Full Paper of the WasteSafe 2025 – 9th International Conference on
Sustainable Solutions for Waste Management in Developing Countries
22-23 February 2025, Dhaka, Bangladesh

A NOVEL MACHINE LEARNING APPROACH FOR PREDICTING COMPRESSIVE STRENGTH OF CONCRETE REINFORCED WITH WASTE PET FIBERS

Bappi Sur[1], Md. Kawsarul Islam Kabbo[2] and Md. Sagidul Islam[3]

[1,2,3]Building Engineering and Construction Management, Khulna University of Engineering & Technology, Khulna, Khulna-9203, Bangladesh.

ABSTRACT

This paper investigates the potential of machine learning (ML) to predict the compressive strength of concrete reinforced with waste polyethylene terephthalate (PET) fibers, contributing to the development of sustainable construction materials. Six ML models (Random Forest (RF);XGBoost; LightGBM; GBM; Decision Tree; and AdaBoost) were evaluated using a derived 204 datasets based on fiber length, cement, sand, aggregate, superplasticizer content, water/cement ratio, PET volume percentage and curing days. Model performance was assessed via statistical metrics, with XGBoost, GBM and RF exhibiting the reliable performance ($R^2 = 0.925$,0.926 and 0.932 respectively) while AdaBoost shows the least accuracy ($R^2 = 0.838$). The most influential parameters identified via SHAP analysis were curing days and cement content. ML is a powerful tool for enhanced performance design, resulting in the development of affordable and green materials for future applications. This method enables the use of PET waste in durable, high-performance concrete, thus addressing the environmental issues caused by PET waste disposal.

INTRODUCTION

Concrete is the vast used construction material in the globe due to its high compressive strength, durability and cost effectiveness. But it has some limitations of low tensile strength and poor crack resistance. To overcome these drawbacks extensive research has been done on fiber reinforced concrete. Fibers made from various polymers such as nylon, aramid, polyesters (Amalraj & Ilangovan, 2023; Qin et al., 2021) and materials like glass, carbon and other naturally occurring substances are commonly used in concrete. Also the mechanical properties of the fibers and the bond between the cement and fiber matrices plays a significant role in the structural performance of the concrete (Foti, 2019; Sadeghian, Tanyous, & Mirshekar, 2020). Incorporation of fibers as reinforcement is one of the most effective way to enhance the mechanical properties of cement based composites. Moreover, “Fiber spacing theory” (Liu & Lu, 2012) and “Composite material theory” (Hashin & Wendt, 1970) reveals that, fibers improve crack resistance and toughness in cement based composites and prevents crack initiation and propagation and enhances the mechanical strength and durability.

Recently the problem of reusing waste plastics has been getting a lot of attention from researchers because of the environmental and economic impact of the construction industry (Ahmed, 2023; Lamba, Kaur, Raj, & Sorout, 2022). Globally Polyethylene Terephthalate (PET) production has surpassed 6.7 million tons a year and especially in Asia, China and India are seeing a big surge due to increasing demand (Lai et al., 2022). In South Korea, PET bottle production is around 130,000 tons a year (J. Zhang, Wang, & Kannan, 2021). But most PET bottles used for beverage are used once and discarded in mostly landfill and incineration, which is causing big environmental problems (Marsh & Bugusu, 2007). In these circumstances, a cost effective solution is needed for PET bottle wastes. As PET fibers have gained worldwide use in construction applications including tunnel linings, selfcompacting concrete, making light RCC beam and pavements (Mwangi, 2001; W. Zhang et al., 2024), Using recycled PET as short fiber reinforcement in construction industry is one potential approach.

In case of implementation of PET fiber in concrete mix design, there are many factors, as example, length of the fiber, the percentage of fibers, and their shape and orientation, and bond within the cement mixing as well as concrete mix design affect in improving the mechanical and durability properties of concrete. Some researchers have attempted to develop cost-effective and sustainable methods by reducing the amount of cement and varying the parameter of PET fiber in different size, shape and volume as substitutes to achieve the desirable strength. However, in order to precisely estimate the concrete's strength, varied mixes frequently need a large number of tests, analysis and resources. It takes a significant expense commitment and time to precisely determine the relationship between concrete's properties and mix design using experimental approaches. That is why in order to estimate values of compressive strength with ease and accuracy, a modeling approach based on the components of concrete along with PET fiber must be developed.

In construction industry, Several researches have revealed vast use of machine learning to make model between components of mix design and mechanical properties (W. Ahmad et al., 2021; Pakzad, Roshan, & Ghalehnovi, 2023; Peng & Unluer, 2022), and also attempted to predict the mechanical characteristics using various substitutes like cementitious waste materials (Aslam & Shahab, 2024), steel fibers (Kang, Yoo, & Gupta, 2021), geopolymers (Rahmati & Toufigh, 2022) etc. in various types of concrete like self-compacting concrete (Ali, Faraj, Saeed, Ahmed, & Ahmed, 2024), high-performance concrete (Yang et al., 2023), recycled aggregate concrete (Sau, Shiuly, & Hazra, 2023) etc. Apart from these, there are a few studies on making model to predict Compressive Strength (CA) of concrete using PET fibers as a reinforcement. To overcome this limitation, this study utilizes six machine learning models named Random Forest (RF), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Gradient Boosting Machine (GBM), Decision Tree (DT) and AdaBoost (Adaptive Boosting) using 204 sets of data collected from previous studies on PET fiber utilization in concrete as reinforcement. To build the model, 7 parameters were used as inputs which are length of PET, amount of cement, sand and aggregate in kg/m^3, water to cement ratio (w/c), percentage of PET of total volume, amount of super plasticizer (sp) and curing days. The target value was compressive strength (CS). All models were trained with 70% datasets, tested with 15% and validated with the rest 15%. The performance of this models are evaluated by statistical and graphical approaches. Additionally, ML algorithms were paired with Shapley Additive Explanations (SHAP) to examine the importance and influence of both individual and combination variables on CS predictions.

METHODOLOGY

Extreme Gradient Boosting (XGBoost OR XGB)

This method is based on the extreme gradient boosting decision tree and was first described by Chen and Guestrin (Duong, Tran, Satomi, & Takahashi, 2022) as part of the overall gradient boosting machine (GBM) approach and has thus far proven to be a highly accurate, efficient and adaptable algorithm. Meaning, XGB uses the boosting method and constructs its models in a sequential manner where the weak learners are improved to build a very accurate final forecastor (Duong et al., 2022). A prominent aspect is its second-order Taylor expansion for loss functions, which takes advantage of first and second-order derivatives for better predictions. To avoid overfitting regularization is added, providing robust performance according to Eq. (1) (Ahmed, 2023; N.-H. Nguyen, Abellán-García, Lee, Garcia-Castano, & Vo, 2022).

$$f_{ip} = \sum_{lk=1} f_k(x_i) = f_{i(p-1)} + f_p(x_i) \qquad (1)$$

Where the$f_p(x_i)$acts as the learner at phase p, f_i^p and $f_i^{(p-1)}$ represent the predicted value p − 1 at phase p, and the input variable is x_i acts.

The XGB model develops a method to determine the "goodness" of the chosen model (Ekanayake, Meddage, & Rathnayake, 2022), preventing overfitting issues without sacrificing data processing speed.

$$obj_{(p)} = \sum_{nk=1} l\,(\overline{y_i}\,, y_i) + \sum_{pk=1} \sigma\, f_i \qquad (2)$$

where I stands for the following equation defines LF, n for the number of observations, and σ used for the term of uniformity

$$\sigma(f) = \gamma T + \frac{1}{2}\ \lambda||\omega||2 \qquad (3)$$

In this context, **ω** represents the vector scores in the leaves, **ɣ** denotes the minimum loss used for splitting a leaf node, and **λ** refers to the regularization parameters.

Light Gradient Boosting Machine (LightGBM OR LGB)

The LightGBM (LGB) algorithm is another tree-based learning method combining gradient boosting and decision trees, achieving high accuracy performance, efficiency and scalability for regression tasks. Its unique tree building algorithm using histogram-based discretization enables finer tree construction using leaf-wise growth. The advantages of LGB are faster training, lower memory, and supporting parallel training (Ahmed, 2023; N.-H. Nguyen et al., 2022). The predicted decline in specifically termed this details gain, happens when the nodes are split based on their functions and can easily be computed as. $IG(C,A) = Fn(C) - \sum V \in Values(V) \frac{|Cv|}{C}\ Fn(Cv)$ (4)

$$Fn(C) = \sum_{b=1}^{B} -pblog_{2}pd \tag{5}$$

Here subset of C represented by Cv where the feature takes the value v, v is the attribute value, B denotes the classes, pb is the proportion of C in class b, and C refers to the entropy information of the collection

Gradient Boosting Machine (GBM)

Gradient Boosting Machine (GBM) is a forward learning ensemble method in machine learning for regression and classification (Yang et al., 2023). It builds models by combining decision trees and learning from residual errors unlike random methods (Kaloop, Kumar, Samui, Hu, & Kim, 2020). Similar to AdaBoost, GBM adds classifiers to correct prior errors but optimizes using gradient based techniques. Regression trees are the weak learners in a stage wise additive process and the model is improved by iterative error minimization (Kaloop et al., 2020; H. Nguyen, Vu, Vo, & Thai, 2021).

$$f_{m}(x) = f_{(m-1)}x + K_{m}(x) \tag{6}$$

$$f_{m}(x) = f_{(m-1)}x + \alpha K_{m}(x) \tag{7}$$

The step-by-step procedure used in GB to train additive algorithms is described in Eq.(6). Regression decision trees, often known as Km(x), are algorithms' weakest learners. Each iteration of the GBM algorithms adds a new estimator to the network by aggregating "m" weak learners. A regulatory variable called "learning rate" is also used in the process, as shown in Eq. (7). training the GB model minimizes the influence of each decision tree on the final prediction.

Adaptive Boosting (AdaBoost)

In 2003, Yoav Freund and Robert Schapire introduced AdaBoost, or adaptive boosting, for statistical classification (A. Ahmad, Ahmad, Aslam, & Joyklad, 2022; Rathakrishnan, Bt. Beddu, & Ahmed, 2022). This ensemble method reduces bias and variance by iteratively training weak learners. Each sub model adjusts sample weights based on prior model errors and creates a series of learners trained on weighted data (Khan, Lao, & Dai, 2022). After N iterations, weak learners are combined with weighted contributions to form a strong model. Although individual weak learners perform near random, their combined weighted predictions are strong. The weak learner Gk(X) estimates the forecasting errors iteratively according to Eq. (8),

$$e_{i} = |\underline{\qquad\quad}_{Y_{i}-G_{E}(X_{i})|} \tag{8}$$

Where, E = max$|Y_{i} - G(K_{i})|$ represents the maximum deviation from the actual value across all forecasts, where Y_i is the outcome, X_i the input vector, and the error ratio e_k is calculated as

$e_{k} = \sum_{i=1}^{m} e_{ki}$ (9) $^{e^{k}}$ (10)
$\alpha_{k} = {}_{1-e_{k}}$

The weight w_k=1,i for each data sample in the next training step is updated as.

$$w_{k} = 1,\ i = \sum\underline{\qquad\qquad}_{mi=1\ w_{k,ki}\ \alpha 1-e_{ki}} \overset{w_{k,i}\ \alpha 1-e_{ki}}{} \underset{k}{} \tag{11}$$

H_x is defined A strong learner as follows

$$H_{x} = v \sum{}^{N}{}_{k=1}(In^{1}{}_{\alpha_{k}})\ g(X) \tag{12}$$

Here, g(X) is the median of $\alpha_k G_k(X)$ for k = 1, 2, 3, …, where α_k represents the weights of weak learners. The regularization parameter $v \in (0, 1]$ helps prevent overfitting.

Decision Tree (DT)

DT uses data from training to create a model that resembles a tree. The root node, decision node (sometimes it also called the internal node), and terminal node (also called the leaf node) are the three nodes that make up the DT model. The approach divides into decision nodes after starting at the root node, which contains all training data. The method starts at the root node, which contains all of the training data, and then divides into decision nodes. The process is continued for each subsequent level until the tree reaches a predetermined maximum depth or the nodes contain only a single training sample (Tran, 2023).

Random Forest (RF)

Bagging with decision trees leads to the Random Forest (RF) algorithm, which averages the predictions of many trees are combined for higher accuracy (Chopra, Sharma, Kumar, & Chopra, 2018). RF performs well in regression and classification (Han, Gui, Xu, & Lacidogna, 2019; Xu et al., 2021), particularly when applied to high-dimensional data(Han et al., 2019), as it develops trees by randomly selecting predictors. By training models from bootstrapped samples of the dataset, and aggregation using the bagging algorithm, RF successfully minimizes model instability and avoids overfitting, thereby producing robust and reliable predictions (Rahman, Ahmed, Khan, Islam, & Mangalathu, 2021).

$$Y = \frac{1}{B}\sum_{j=i}^{b}(Y_b)\,X' \tag{13}$$

Shapley Additive Explanations (SHAP)

SHAP is based on principles from cooperative game theory, concretely based on a mathematical approach for determining the added value of each player on a joint outcome. The participants in this game are called input features in the field of machine learning, and the game itself is called predictive outcome. It analyzes how much each feature contributes to a model by observing how SHAP defines an explanation model as a linear combination of input features, showing how model predictions change when a feature is included or excluded

$$g(x') = \emptyset_0 + \sum^{M}_{i=1}(\emptyset_i\, x_i') \tag{14}$$

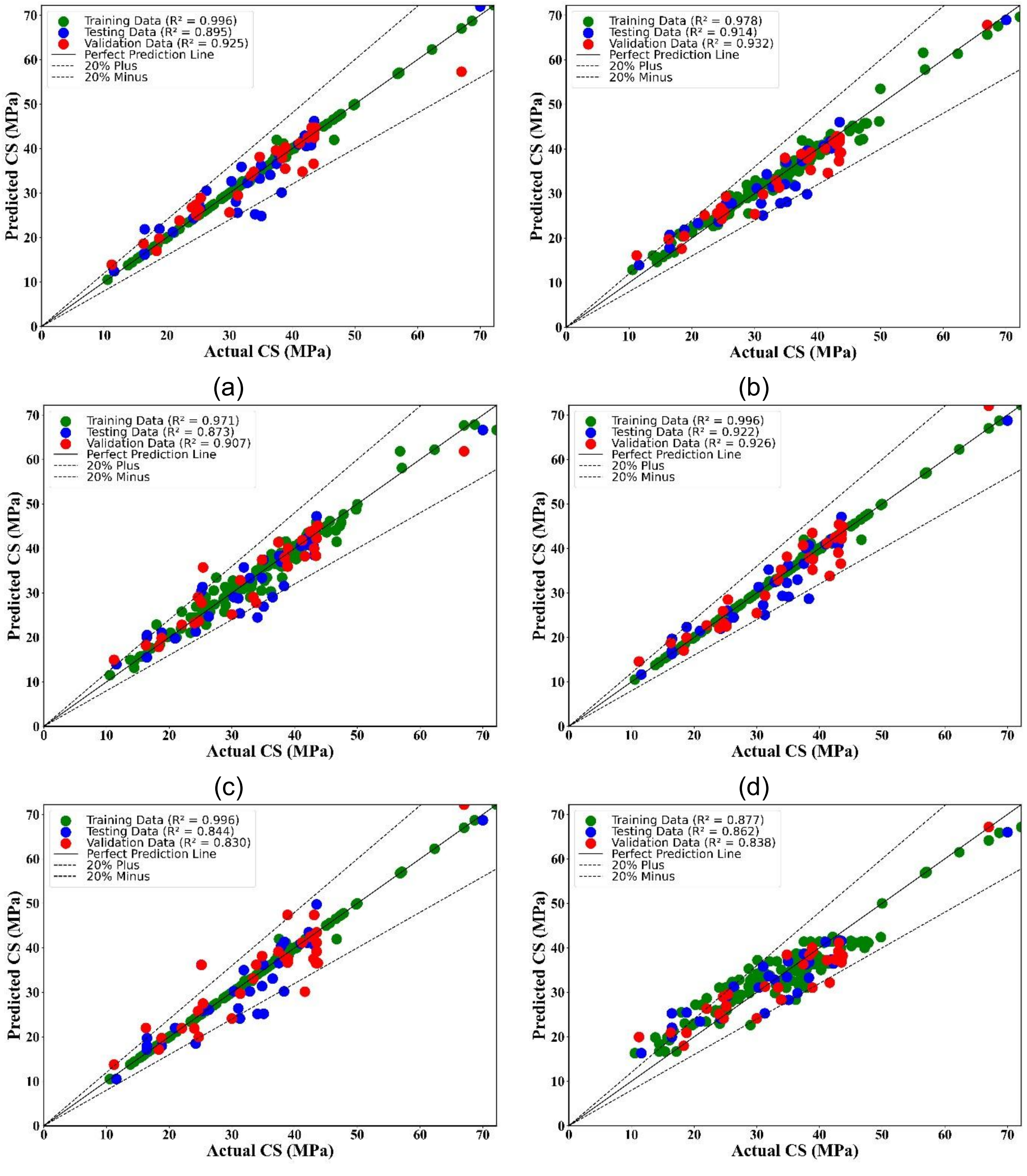
In this context, Shapley values (ϕi) act as coefficients in a linear model, where x′ represents the input and g is the explanation model

$$\phi_i = \sum_{s\varepsilon F\{i\}} \frac{S!(F-S-1)!}{F!}\,[f_{SU\{i\}}(x_{SU\{i\}} - f_s(x_s)] \tag{15}$$

Here, F signifies the complete set of input features, while S is a subset of F, excluding the parameters with index i .

RESULTS AND DISCUSSION

After predicting the values of CS on 6 models, “Actual CS vs Predicted CS” graphs were plotted. We can see in figure 1(a) that, for XGBoosting model, with an R^2=0.996, the training data is represented by the green points, suggesting a nearly perfect match although this tight match between the actual and projected could also be an indication of excessive fitting. The testing dataset is indicated by the blue points, and the model's capacity to generalize to new data is demonstrated by its R^2=0.895. In a similar vein, the validation dataset, represented by the red points, exhibits good and reliable accuracy with an R^2=0.925. With the values of R^2 about 0.978, 0.914 and 0.932 for training, testing and validation data respectively in Random Forest model, all points are staying at near the perfect prediction (PP) line in figure 1(b) which represents the model's accuracy and robustness also. GBM also performed reliable prediction with the values of R^2 about 0.996, 0.922 and 0.926 for training, testing and validation data respectively and figure 1(d) representing that almost all of the points are staying near at the PP line. LGB model predicted less reliable than XGB, GBM and RF with less R^2 value. Figure 1(e) reflects the overfitting in training data with DT model reasoning reduction in values of R^2 in testing and validating data. AdaBoost gives the worst prediction level which is indicated by figure 1(f).



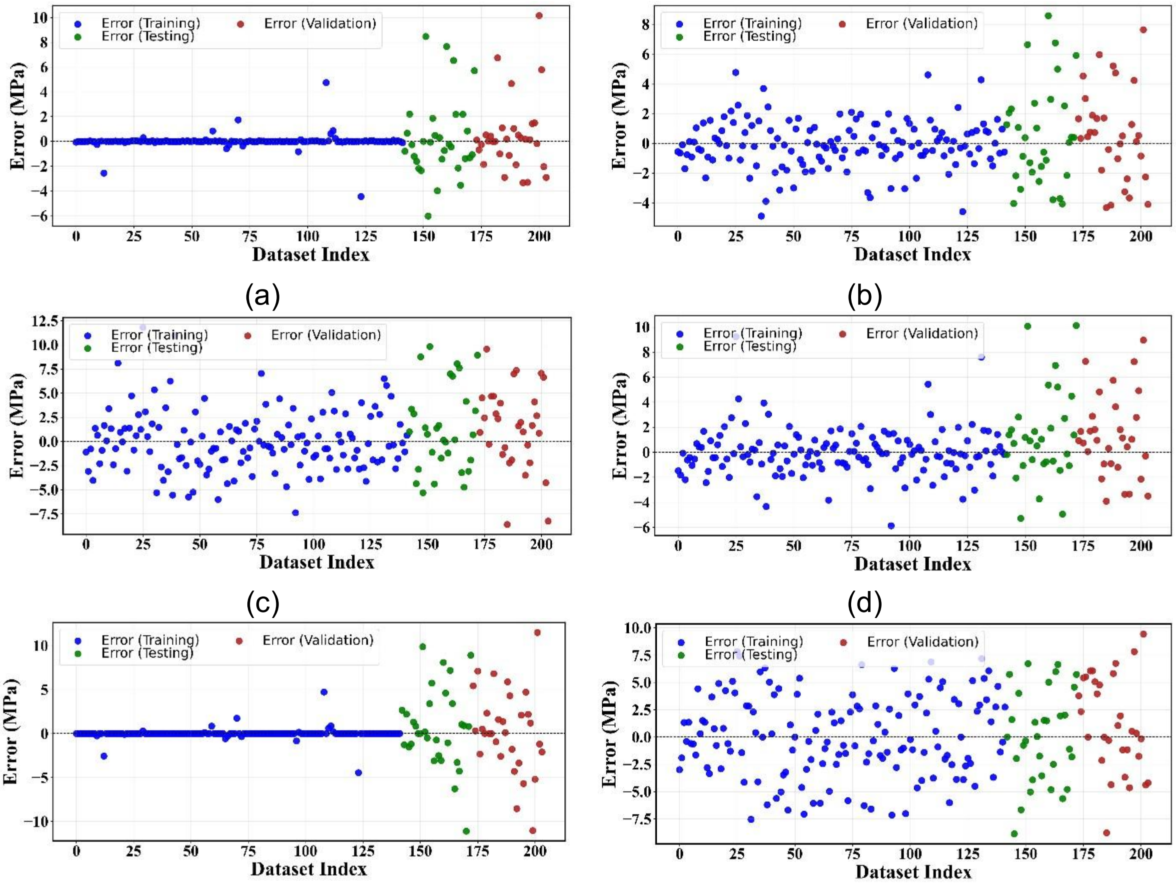
(a) (b)

(c) (d)

(e) (f)

Figure 1 Relationship between observed and predicted CS values for (a) XGBoosting, (b) RF, (c) LightGBM , (d) GBM , (e) DT , (f) AdaBoosting

Almost 75-80% data has error less than 6 MPa for XGBoosting, RF, GBM and DT model while rest two has more error as the points stay far behind from the zero error line in errors dispersion graphs shown in figure 02. As the error increases, the values of Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE) also increase. Statistical values of developed models are given in table 01. Higher values of these parameters indicate lower prediction level and it is clear from the comparison that, XGB, GBM and RF performed better prediction than others.



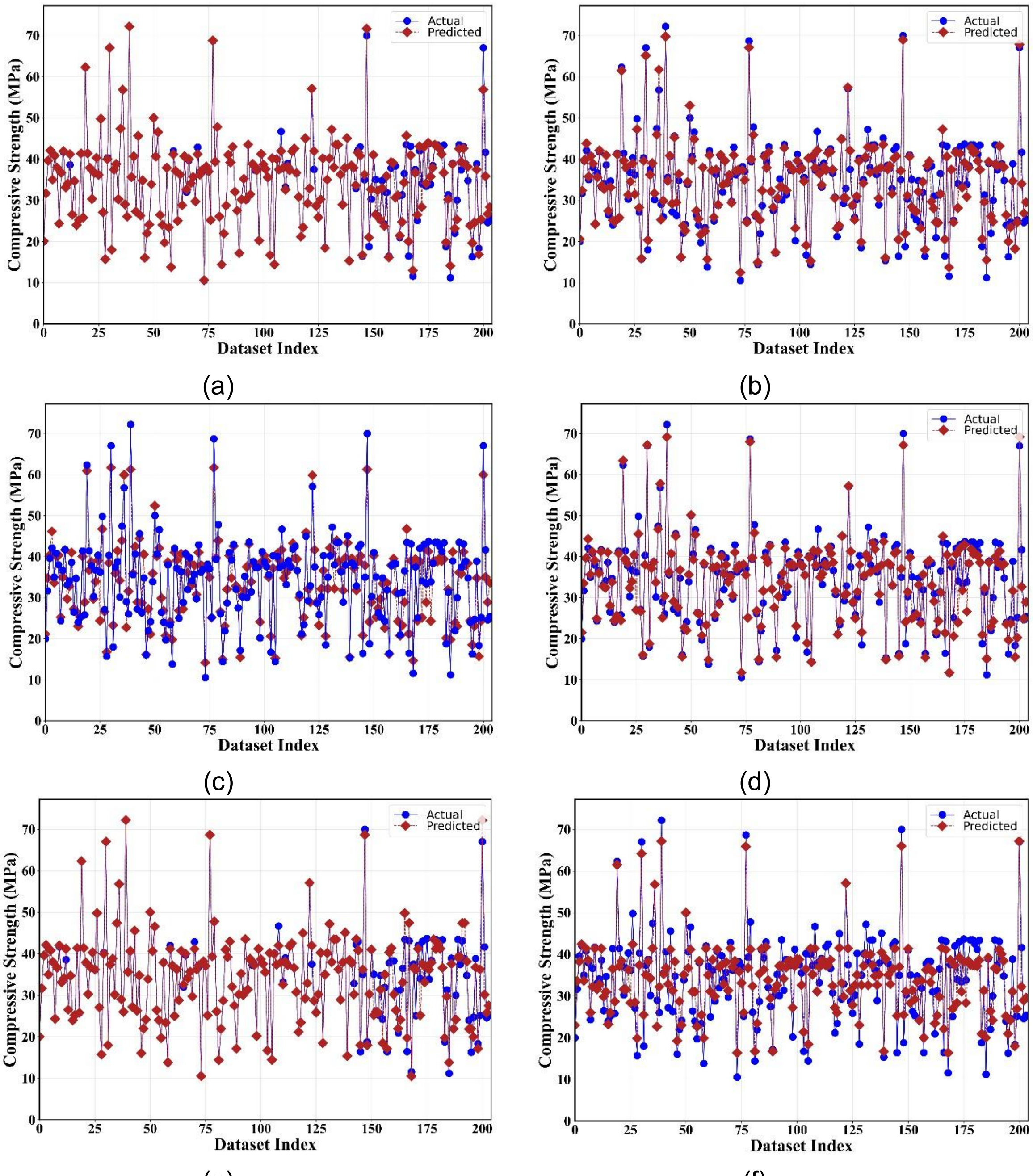
(a)
(b)
(c)
(d)
(e)
(f)

Figure 2 Errors dispersion graphs for (a) XGBoosting, (b) RF, (c) LightGBM , (d) GBM , (e) DT , (f) AdaBoosting Table 1. Statistical description of developed models.

Parameter	XGBoosting	RF	LightGBM	GBM	DT	AdaBoost
R^2(Training)	0.996	0.978	0.971	0.996	0.996	0.877
R^2 (Testing)	0.895	0.914	0.873	0.922	0.844	0.862
R^2(Validation)	0.925	0.932	0.907	0.926	0.830	0.838
MAE(Training)	0.150	1.204	1.238	0.134	0.134	3.061
MAE(Testing)	2.259	2.676	3.154	2.353	3.258	3.483
MAE(Validation)	1.814	2.424	2.672	2.366	3.433	3.659
MAPE(Training)	0.388%	3.848%	3.864%	0.335%	0.335%	10.256%
MAPE(Testing)	7.455%	9.521%	10.875%	7.535%	10.631%	13.389%
MAPE(Validation)	5.939%	8.459%	8.930%	7.692%	10.945%	12.668%
RMSE(Training)	0.625	1.611	1.804	0.624	0.624	3.687
RMSE(Testing)	3.201	3.388	4.004	3.150	4.443	4.184
RMSE(Validation)	2.920	3.106	3.413	3.046	4.615	4.510

Figure 03 presents line graphs that shows the scattering of predicted and actual values for all models. It is clear in the graphs that, the points for actual and predicted values are almost together for models

Such as XGBoosting, RF, LightGBM, GBM compared to the rest two models. The difference between these two points clearly noticeable from this graph, which leads the prediction less reliable and makes those statistical values larger.



(a)

(b)

(c)

(d)

(e)

(f)

Figure 3 Actual and Predicted values' distribution for (a) XGBoosting, (b) RF, (c) LightGBM, (d) GBM, (e) DT, (f) AdaBoosting

The SHAP value analysis highlights key factors influencing model performance, providing insights into parameter effects on predictions. "Days" significantly enhance predictions with high SHAP values (~25), emphasizing curing length's critical role. "Cement_(kg/m3)" exhibits a dual effect, revealing a non-linear relationship where interactions determine its positive or negative impact. The "w/c" (water-to-cement) ratio shows sensitivity, with lower values improving predictions, while higher values detract. Aggregates like "sand_(kg/m3)" and "aggregate_(kg/m3)" contribute minimally, remaining close to zero. Additives, notably "sp" (superplasticizer), show periodic impacts. PET fiber characteristics, such as "length_of_pet(mm)" and "pet fiber__(%volume)," demonstrate mixed trends, with fiber length having a more substantial influence. These findings underscore curing time and cement optimization's importance while identifying secondary contributions from additives and fibers.

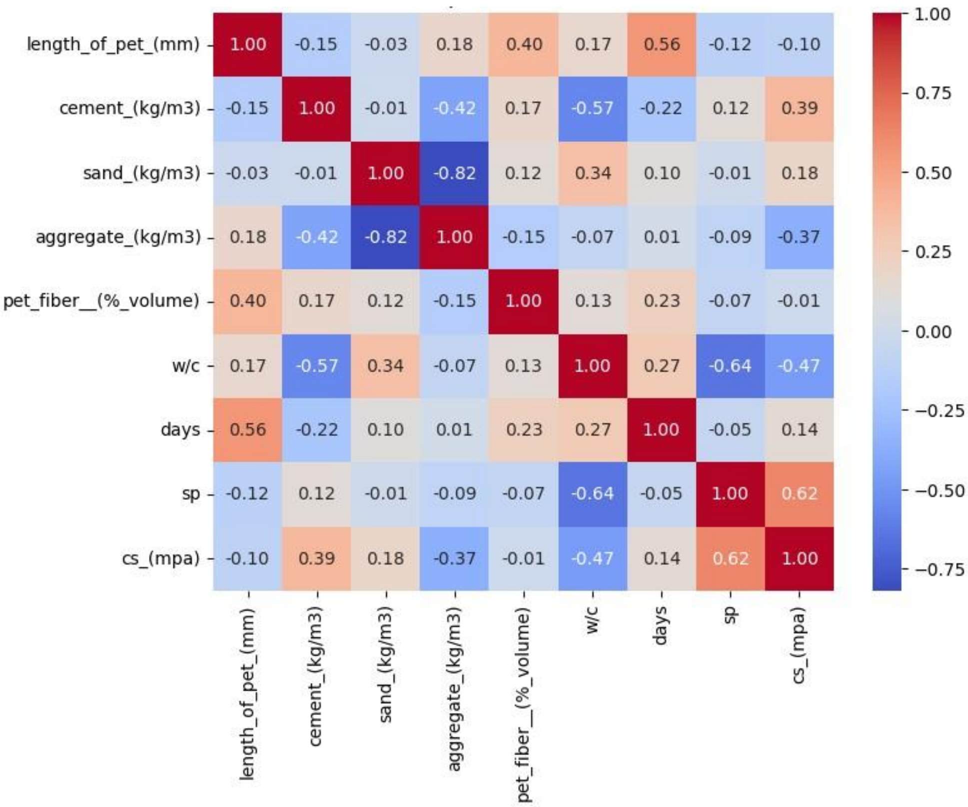


Figure 04 Heatmap of correlation matrix

The correlations between different concrete parameters, such as PET fiber qualities, cement, sand, and aggregate concentrations, and concrete compressive strength (cs_mpa), are shown in figure 04 correlation matrix heatmap. Though the water-to-cement ratio (w/c) and cs_mpa show an opposite relationship (-0.47), the matrix shows a significant positive relationship (0.62) between compressive strength and superplasticizer (sp) content. The opposing volumetric relationship is highlighted by the negative correlation (-0.82) between aggregate content and sand content that shows in figure 4. Also (PET) fiber length may have an impact on long-term strength, indicated by the moderate relationship (0.56) between it and curing days. Cement content's importance in strength development is reinforced by its positive correlation with cs_mpa (0.39).

CONCLUSIONS

The study found that machine learning (ML) models like XGBoost, Random Forest, and Gradient Boosting Machine (GBM) can accurately predict the compressive strength (CS) of concrete reinforced using waste polyethylene terephthalate (PET) fibers. The models' ability to represent deep correlations between input variables and CS parameters is shown by their excellent prediction accuracy, as seen by their R^2 values of 0.996, 0.932, and 0.926. Cement content and curing time were found to be significant predictors of CS by the SHAP value analysis. Moreover, draws attention to the need to optimize mix design parameters is the effect of PET fiber qualities like length and volume. The study also highlights the advantages of using thrown-away PET fibers in concrete manufacturing for the environment. Using recycled materials contributes to sustainable building practices by improving concrete's mechanical qualities and addressing the urgent problem of disposing of PET waste.

REFERENCES

Ahmad, A., Ahmad, W., Aslam, F., & Joyklad, P. (2022). Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques. *Case Studies in Construction Materials, 16*, e00840.

Ahmad, W., Ahmad, A., Ostrowski, K. A., Aslam, F., Joyklad, P., & Zajdel, P. (2021). Application of advanced machine learning approaches to predict the compressive strength of concrete containing supplementary cementitious materials. *Materials, 14*(19), 5762.

Ahmed, N. (2023). Utilizing plastic waste in the building and construction industry: A pathway towards the circular economy. *Construction and Building Materials, 383*, 131311.

Ali, B. H. S. H., Faraj, R. H., Saeed, M. A. H., Ahmed, H. U., & Ahmed, F. W. (2024). Innovative machine learning approaches to predict the compressive strength of recycled plastic aggregate selfcompacting concrete incorporating different waste ashes. *Multiscale and Multidisciplinary Modeling, Experiments and Design*, 1-20.

Amalraj, E. F. P., & Ilangovan, P. (2023). Experimental behavior of high-strength concrete reinforced with aramid fiber and polyurethane resin. *Buildings, 13*(7), 1713.

Amin, N.-u. (2011). Use of bagasse ash in concrete and its impact on the strength and chloride resistivity. *Journal of materials in civil engineering, 23*(5), 717-720.

Aslam, F., & Shahab, M. Z. (2024). Supplementary cementitious materials in blended cement concrete: Advancements in predicting compressive strength through machine learning. *Materials Today Communications, 38*, 107725.

Chopra, P., Sharma, R. K., Kumar, M., & Chopra, T. (2018). Comparison of machine learning techniques for the prediction of compressive strength of concrete. *Advances in Civil Engineering, 2018*(1), 5481705.

Duong, N. T., Tran, K. Q., Satomi, T., & Takahashi, H. (2022). Effects of agricultural by-product on mechanical properties of cemented waste soil. *Journal of Cleaner Production, 365*, 132814.

Ekanayake, I. U., Meddage, D. P. P., & Rathnayake, U. (2022). A novel approach to explain the blackbox nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). *Case Studies in Construction Materials, 16*, e01059. doi:https://doi.org/10.1016/j.cscm.2022.e01059

Foti, D. (2019). Recycled waste PET for sustainable fiber-reinforced concrete. In *Use of recycled plastics in eco-efficient concrete* (pp. 387-410): Elsevier.

Han, Q., Gui, C., Xu, J., & Lacidogna, G. (2019). A generalized method to predict the compressive strength of high-performance concrete by improved random forest algorithm. *Construction and Building Materials, 226*, 734-742.

Hashin, Z., & Wendt, F. (1970). Theory of composite materials. *Mechanics of composite materials*, 201242.

Kaloop, M. R., Kumar, D., Samui, P., Hu, J. W., & Kim, D. (2020). Compressive strength prediction of high-performance concrete using gradient tree boosting machine. *Construction and Building Materials, 264*, 120198.

Kang, M.-C., Yoo, D.-Y., & Gupta, R. (2021). Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. *Construction and Building Materials, 266*, 121117.

Khan, M., Lao, J., & Dai, J.-G. (2022). Comparative study of advanced computational techniques for estimating the compressive strength of UHPC. *Journal of Asian Concrete Federation, 8*(1), 5168.

Lai, W. L., Sharma, S., Roy, S., Maji, P. K., Sharma, B., Ramakrishna, S., & Goh, K. L. (2022). Roadmap to sustainable plastic waste management: a focused study on recycling PET for triboelectric nanogenerator production in Singapore and India. *Environmental Science and Pollution Research, 29*(34), 51234-51268. doi:10.1007/s11356-022-20854-2

Lamba, P., Kaur, D. P., Raj, S., & Sorout, J. (2022). Recycling/reuse of plastic waste as construction material for sustainable development: a review. *Environmental Science and Pollution Research, 29*(57), 86156-86179.

Liu, H., & Lu, Z. (2012). *Reinforced Mechanism and Cracking Strength Study of Fiber Concrete Based on Spacing Theory.* Paper presented at the Proceedings of the 2012 Second International Conference on Electric Technology and Civil Engineering.

Marsh, K., & Bugusu, B. (2007). Food packaging—roles, materials, and environmental issues. *Journal of food science, 72*(3), R39-R55.

Mwangi, J. P. M. (2001). *Flexural behavior of sisal fiber reinforced concrete beams*: University of California, Davis.

Nguyen, H., Vu, T., Vo, T. P., & Thai, H.-T. (2021). Efficient machine learning models for prediction of concrete strengths. *Construction and Building Materials, 266*, 120950.

Nguyen, N.-H., Abellán-García, J., Lee, S., Garcia-Castano, E., & Vo, T. P. (2022). Efficient estimating compressive strength of ultra-high performance concrete using XGBoost model. *Journal of Building Engineering, 52*, 104302.

Pakzad, S. S., Roshan, N., & Ghalehnovi, M. (2023). Comparison of various machine learning algorithms used for compressive strength prediction of steel fiber-reinforced concrete. *Scientific Reports, 13*(1), 3646.

Peng, Y., & Unluer, C. (2022). Analyzing the mechanical performance of fly ash-based geopolymer concrete with different machine learning techniques. *Construction and Building Materials, 316*, 125785.

Qin, Y., Li, M., Li, Y., Ma, W., Xu, Z., Chai, J., & Zhou, H. (2021). Effects of nylon fiber and nylon fiber fabric on the permeability of cracked concrete. *Construction and Building Materials, 274*, 121786.

Rahman, J., Ahmed, K. S., Khan, N. I., Islam, K., & Mangalathu, S. (2021). Data-driven shear strength prediction of steel fiber reinforced concrete beams using machine learning approach. *Engineering Structures, 233*, 111743.

Rahmati, M., & Toufigh, V. (2022). Evaluation of geopolymer concrete at high temperatures: An experimental study using machine learning. *Journal of Cleaner Production, 372*, 133608.

Rathakrishnan, V., Bt. Beddu, S., & Ahmed, A. N. (2022). Predicting compressive strength of highperformance concrete with high volume ground granulated blast-furnace slag replacement using boosting machine learning algorithms. *Scientific Reports, 12*(1), 9539.

Sadeghian, V., Tanyous, M., & Mirshekar, S. (2020). *Modelling FRP-strengthened beam-column joints in performance assessment of RC frames.* Paper presented at the Proceedings of the 6th International Conference on Construction Material (ConMat 20), Fukuoka, Japan.

Sau, D., Shiuly, A., & Hazra, T. (2023). Study on green concrete replacing natural fine and coarse aggregate by plastic waste–An experimental and machine learning approach. *Materials Today: Proceedings*.

Tran, V. Q. (2023). Data-driven approach for investigating and predicting of compressive strength of fly ash–slag geopolymer concrete. *Structural Concrete, 24*(6), 7419-7444.

Xu, Y., Ahmad, W., Ahmad, A., Ostrowski, K. A., Dudek, M., Aslam, F., & Joyklad, P. (2021). Computation of high-performance concrete compressive strength using standalone and ensembled machine learning techniques. *Materials, 14*(22), 7034.

Yang, S., Chen, H., Feng, Z., Qin, Y., Zhang, J., Cao, Y., & Liu, Y. (2023). Intelligent multiobjective optimization for high-performance concrete mix proportion design: A hybrid machine learning approach. *Engineering Applications of Artificial Intelligence, 126*, 106868.

Zhang, J., Wang, L., & Kannan, K. (2021). Quantitative analysis of polyethylene terephthalate and polycarbonate microplastics in sediment collected from South Korea, Japan and the USA. *Chemosphere, 279*, 130551.

Zhang, W., Wang, Y., Nan, X., Sun, S., Ma, Y., & Wu, Y. (2024). An Experimental Study on the Performance of Materials for Repairing Cracks in Tunnel Linings under Erosive Environments. *Buildings, 14*(8), 2427.