal MSE of 0.003 during training, which escalates to 31.74 in testing. For the XGB model, the substantial increase in MSE and MAE from training (0.003 and 0.03, respectively) to testing (31.74 and 3.21) indicates moderate overfitting. This phenomenon may arise from XGB's ability to capture complex patterns during training, which, without proper regularization, could lead to overfitting of the training data and reduced generalization during testing. Regularization techniques, such as adjusting the learning rate or increasing tree constraints, could help mitigate this issue.

TABLE II. SUMMARY OF THE OUTPUT OF MODEL'S PERFORMANCE

Model	R^2 (Train)	R^2 (Test)	MSE (Train)	MSE (Test)	MAE (Train)	MAE (Test)
XGBoost	0.999	0.88	0.003614935	31.78424911	0.038395839	3.210272489
LR	0.544	0.476	135.0907087	139.0844915	9.399479655	9.039458656
DT	0.999	0.807	~0.00	51.217725	~0.00	4.772777778
kNN	0.779	0.652	65.35702198	92.27886189	6.192940767	7.119833333
Bagging	0.981	0.854	5.526606265	38.87828928	1.560139373	4.046611111
AdaBoost	0.89	0.802	32.48546927	52.43817956	4.734921248	5.593612652

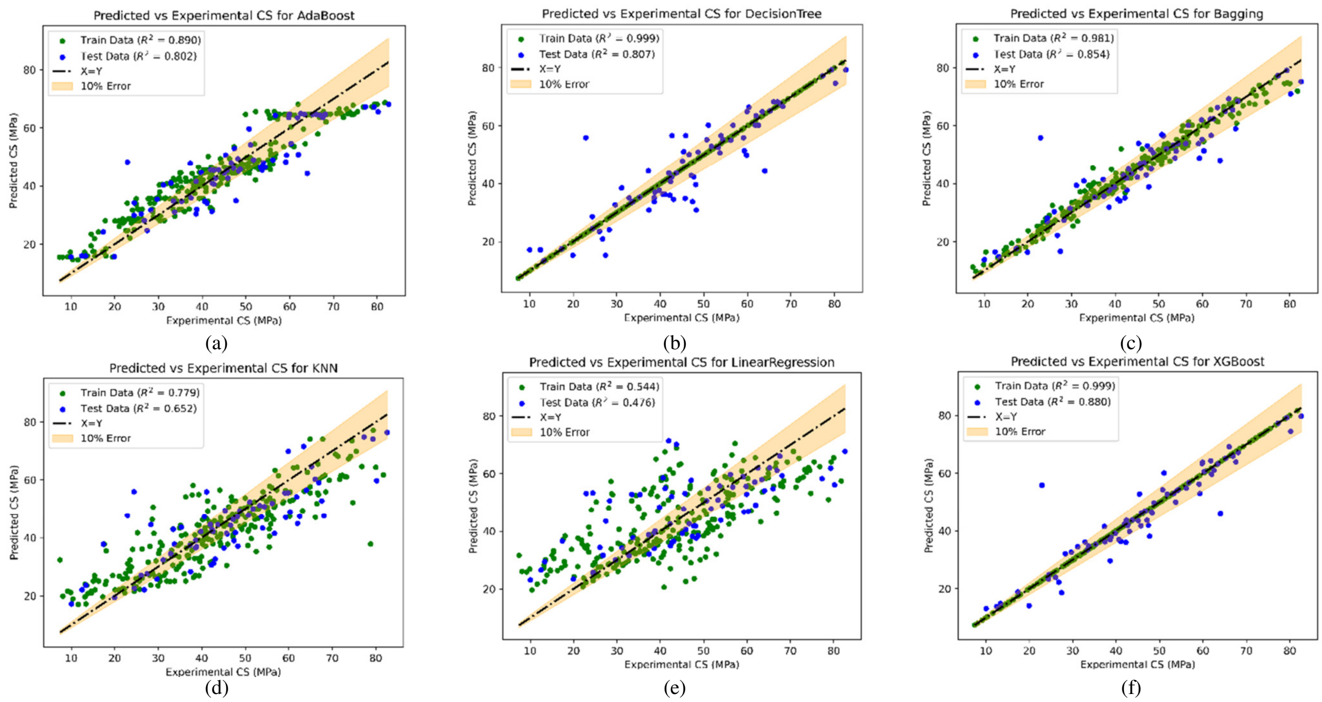

Fig. 5. Scatter plot of the proposed ML models concrete CS.

(a)

(b)

(c)

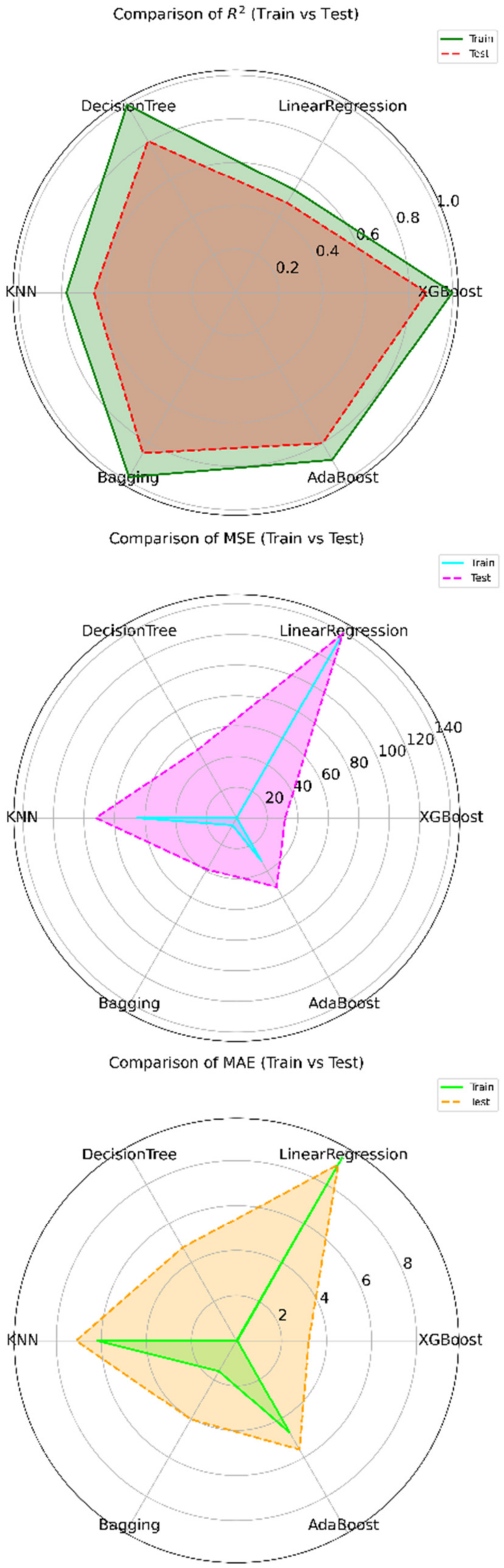

Fig. 6. Performance evaluation using (a) R^2, (b) MSE, and (c) MAE.

The DT model demonstrates near-perfect training performance, with MSE and MAE values approaching zero. However, its performance significantly deteriorates in testing, with MSE and MAE values of 51.21 and 4.77, respectively, highlighting a pronounced overfitting problem. This issue is inherent to DT, as they tend to form highly specific models tailored to training data. Ensemble techniques such as RF or GBR could address this challenge by reducing variance.

The GBR model strikes a better balance between training and testing performance, with moderate increases in MSE (from 5.52 to 38.87) and MAE (from 1.56 to 4.04). This suggests that GBR effectively captures the data's underlying structure without overly complex patterns that compromise generalization. Its stability stems from iterative optimization and inherent regularization mechanisms.

AdaBoost also demonstrates commendable performance, with relatively stable error metrics during both training (MSE: 32.48, MAE: 4.73) and testing (MSE: 52.43, MAE: 5.59). This performance stability reflects AdaBoost's ability to focus on difficult-to-predict instances while maintaining robustness across datasets.

In contrast, kNN and LR models exhibit relatively poor performance. kNN's higher MSE (65.35 in training, 92.27 in testing) and MAE (6.19 in training, 7.11 in testing) suggest it struggles with high-dimensional data or insufficiently optimized hyperparameters (e.g., number of neighbors). LR's extreme MSE (135.09 in training, 139.08 in testing) indicates an inability to capture non-linear relationships in the data, making it unsuitable for complex datasets.

Notably, the Bagging regressor emerges as the most balanced model, with consistently low errors across training and testing. This suggests that Bagging effectively reduces variance through ensemble learning, making it reliable for both training and unseen data.

E. Feature Analysis

The importance of each feature to the model's predictions of compressive strength of SCMC is represented graphically by the SHAP summary plots in Figure 7. Each feature's importance is displayed as a range of violin plots on the graph in the summary plot, where the x-axis indicates the SHAP level and the y-axis indicates feature relevance. Generally, dark red indicates lower importance, and light red points indicate higher importance. This figure shows the trend of various input materials for each data point with a particular SHAP value and illustrates how an input variable affects the result of a prediction algorithm. Ambient curing and cement have a significantly high influence on the CS of SCMC, with a SHAP value of over +10 at high feature values. This means they significantly enhance the strength of SCMC. Water showed a similar inverse trend, and a SHAP value of -10 was achieved at high feature values. Superplasticizers can be effective in reducing the required quantity of water, which enhances the compressive strength of concrete. Although coarse and fine aggregates have low SHAP values, they still positively impact the CS. Finally, supplemental materials have an almost neutral SHAP feature, which indicates they have no significant impact.

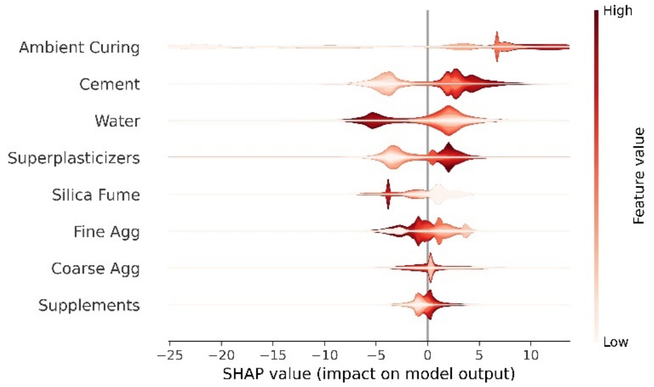

Fig. 7. SHAP feature analysis of variables.

IV. CONCLUSIONS

This study aimed to introduce predictive machine learning models for assessing the compressive strength of concrete incorporating supplements. XGB illustrated superior accuracy compared to other techniques, achieving the lowest mean errors of 0.03 during training and 3.21 during testing. In contrast, linear regression exhibited the highest mean errors, indicating problems with overfitting. The Bagging regressor demonstrated competitive performance, whereas models such as decision trees, k-nearest neighbors, and AdaBoost attained perfect minimum errors. XGB demonstrated outstanding accuracy with R^2 values of 0.999 for training and 0.88 for testing, although some overfitting was observed. AdaBoost and kNN exhibited satisfactory R^2 values of approximately 0.89 and 0.77 during testing. In contrast, Bagging and decision trees demonstrated effective performance, maintaining a balance between training and testing R^2 scores.

The analysis of mean SHAP values indicates that the two primary factors influencing the strength of SCMC are cement and the duration of ambient curing. The SHAP feature plot indicates that water and fine aggregate negatively affect compressive strength, whereas superplasticizers and cement significantly enhance strength at elevated feature values, as evidenced by SHAP values. In comparison to other variables, supplements exert a lesser influence on strength enhancement.

Future research may explore the incorporation of additional parameters or data sources into the machine learning models. Enhancing the predictability of compressive strength may involve incorporating data regarding the specific characteristics of the supplements utilized or the prevailing local environmental conditions. This study's findings provide an effective means to reduce experimentation costs and labor hours, thereby enhancing concrete production efficiency within the USA construction industry.

DATA AVAILABILITY

Data will be available on request.

REFERENCES

[1] D. Chen, Y. Chen, L. Ma, M. H. R. Sobuz, M. K. I. Kabbo, and M. M. H. Khan, "A state of review on manufacturing and effectiveness of ultra-high-performance fiber reinforced concrete for long-term integrity of concrete structures," *Advances in concrete construction*, vol. 17, no. 5, pp. 293–310, 2024, https://doi.org/10.12989/acc.2024.17.5.293.

[2] M. Mishra, "Review of Experimental and FE Parametric Analysis of CFRP-Strengthened Steel-Concrete Composite Beams," *Journal of Mechanical, Civil and Industrial Engineering*, vol. 3, no. 3, pp. 92–101, Dec. 2022, https://doi.org/10.32996/jmcie.2022.3.3.12.

[3] F. S. Aditto *et al.*, "Fresh, mechanical and microstructural behaviour of high-strength self-compacting concrete using supplementary cementitious materials," *Case Studies in Construction Materials*, vol. 19, Dec. 2023, Art. no. e02395, https://doi.org/10.1016/j.cscm.2023.e02395.

[4] A. E. Oke, J. Aliu, M. Z. Farhana, O. T. Jesudaju, and H.-P. Lee, "A capability assessment model for implementing digital technologies in Nigerian heavy construction firms," *Smart and Sustainable Built Environment*, Jul. 2024, https://doi.org/10.1108/SASBE-04-2024-0112.

[5] N. Farhan *et al.*, "Effect of silica fume on the microstructural and mechanical properties of concrete made with 100% recycled aggregates," *Revista de la construccion*, vol. 23, no. 2, pp. 413–435, 2024, https://doi.org/10.7764/rdlc.23.2.413.

[6] N. M. S. Hasan *et al.*, "Eco-friendly concrete incorporating palm oil fuel ash: Fresh and mechanical properties with machine learning prediction, and sustainability assessment," *Heliyon*, vol. 9, no. 11, Nov. 2023, Art. no. e22296, https://doi.org/10.1016/j.heliyon.2023.e22296.

[7] H.-P. Lee, E. Jusli, J. H. Ling, W. P. Low, and N. H. M. Kamaruddin, "Concrete Paving Blocks incorporating palm oil boiler ash and palm oil clinker as substitute concrete materials," *Journal of Engineering Science and Technology*, vol. 18, no. 6, pp. 75–81, Dec. 2023.

[8] Md. K. Kabbo, Md. H. Sobuz, and M. Khan, "Combined influence of Waste Marble Powder and Silica Fume on the Mechanical Properties of Structural Cellular Lightweight Concrete," in *International Conference on Planning, Architecture and Civil Engineering*, Rajshahi, Bangladesh, Oct. 2023, pp. 1–6.

[9] M. J. Rana, M. R. Hasan, and M. H. R. Sobuz, "An investigation on the impact of shading devices on energy consumption of commercial buildings in the contexts of subtropical climate," *Smart and Sustainable Built Environment*, vol. 11, no. 3, pp. 661–691, Mar. 2021, https://doi.org/10.1108/SASBE-09-2020-0131.

[10] Md. H. Rahman Sobuz *et al.*, "Experimental and analytical studies of size effects on compressive ductility response of Ultra-High-Performance Fiber-Reinforced concrete," *Construction and Building Materials*, vol. 409, Dec. 2023, Art. no. 133864, https://doi.org/ 10.1016/j.conbuildmat.2023.133864.

[11] A. Gogineni, I. K. Panday, P. Kumar, and R. Kr. Paswan, "Predicting compressive strength of concrete with fly ash and admixture using XGBoost: a comparative study of machine learning algorithms," *Asian Journal of Civil Engineering*, vol. 25, no. 1, pp. 685–698, Jan. 2024, https://doi.org/10.1007/s42107-023-00804-0.

[12] D. P. P. Meddage, I. Fonseka, D. Mohotti, K. Wijesooriya, and C. K. Lee, "An explainable machine learning approach to predict the compressive strength of graphene oxide-based concrete," *Construction and Building Materials*, vol. 449, Oct. 2024, Art. no. 138346, https://doi.org/10.1016/j.conbuildmat.2024.138346.

[13] Md. M. H. Khan, Md. H. R. Sobuz, M. M. Meraz, V. W. Y. Tam, N. Md. S. Hasan, and N. M. N. Shaurdho, "Effect of various powder content on the properties of sustainable self-compacting concrete," *Case Studies in Construction Materials*, vol. 19, Dec. 2023, Art. no. e02274, https://doi.org/10.1016/j.cscm.2023.e02274.

[14] M. K. Elshaarawy, M. M. Alsaadawi, and A. K. Hamed, "Machine learning and interactive GUI for concrete compressive strength prediction," *Scientific Reports*, vol. 14, no. 1, Jul. 2024, Art. no. 16694, https://doi.org/10.1038/s41598-024-66957-3.

[15] S. A. Ahmad, S. K. Rafiq, H. D. M. Hilmi, and H. U. Ahmed, "Mathematical modeling techniques to predict the compressive strength of pervious concrete modified with waste glass powders," *Asian Journal of Civil Engineering*, vol. 25, no. 1, pp. 773–785, Jan. 2024, https://doi.org/10.1007/s42107-023-00811-1.

[16] Md. A. Uddin, M. Jameel, H. R. Sobuz, Md. S. Islam, and N. Md. S. Hasan, "Experimental study on strength gaining characteristics of concrete using Portland Composite Cement," *KSCE Journal of Civil*

Engineering, vol. 17, no. 4, pp. 789–796, May 2013, https://doi.org/10.1007/s12205-013-0236-x.

[17] E. Ahmed and H. R. Sobuz, "Flexural and time-dependent performance of palm shell aggregate concrete beam," *KSCE Journal of Civil Engineering*, vol. 15, no. 5, pp. 859–865, May 2011, https://doi.org/10.1007/s12205-011-1148-2.

[18] A. Saxena, C. Sabillon-Orellana, and J. Prozzi, "Prediction of compressive strength in sustainable concrete using regression analysis," *Journal of Material Cycles and Waste Management*, vol. 26, no. 5, pp. 2896–2909, Sep. 2024, https://doi.org/10.1007/s10163-024-02010-9.

[19] M. Arif *et al.*, "Data-driven models for predicting compressive strength of 3D-printed fiber-reinforced concrete using interpretable machine learning algorithms," *Case Studies in Construction Materials*, vol. 21, Dec. 2024, Art. no. e03935, https://doi.org/10.1016/j.cscm.2024.e03935.

[20] M. Abdellatief, Y. M. Hassan, M. T. Elnabwy, L. S. Wong, R. J. Chin, and K. H. Mo, "Investigation of machine learning models in predicting compressive strength for ultra-high-performance geopolymer concrete: A comparative study," *Construction and Building Materials*, vol. 436, Jul. 2024, Art. no. 136884, https://doi.org/10.1016/j.conbuildmat.2024.136884.

[21] M. Alyami *et al.*, "Innovative hybrid machine learning models for estimating the compressive strength of copper mine tailings concrete," *Case Studies in Construction Materials*, vol. 21, Dec. 2024, Art. no. e03869, https://doi.org/10.1016/j.cscm.2024.e03869.

[22] Md. H. R. Sobuz, M. Khatun, Md. K. I. Kabbo, and N. M. Sutan, "An explainable machine learning model for encompassing the mechanical strength of polymer-modified concrete," *Asian Journal of Civil Engineering*, pp. 1–23, Nov. 2024, https://doi.org/10.1007/s42107-024-01230-6.

[23] W. Ben Chaabene, M. Flah, and M. L. Nehdi, "Machine learning prediction of mechanical properties of concrete: Critical review," *Construction and Building Materials*, vol. 260, Nov. 2020, Art. no. 119889, https://doi.org/10.1016/j.conbuildmat.2020.119889.

[24] M. H. R. Sobuz *et al.*, "Optimization of recycled rubber self-compacting concrete: Experimental findings and machine learning-based evaluation," *Heliyon*, vol. 10, no. 6, Mar. 2024, Art. no. e27793, https://doi.org/10.1016/j.heliyon.2024.e27793.

[25] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, Aug. 1996, https://doi.org/10.1007/BF00058655.

[26] S. D. Datta *et al.*, "Analysis of the characteristics and environmental benefits of rice husk ash as a supplementary cementitious material through experimental and machine learning approaches," *Innovative Infrastructure Solutions*, vol. 9, no. 4, Mar. 2024, Art. no. 121, https://doi.org/10.1007/s41062-024-01423-7.

[27] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, Aug. 1997, https://doi.org/10.1006/jcss.1997.1504.

[28] B. V. Varma, E. V. Prasad, and S. Singha, "Study on predicting compressive strength of concrete using supervised machine learning techniques," *Asian Journal of Civil Engineering*, vol. 24, no. 7, pp. 2549–2560, Nov. 2023, https://doi.org/10.1007/s42107-023-00662-w.

[29] W. Z. Taffese and L. Espinosa-Leal, "Multitarget regression models for predicting compressive strength and chloride resistance of concrete," *Journal of Building Engineering*, vol. 72, Aug. 2023, Art. no. 106523, https://doi.org/10.1016/j.jobe.2023.106523.

[30] F. Aslam and M. Z. Shahab, "Supplementary cementitious materials in blended cement concrete: Advancements in predicting compressive strength through machine learning," *Materials Today Communications*, vol. 38, Mar. 2024, Art. no. 107725, https://doi.org/10.1016/j.mtcomm.2023.107725.

[31] R. K. Tipu, Suman, and V. Batra, "Development of a hybrid stacked machine learning model for predicting compressive strength of high-performance concrete," *Asian Journal of Civil Engineering*, vol. 24, no. 8, pp. 2985–3000, Dec. 2023, https://doi.org/10.1007/s42107-023-00689-z.

[32] F. Alsharari, "Predicting the compressive strength of fiber-reinforced recycled aggregate concrete: A machine-learning modeling with SHAP analysis," *Asian Journal of Civil Engineering*, vol. 26, no. 1, pp. 179–205, Jan. 2025, https://doi.org/10.1007/s42107-024-01183-w.

[33] T. M. Cover, "Nearest neighbor pattern classification," *IEEE Trans. Information Theory*, vol. 4, no. 5, pp. 515–516, 1968.

[34] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, Dec. 2016, pp. 785–794, https://doi.org/10.1145/2939672.2939785.

[35] M. Mishra, "Quantifying compressive strength in limestone powder incorporated concrete with incorporating various machine learning algorithms with SHAP analysis," *Asian Journal of Civil Engineering*, Dec. 2024, https://doi.org/10.1007/s42107-024-01219-1.

[36] K. I. M. Ibrahim, "Recycled waste glass powder as a partial replacement of cement in concrete containing silica fume and fly ash," *Case Studies in Construction Materials*, vol. 15, Dec. 2021, Art. no. e00630, https://doi.org/10.1016/j.cscm.2021.e00630.

[37] A. S. M. Akid, S. M. A. Shah, M. D. H. R. Sobuz, V. W. Y. Tam, and S. H. Anik, "Combined influence of waste steel fibre and fly ash on rheological and mechanical performance of fibre-reinforced concrete," *Australian Journal of Civil Engineering*, vol. 19, no. 2, pp. 208–224, Jul. 2021, https://doi.org/10.1080/14488353.2020.1857927.

[38] M. H. R. Sobuz, S. D. Datta, and A. S. M. Akid, "Investigating the combined effect of aggregate size and sulphate attack on producing sustainable recycled aggregate concrete," *Australian Journal of Civil Engineering*, vol. 21, no. 2, pp. 224–239, Jul. 2023, https://doi.org/10.1080/14488353.2022.2088646.

[39] Md. H. R. Sobuz *et al.*, "High-Strength Self-Compacting Concrete Production Incorporating Supplementary Cementitious Materials: Experimental Evaluations and Machine Learning Modelling," *International Journal of Concrete Structures and Materials*, vol. 18, no. 1, Oct. 2024, Art. no. 67, https://doi.org/10.1186/s40069-024-00707-7.

[40] Md. H. R. Sobuz *et al.*, "Enhancing Sustainable Concrete Production by Utilizing Fly Ash and Recycled Concrete Aggregate with Experimental Investigation and Machine Learning Modeling," *Journal of Building Pathology and Rehabilitation*, vol. 9, no. 2, Aug. 2024, Art. no. 134, https://doi.org/10.1007/s41024-024-00474-8.