Fusion Journal of Engineering and Sciences

Article

Predicting the Effect of Fly Ash Dosage on the Compressive Strength of Various Concretes using Machine Learning (ML) Techniques

Muhammad Sufian Naeem [1], Malik Sarmad Riaz [1*] and Zain Ahmad [2]

[1] Civil Engineering Department, National University of Technology (NUTECH), Islamabad, Pakistan. muhammadsufian.naeem1@gmail.com;sarmadriaz@nutech.edu.pk

[2] Computer Science Department, Air University, Kamra, Pakistan. 235085@aack.au.edu.pk

Abstract

Fly Ash (FA) is a widely used pozzolanic supplement in the production of sustainable concrete in the construction industry, presenting evident advantages, such as reducing CO_2 emissions, cost-efficient mixes, improved strength. In this research, to develop machine learning ML-based models to predict the compressive strength of various fly ash-based concretes. Three ensemble ML techniques include Extra Trees regression (ETR), XGBoost (XGB), and Random Forest (RF) were applied to a credible dataset of 545 points collected from existing literature. Ten input parameters include cement, fine aggregate, coarse aggregate, type of concrete, admixtures, Fly Ash, water-to-binder, temperature, curing days, and relative humidity. The performance of the models was assessed using various statistical parameters including Mean Square Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). Extra Trees regression model has achieved high predictive accuracy with MSE is 12.21 and R^2 is 0.99, outperforming both RF and XGB for predicting the compressive strength in various FA based concretes. This study emphasizes the potential of ML approaches predictive concrete compressive strength capacity, enabling optimized mix-designs that contribute to structural integrity and environmental sustainability.

Keywords: Various FA Concretes; Fly Ash; Compressive Strength; Prediction Modeling; Machine Learning Techniques

1. Introduction

Cement is essential to global infrastructure development. In 2020, worldwide cement production reached approximately 4.4 billion tons. Each ton of cement produced emits 0.9 tons of carbon dioxide (CO2). Consequently, the Portland industry is responsible for approximately 8% of global human induced CO2 emissions [1]. Although global concerns toward renewable energy sources, coal is still a major source of electricity generation in certain countries. The reliance leads to the production of significant amount of fly ash, a byproduct of coal combustion. The fly ash exhibits pozzolanic properties, which contribute to enhancing the strength of concretes [2]. Pozzolanic materials, which are rich in SiO2 and often contain Al2O3, are reactive enough to form calcium silicate hydrate (CSH) when mixed with water and CaO at ambient temperature thereby functioning as hydraulic cements [3]. The use of different reactive pozzolanas as replacement cementitious materials is rapidly increasing in the pursuit of more durable and high performance concrete [4]. Certain fly ashes with elevated calcium content can exhibit cementitious properties by reacting with water to form hydrates even without the presence of calcium hydroxide. These pozzolanic reactions benefit concrete by increasing the amount of cementitious binder phases such as calcium-silicate-hydrate (C-S-H) and, to a lesser degree, calcium-aluminate hydrates. This contributes to enhanced long-term strength and decreased permeability of the concrete, both of which improve its overall durability [5].

Fly ash exhibits a variable composition based on the coal variety used and the design of boilers, resulting in different classifications such as siliceous, calcareous, or silica calcareous fly ash [6]. Experiment using 0 – 50 % fly ash as a partial replacement of cement at 7,14,28 and 56 days after curing yielded optimal result for compressive strength [7]. However, using fly ash in concrete applications may result in slower early compressive strength development [8].

Conventional formulas and through experimental testing are commonly used to forecast for compressive strength but can be time-consuming and expensive. Machine learning has gained popularity for its capacity to examine large datasets to make accurate predictions. This study focused on developing the machine learning model to predict compressive strength using fly ash-based concretes as a partial replacement of cement in various types of concretes. Using fly ash to replace cement in concrete can enhance durability, permeability, and decrease carbon emissions. To accurately quantify the impact of variations in fly ash characteristics on compressive strength, complex predictive models are required.

In the pursuit of sustainable practices, the use of alternative cementing materials has gained significant momentum. These materials, when used in partial replacement of cement, not only contribute to waste reduction but also enhance the mechanical properties of concrete. The material benefits are shown in Fig 1. Among various SCMs, fly ash has emerged as a widely used and effective material due to its pozzolanic nature [9].

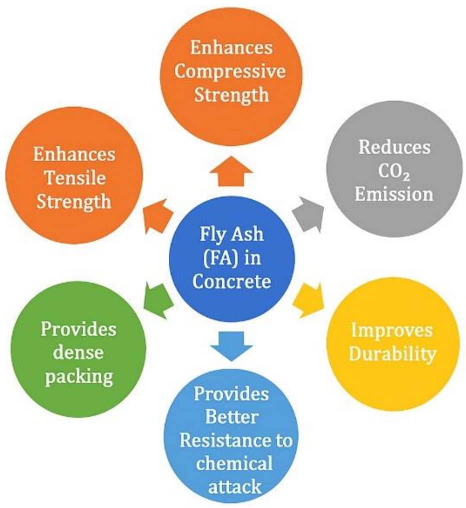

Figure 1. key benefits of incorporating Fly Ash (FA) in concrete, highlighting its role in enhancing strength, durability and sustainability.

The application of machine learning (ML) to civil engineering has transformed the prediction of concrete qualities. ETR, XGB, and RF algorithms can describe complex interactions between input factors and compressive strength. ML models can manage nonlinear interactions among variables, resulting in more accurate and efficient predictions than standard empirical methods. A study compared machine learning model techniques (ETR, XGB, and RF) to predict the compressive strength of concrete with fly ash [9]. The models were refined using a grid search technique. The revised ETR model's superior predictive performance suggests its potential for pre-assessing concrete based on mixed design proportions and other features. The choice of ML algorithm has a considerable impact on predictions accuracy. Ensemble approaches effectively manage variability in concrete mix designs.

Research is using ML algorithms, genetic algorithms, and particle swarm optimization to optimize hyperparameters and increase model reliability. Using models trained on related data sets to solve new problems has helped alleviate data shortage concerns. This approach creates robust predictions models with minimum experimental data, making it applicable to a wide range of concrete mix design options. Table 1 summarizes some of the studies on machine learning applications for industrial waste.

Table 1. Prediction of concrete properties by using different industrial raw materials

Sr No.	Algorithm Name	Notation	Dataset	Prediction Properties	Material Used	Year	References
1	Random Forest	RF	131	Compressive Strength	GGBS	2019	[10]
2	Artificial neural network	ANN	69	Compressive Strength	FA, GGBS, RHA	2016	[11]
3	Adaptive neuro fuzzy inference system	ANFIZ	7	Compressive Strength	POFA	2020	[12]
4	Multivariate	MV	21	Compressive Strength	Crumb rubber with SF	2020	[13]
5	Gene Expression Programming	GEP	351	Compressive Strength	GGBS	2020	[14]
6	Gene Expression Programming	GEP	351	Compressive Strength	NZ (Natural Zeolite)	2019	[15]
7	Artificial neural network	ANN	169	Compressive Strength	SF, GGBS	2016	[16]
8	Response Surface Method, Gene expression programming	RSM, GEP	108	Compressive Strength	Steel Fibers	2020	[17]
9	Random Forest	RF	321	Compressive Strength	GGBS, Alkali	2023	[18]
10	Artificial neuron network	ANN	205	Compressive Strength	GGBFS, SF, RHA	2019	[19]

As presented in Table 1, previous research has demonstrated the effectiveness of machine learning approaches such as Random Forest (RF) [20], Artificial Neural Networks (ANN) [21], and Boosting techniques [22] were adopted to predict the compressive strength of specific concrete incorporating industrial by-product. However, most of these studies were limited by small datasets, focused on specific types of concrete such as high-strength concrete [23], normal concrete [24], eco-friendly concrete [25], and applied algorithms with limited capability and lacking comprehensive models. In this research, addresses these gaps by utilizing a large and more diverse dataset (n =545) incorporating various concretes samples. By implementing advanced ensemble ML models ETR, XGB, and RF are equipped to model non-linear interactions, minimize statistics measurement errors, and enhance the reliability of compressive strength prediction of various FA-based concretes, non-linear interactions minimize statistics measurement errors, and enhance the reliability of compressive strength prediction of various FA-based concretes.

2. Materials and Methods

The aim of this research is to create machine learning models that can accurately and reliably forecast the compressive strength of different concretes that include fly ash in place of some of the cement. Data collection, data preprocessing, feature selection, model creation, training, and other stages comprise the approach used in this study. Testing, assessment, and ultimate choice of the best model. To assess compressive strength values using a variety of input variables, a dataset was assembled utilizing experimental data gathered from numerous published publications.

To guarantee consistency and applicability for machine learning models, the gathered dataset underwent processing steps in preprocessing. The methodology used a methodical, multi-phase approach to guarantee accuracy, reproducibility, and consistency. These stages consist of training the model, preparing the database, and assessing performance with statistical measures.

In the initial stage, the database was assembled from published experimental research involving concrete mixtures and different proportions of fly ash employed as a partial cement substitute. Critical input characteristics such as cement and fly ash concentrations, water-to-cement ratio, proportions of fine and coarse aggregates, curing age, and measured compressive strength were included in every dataset entry.

Table 2. Descriptive statistical of input and output parameter.

Parameters	Cement (kg/m^3)	Fine Aggregate (kg/m^3)	Coarse Aggregate (kg/m^3)	ADMIXTURE (%)	Fly Ash (%)	W/B	Temperature (°C)	Curing (Days)	Relative Humidity (%)	Compressive Strength (MPa)
	Input	Input	Input	Input	Input	Input	Input	Input	Input	Output
Mean	406.38	672.59246	972.58092	1.477449541	10.753	0.44	28.30825688	31.35	81.601835	44.34590826
Median	374	695	1000	1.2	10	0.4	25	28	90	36.78
Mode	360	600	1000	1	10	0.45	27	28	95	45
Standard deviation	184.948	219.95581	269.79084	1.174952561	11.135	1.15	13.02643049	32.225	16.582603	30.62117451
Variation	34205.8	48380.558	72787.095	1.38051352	124	1.31	169.6878913	1038.5	274.98272	937.6563283
Kurtosis	4.44028	1.3852421	2.4330448	3.867383277	27.427	535	17.95181319	6.1877	-1.350765	3.482628656
Skewness	1.66659	-0.998818	-1.081946	2.084807159	4.1546	23	3.905943721	2.27	-0.61625	1.85599127
Min	0	10	8.2	0.27	0	0.17	20	1	50	2.5
Max	1120	1030	1525	6	100	27	120	180	100	160
Range	1120	1020	1516.8	5.73	100	26.8	100	179	50	157.5
Sum	221477	366562.89	530056.6	805.21	5860.5	241	15428	17086	44473	24168.52
Count	545	545	545	545	545	545	545	545	545	545

The model was trained 80% of the dataset, while 20% of the dataset was used to test the model's performance. Among the algorithms tested, Extra Trees regression (ETR) was selected for its strong predictive capacity. XG Boost is resilient to multi-collinearity, while Random Forest (RF) is good at training and lowering variance. The trained models were evaluated for the test dataset, and predictive performance evaluation using statistical measures such as mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2) were the main metrics employed are shown in Figure 2.

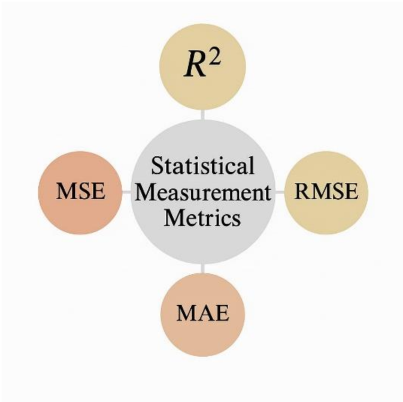

Figure 2. Statistical measurements metrics for the model evaluation.

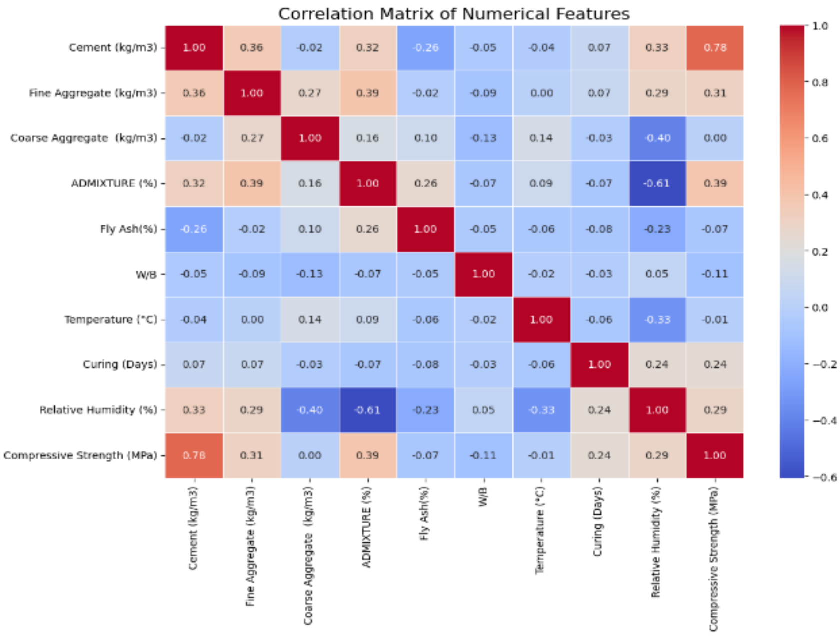

Figure 3. Correlation Heatmap matrix.

2.1 Overview of AI in Civil Engineering

Artificial Intelligence has become more popular and has been used in many different scientific fields. The Extra Trees Regression is one of the most potent machine learning algorithms in data science, and it is widely utilized in the building sector. Numerous technical issues in earth science, environmental science, geotechnical engineering, and civil engineering have been successfully resolved by the ET model [26].

2.2. Machine Learning Algorithms

Leveraging machine learning makes it possible for machines to process data more efficiently. Sometimes extracting valuable insights from data requires more than just looking at it. In these situations, machine learning methods are used to find hidden patterns and comprehend the data more thoroughly [27]. In civil engineering, machine learning has emerged as a potent tool, particularly for mix design optimization and concrete behavior prediction. In this study, machine learning techniques including Extra Trees regression (ETR)**,** *X*GBoost (XGB), and Random Forest (RF), were used to predict compressive strength. An ensemble-based technique called Extra Trees is well known for lowering variance and enhancing model stability. By combining the output of several decision trees, random forest, a popular ensemble technique, offers reliable overfitting. Because it can manage correlated input features with little effect on model performance, XGBoost is resilient to multicollinearity. The effectiveness of these algorithms in determining the mechanical behavior of concrete and improving computational speed has been highlighted by recent studies. Using these methods, the current study creates a predictive model for sustainable concrete, evaluates its performance indicators, and determines which model is the most accurate and computationally efficient for real-world applications.

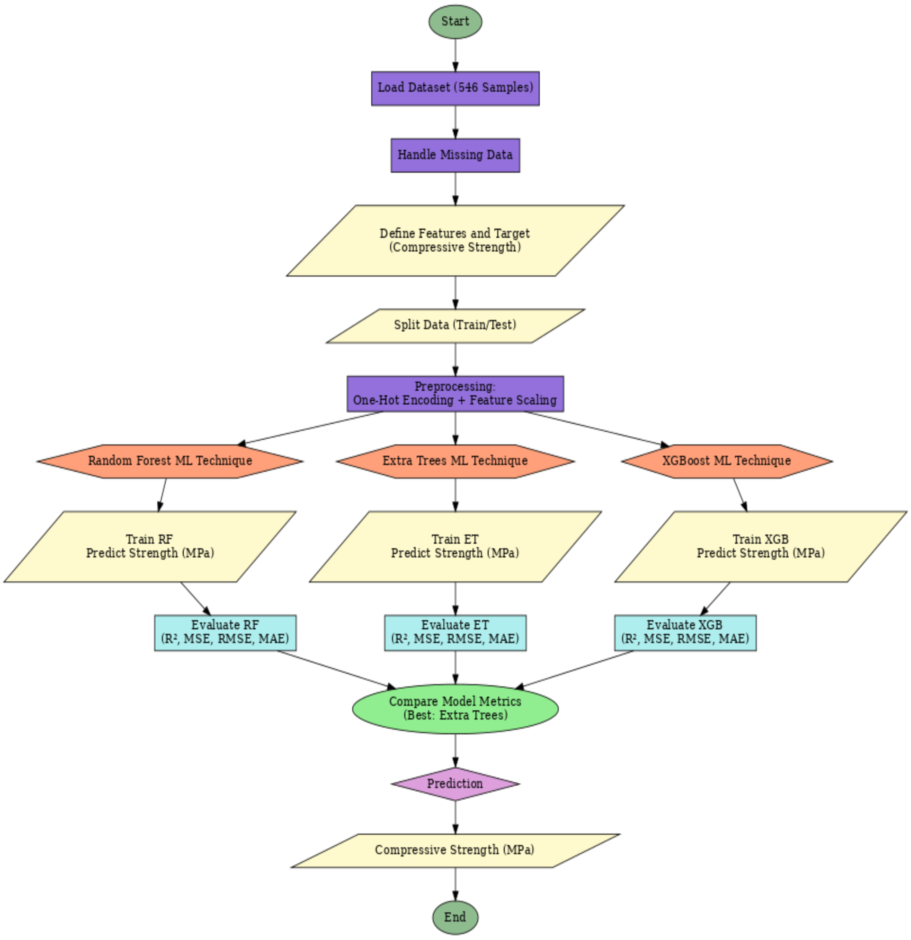

Figure 4. Flowchart of Machine Learning Algorithms.

2.2.1. Extra Tree Regression (ETR)

The tree-based ensemble approaches like Random Forest and Bootstrap Aggregation (Bagging) are closely connected to the Extra Trees (ETR) Regressor, an ensemble method constructed from numerous decision trees. Using the training data, the Extra Trees algorithm builds a numerous number of unpruned decision trees. It averages the outputs of each individual tree to produce final predictions in regression problems. Using the training data, the Extra Trees algorithms create an unpruned decisions tree. The average of the outputs from each tree in the ensemble is used to provide the final predictions for regression problems [28]. There are two main differences between Random Forests (RFs) and Extra Trees Regressors (ETRs). First, instead of picking the best split, ETRs select split points at random from a range of potential thresholds. Second, they reduce bias by growing each tree using entire training dataset, rather than relying on bootstrap sampling [27]. The splitting process in the Extra Trees Regressor is controlled by two parameters: k and n_{min}. Here, k represents the number of features randomly selected at each node, while n_{min} denotes the minimum number of samples required to split a node. These parameters influence both the effectiveness of attribute selection and the level of noise in the average output. Using these parameters improves the models precision and reduces overfitting [29].

2.2.2. Random Forest (RF)

The random forest algorithm is a machine learning method commonly applied to both classification and regression problems. It is a clever application of bootstrap bagging (or simply bagging). This method involves two randomization steps: first it uses the bootstrap technique to create random samples with replacement, and second, it selects random subsets of features from the original dataset in random order. These two diverse datasets are highly varied, which helps reduce variance. Using these datasets, classification and regression trees (CARTs) are built. A forest is then formed by combining multiple CARTs, with different data randomly assigned to these trees in terms of both rows and columns.

It consists of multiple decision trees (DTs). For classification purposes, the final prediction is determined by the significant decision across all trees, while for regression tasks, the predictions the average of the outputs from all the trees. RF is highly versatile and can be used for large-scale problems [30], [31].

2.2.3. XGBoost

XGBoost is a robust and efficient machine learning algorithm known for builds a strong predictive models by combining multiple weak learners, typically decision trees, in a sequential manners [32]. Each tree is trained to correct the errors made by the previous ones, improving the model's overall performance. Large dataset handling, regularization strategies to avoid overlifting, and parallel processing to speed up training. Additionally, it provides early halting to prevent needless computation and efficiently handles missing data. Because of these characteristics, XGB is a well-known method for both regression and classification problems in a wide range of challenges.

2.3. Model Evaluation Criterion

Several statistical errors, including Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of Determination (R^2), can be used to evaluate how well a created model performs on training and trial datasets. R^2 is generally thought to be the best of these for assessing model correctness. A variety of modelling techniques have been used to forecast the mechanical properties of concrete due to the quick development of artificial intelligence. These error measures, which provide distinct insights into the type and magnitude of prediction mistakes, are calculated as part of the statistical analysis utilized in this work to assess model performance.

The average absolute difference between the expected and actual values is determined by MAE, which calculates the average absolute difference between the predicted and actual values. It is less sensitive to outliers than MSE. MAE provides another perspective on model accuracy, especially useful for models dealing with real-world noisy concrete data. MAE value can be calculated with equation 1:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |x_i - y_i| \quad (1)$$

The errors on average of the squares between predicted values f_i and actual (true) values p_i measured using MSE. A lower MSE indicates that the model's predictions are closer to the actual values. MSE helps evaluate the prediction accuracy of the XGB, Random Forest, and Extra Trees models for compressive strength. The MSE can be calculated using equation 2:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - y)^2 \quad (2)$$

RMSE penalizes underestimates more than overestimates and is helpful when actual values span multiple orders of magnitude. Useful in civil engineering datasets where compressive strength can vary greatly. RMSE supports evaluating model robustness when dealing with highly skewed or non-linear concrete strength data. The evaluation of RMSLE using equation 3:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=0}^{n} (xi - yi)^2}{n}} \quad (3)$$

R^2 explains how well the model's predictions approximate the actual data. The value of R^2 (closer to 1) indicates better performance. R^2 quantifies how well fly ash data inputs explain variability in concrete strength predictions. The R^2 value is calculated by equation 4:

$$R^2 = 1-\frac{\sum_{i=1}^{n}(xi-yi)^2}{\sum_{i=1}^{n}(xi-\bar{x})^2} \quad (4)$$

3. Results and Discussion

In this research study, a machine learning (ML) model was developed to predict the compressive strength of concretes incorporating fly ash as a partial replacement for cement. The algorithms, including ETR, XGB, and RF, were tested for accuracy, with extra trees regression model demonstrating the highest prediction performance based on R^2 score, MAE, RMSE and MSE. The dataset used was compiled from existing experimental studies and pre-processed for optimal training. The model effectively captured nonlinear relationship between input variables such as mix Cement (Kg/m^3), Fine Aggregates (Kg/m^3), Coarse Aggregates (Kg/m^3), Concrete Type, admixture (%), Fly ash (%), W/B, Temperature (°C), curing (days), relative humidity (RH)

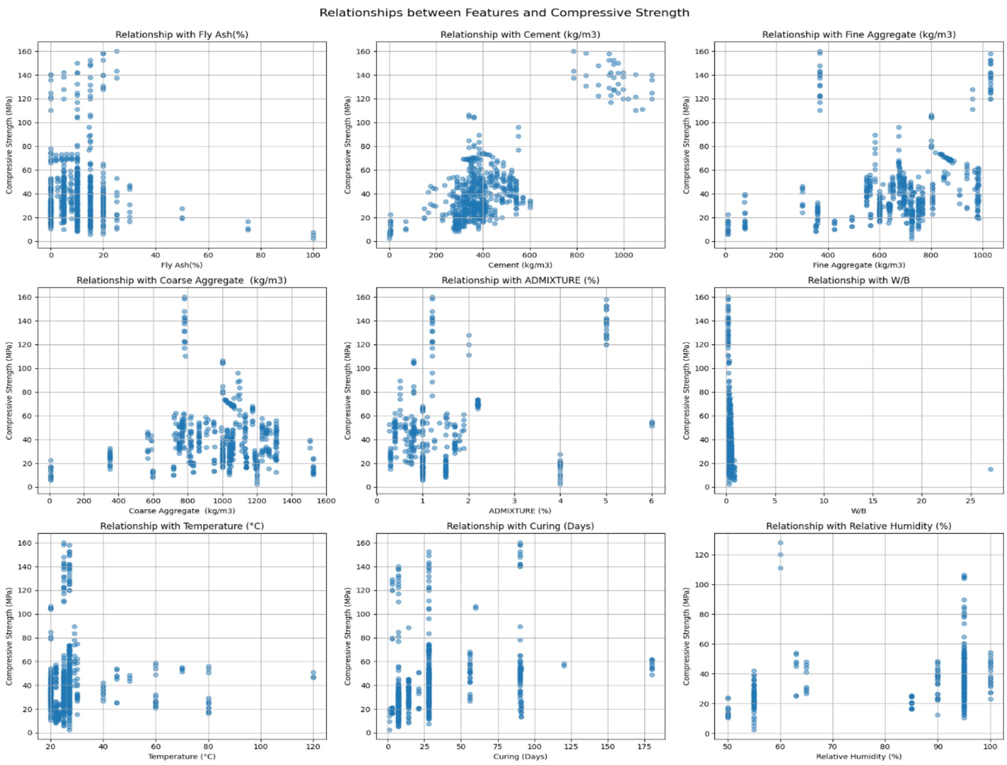

Figure 5. Non-linear relationships between input features and compressive strength (MPa).

3.1. Comparison between applied machine learning models

The model's performance was evaluated based on MAE, RMLSE, MSE, and R^2 values. The findings reveal that the ETR model had the highest accuracy, with an MAE of 2.85, the lowest RMSE of 3.99, the lowest MSE of 12.21, and the highest R2 of 0.99, this demonstrates strong predictive power. Other models, such as XGB and RF also have high R^2 values 0.98 and 0.97, indicating occasional prediction potential. Ensemble based models ET have higher predictions accuracy for this dataset. These studies show that ensemble learning can accurately forecast the strength of various concretes. The Extra Trees model outperformed other models, demonstrating the effectiveness of ensemble-based approaches in dealing with nonlinear relationships and complicated feature interactions.

Figure 6. Comparison of Model Performance based on Coefficient of Correlation (R^2)

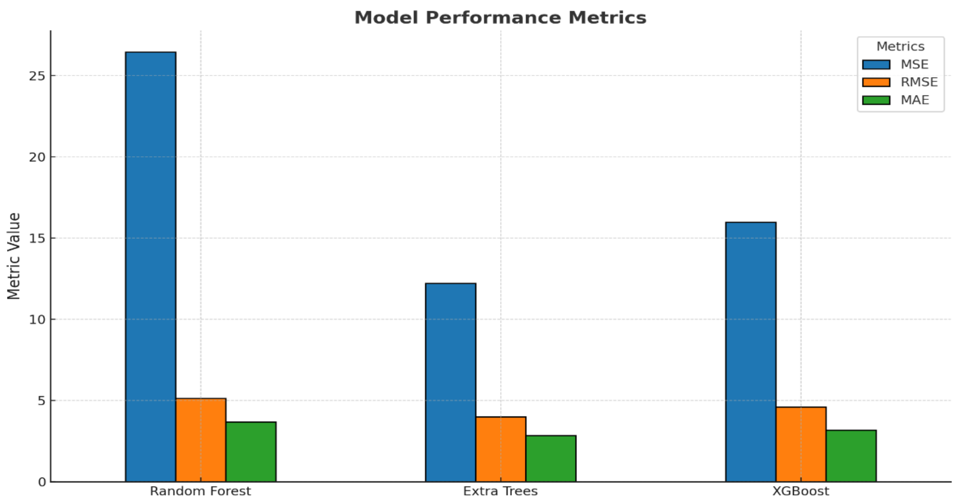

Figure 7. Comparison of Model Performance metrics errors based on MSE, RMSE, MAE.

The comparison highlights how each algorithm performs in terms of predictive accuracy, likely based on metrices such as R^2, MSE, RMSLE, and MAE. Among these, ensemble models ETR shows superior performance, indicating their effectiveness in handling nonlinear and complex relationships in concretes strength prediction.

Table 3. Performance Comparison of Extra Tree, Random Forest, and XGBoost models for Predicting Compressive Strength Based on MAE, RMLSE, MSE, and R^2 Value Metrics

Models	MSE	RMSE	MAE	R^2 Value
Extra Trees Regression	12.21	3.99	2.85	0.99
XGBoost	15.97	4.6	3.19	0.982
Random Forest	26.45	5.14	3.69	0.97

The extra trees model demonstrates the strongest performance among the three algorithms based on key evaluation metrics. ETR has the Lowest Mean Absolute Error (MAE) of 2.85, indicating the smallest average prediction error. Additionally, it records the lowest Root Mean Squared Error (RMSE) of 3.99, which suggests better handling of exponential or skewed data distributions. With a Mean Squared error (MSE) of 12.21, it reflects fewer large prediction errors compared to the other models. Moreover, its R^2 value of 0.99 is the highest, signifying that it explains the most variance in the target variable. These results highlight the ETR model as the most accurate and reliable choice among the three.

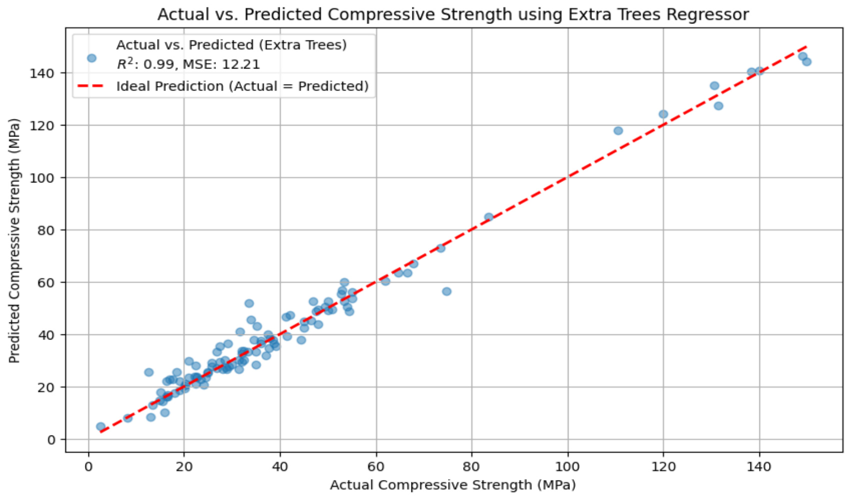

Figure 8. Comparison Actual vs Predicted values graph using Extra Trees Regression technique.

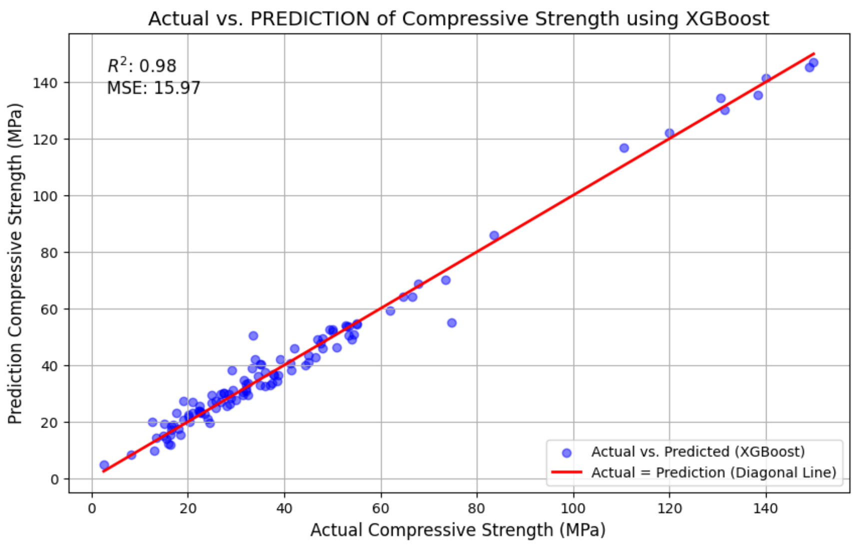

Figure 9. Comparison of Actual vs Predicted values graph using XGBoost technique

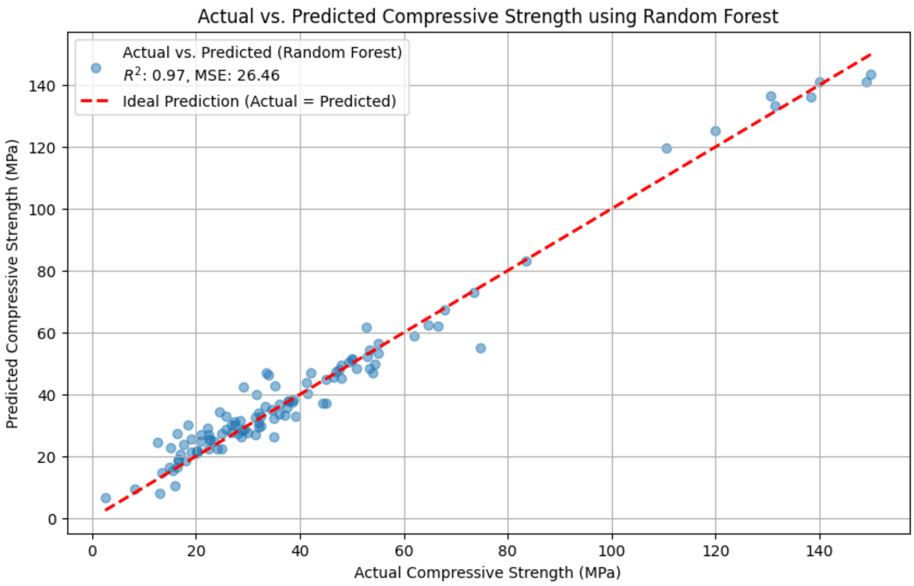

Figure 10. Random Forest Technique Actual vs Predicted values graph

The prediction graphs compare the prediction accuracy of different machine learning algorithms used in the literature to forecast the compressive strength of various concretes incorporating supplementary material such as Fly Ash. Algorithms like ETR, XGB and RF are evaluated. The actual versus prediction value graphs show that the ETR provides more accurate predictions, with data points closely aligned along the reference line, indicating that it effectively captures the fly ash dosage on strength of different concretes. XGBoost, while still showing a strong correlation, demonstrates the largest spread and some deviation from ideal predictions. The Random Forest model also performs well, though there is slightly more dispersion in the data points as compared to Extra Trees. These graphs highlight that ETR model is most effective in predicting compressive strength in various types of concrete. The investigation confirms that ML models, especially tree-based ensembles the influence of fly ash dosage, offering valuable insights for optimizing concrete mix designs in different types of various concretes.

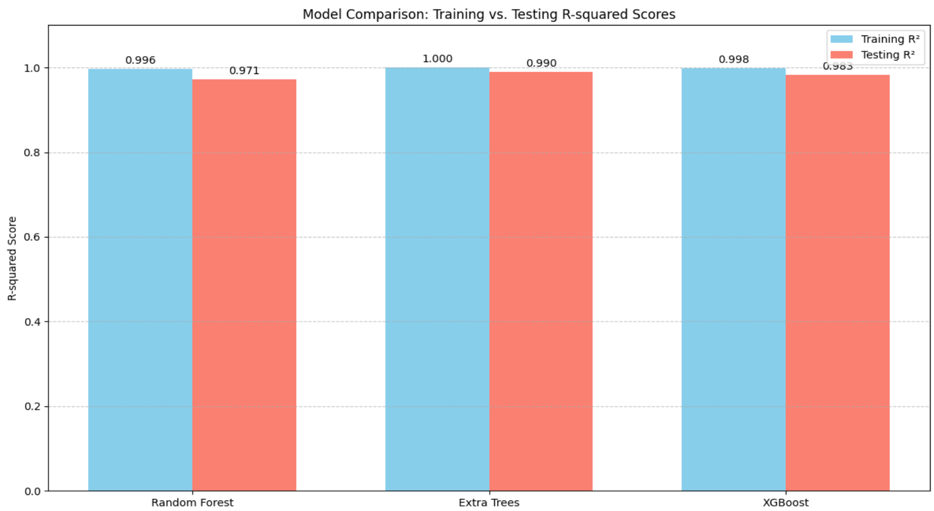

Figure 11. Comparison of ETR, XGB, and RF models on training vs testing with correlation coefficient (R2)

The performance of the ensemble ML models on training and testing evaluation are presented in Fig 6. The Extra Trees Regression (ETR) approach achieved R^2 is 1 on the training set and 0.99 on the testing set, showing potential predictive accuracy and minimal overfitting. Random Forest

(RF) and XGBoost also indicate strong capacity with less difference between training and testing R^2 evaluation. This analysis affirms the predictability and robustness of the models in forecasting the compressive strength of fly-ash based concretes. The consistency of these algorithms demonstrates the effectiveness of ensemble modeling in handling complex non-linear relationships.

3.2. Sensitivity Analysis

Ten parameters, including concrete type, cement (Kg/m^3), fine aggregate (Kg/m^3), coarse aggregate (Kg/m^3), admixtures (%), Fly Ash (%), water-to-binder (W/B), temperature (°C), curing (Days), and relative humidity (%) were used as an input features. shown In Figure 3, show the impact of each input parameter on the model development process. The Concrete type and mix deign have been found to have a greater influence on compressive strength compared to other input feature.

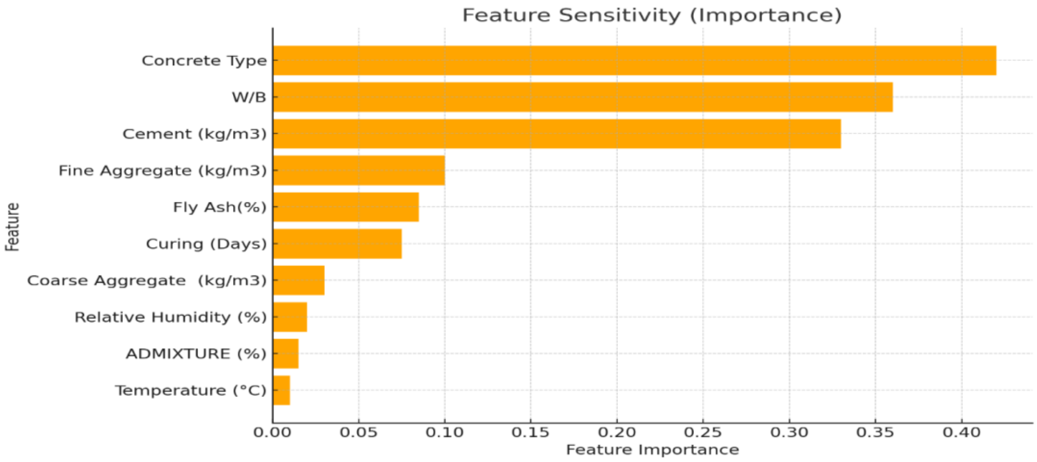

Figure 12. Contribution of input features to predict compressive strength

The above graph presents the sensitivity of different features used in the machine learning models to predict the compressive strength of fly ash-based concretes. The analysis identified, concrete type and mix design have been found to have a greater influence with higher sensitivity score on compressive strength model's predictions as compared to other input features. The more critical parameter is in determining compressive strength of different concretes. This insight helps optimize mix designs by focusing on the most influential components.

4. Conclusions

The use of machine learning to assess compressive strength of various concrete types using FA highlights the potential of data-driven techniques in civil engineering. This technique encourages sustainable construction practices by encouraging the use of the use of industrial by products like fly ash, while simultaneously increasing mix design efficiency. The Extra Trees model performs well, allowing engineers to make sustainable decisions. The ML model accurately predicts compressive strength of various concrete kinds using input values specified in the sensitivity analysis section. Extra Trees outperformed other models, indicating their suitability for capturing non-linear relationships in various concretes mixed data. This approach reduces the need for expensive and time-consuming lab experiments in early design phases. Fly Sash as a partial cement replacement supports eco-friendly construction while maintaining required strength characteristics.

Ensemble machine learning (ML) techniques have proven to be effective in accurately predicting the compressive strength of various types of concretes. To enhance the precision of ML algorithms results by applying the multiple predictions on developed ML model. The integration of ML techniques in civil engineering not only enables efficient prediction of concrete properties but also helps reduce overall project costs and shortens the time required to achieve the target outcomes.

Funding: No internal or external funding.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The data presented in this article are available within the article.

Acknowledgments: The authors gratefully acknowledge the invaluable guidance and support of Eng. Mutahir Abbas, whose expert supervision, insightful suggestions, and constructive feedback greatly contributed to the depth and quality of this research. They also extend heartfelt thanks to their **parents** for their encouragement, and moral support throughout the research journey. Furthermore, sincere appreciation is expressed to faculty members for their academic guidance and inspiration, which have been instrumental in shaping the authors' knowledge and commitment to this work.

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

Abbreviations

The following abbreviations are used in this manuscript:

C-S-H	Calcium-silicate-hydrate
CO_2	Carbon Dioxide
AI	Artificial intelligence
ML	Machine learning
FA	Fly ash
RF	Random Forest
ETR	Extra trees regression
XGB	XGBoost
MSE	Mean square error
MAE	Mean absolute error
RMSE	Root mean square error
CART's	Classification and regression trees
DT	Decision tree
ANN	Artificial neural networks

References

1. K. Kaptan, S. Cunha, and J. Aguiar, "A Review: Construction and Demolition Waste as a Novel Source for CO2 Reduction in Portland Cement Production for Concrete," *Sustain.*, vol. 16, no. 2, 2024, doi: 10.3390/su16020585.
2. S. Abbas, M. A. Saleem, S. M. S. Kazmi, and M. J. Munir, "Production of sustainable clay bricks using waste fly ash: Mechanical and durability properties," *J. Build. Eng.*, vol. 14, pp. 7–14, 2017, doi: 10.1016/j.jobe.2017.09.008.
3. C. a. Strydom and J. C. Swanepoel, "Utilisation of fly ash in a geopolymeric material," *Appl. Geochemistry*, vol. 17, no. 8, pp. 1143–1148, 2002.
4. S. K. Agarwal, "Pozzolanic activity of various siliceous materials," *Cem. Concr. Res.*, vol. 36, no. 9, pp. 1735–1739, 2006, doi: 10.1016/j.cemconres.2004.06.025.
5. Thomas M., "Optimizing the Use of Fly Ash in Concrete," *Portl. Cem. Assoc.*, 2007.
6. N. Longarini, P. G. Crespi, M. Zucca, N. Giordano, and G. Silvestro, "The advantages of fly ash use in concrete structures," *Inz. Miner.*, vol. 15, no. 2, pp. 141–145, 2014.
7. D. N. Subramaniam and N. Sathiparan, "Comparative study of fly ash and rice husk ash as cement replacement in pervious concrete: mechanical characteristics and sustainability analysis," *Int. J. Pavement Eng.*, vol. 24, no. 2, 2023, doi: 10.1080/10298436.2022.2075867.

8. K. Kiattikomol, C. Jaturapitakkul, S. Songpiriyakij, and S. Chutubtim, "A study of ground coarse fly ashes with different finenesses from various sources as pozzolanic materials," *Cem. Concr. Compos.*, vol. 23, no. 4–5, pp. 335–343, 2001, doi: 10.1016/S0958-9465(01)00016-6.

9. H. F. W. Taylor, "Cement chemistry," *Cem. Chem.*, 1997, doi: 10.1680/cc.25929.

10. M. Esmailpour, K. Rahmani, and S. Piroti, "Experimental Evaluation of the Effect of Silica Fume on Compressive, Tensile Strength, Abrasion Resistance, Slump and Impact Test and Water Permability Coefficient of Concrete," *J. Appl. Eng. Sci.*, vol. 8, no. 1, pp. 27–36, 2018, doi: 10.2478/jaes-2018-0004.

11. D. Van Dao, H. B. Ly, H. L. T. Vu, T. T. Le, and B. T. Pham, "Investigation and optimization of the C-ANN structure in predicting the compressive strength of foamed concrete," *Materials (Basel).*, vol. 13, no. 5, pp. 1–17, 2020, doi: 10.3390/ma13051072.

12. T. Al-Mughanam, T. H. H. Aldhyani, B. Alsubari, and M. Al-Yaari, "Modeling of compressive strength of sustainable self-compacting concrete incorporating treated palm oil fuel ash using artificial neural network," *Sustain.*, vol. 12, no. 22, pp. 1–13, 2020, doi: 10.3390/su12229322.

13. R. Bušić, M. Benšić, I. Miličević, and K. Strukar, "Prediction models for the mechanical properties of self-compacting concrete with recycled rubber and silica fume," *Materials (Basel).*, vol. 13, no. 8, 2020, doi: 10.3390/MA13081821.

14. A. A. Shahmansouri, H. Akbarzadeh Bengar, and E. Jahani, "Predicting compressive strength and electrical resistivity of eco-friendly concrete containing natural zeolite via GEP algorithm," *Constr. Build. Mater.*, vol. 229, p. 116883, 2019, doi: 10.1016/j.conbuildmat.2019.116883.

15. F. Farooq *et al.*, "A comparative study of random forest and genetic engineering programming for the prediction of compressive strength of high strength concrete (HSC)," *Appl. Sci.*, vol. 10, no. 20, pp. 1–18, 2020, doi: 10.3390/app10207330.

16. P. Chopra, R. K. Sharma, M. Kumar, and T. Chopra, "Comparison of Machine Learning Techniques for the Prediction of Compressive Strength of Concrete," *Adv. Civ. Eng.*, vol. 2018, 2018, doi: 10.1155/2018/5481705.

17. M. Nematzadeh, A. A. Shahmansouri, and M. Fakoor, "Post-fire compressive strength of recycled PET aggregate concrete reinforced with steel fibers: Optimization and prediction via RSM and GEP," *Constr. Build. Mater.*, vol. 252, p. 119057, 2020, doi: 10.1016/j.conbuildmat.2020.119057.

18. A. Gogineni, I. K. Panday, P. Kumar, and R. kr Paswan, "Predictive modelling of concrete compressive strength incorporating GGBS and alkali using a machine-learning approach," *Asian J. Civ. Eng.*, vol. 25, no. 1, pp. 699–709, 2024, doi: 10.1007/s42107-023-00805-z.

19. H. S. Abdelgader and A. S. El-Baden, "Effect of silica fume on two-stage concrete strength," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 96, no. 1, 2015, doi: 10.1088/1757-899X/96/1/012043.

20. O. A. Mohamed, M. Ati, and O. F. Najm, "Predicting compressive strength of sustainable self-consolidating concrete using random forest," *Key Eng. Mater.*, vol. 744 744 KE, pp. 141–145, 2017, doi: 10.4028/www.scientific.net/KEM.744.141.

21. C. J. Lin and N. J. Wu, "An ann model for predicting the compressive strength of concrete," *Appl. Sci.*, vol. 11, no. 9, 2021, doi: 10.3390/app11093798.

22. M. K. Elshaarawy, M. M. Alsaadawi, and A. K. Hamed, "Machine learning and interactive GUI for concrete compressive strength prediction," *Sci. Rep.*, vol. 14, no. 1, pp. 1–26, 2024, doi: 10.1038/s41598-024-66957-3.

23. S. Lee, N. H. Nguyen, A. Karamanli, J. Lee, and T. P. Vo, "Super learner machine-learning algorithms for compressive strength prediction of high performance concrete," *Struct. Concr.*, vol. 24, no. 2, pp. 2208–2228, 2023, doi: 10.1002/suco.202200424.

24. B. V. Varma, E. V. Prasad, and S. Singha, "Study on predicting compressive strength of concrete using supervised machine learning techniques," *Asian J. Civ. Eng.*, vol. 24, no. 7, pp. 2549–2560, 2023, doi: 10.1007/s42107-023-00662-w.
25. X. Qin and S. Kaewunruen, "Eco-Friendly Design and Sustainability Assessments of Fibre-Reinforced High-Strength Concrete Structures Automated by Data-Driven Machine Learning Models," *Sustain.*, vol. 15, no. 8, 2023, doi: 10.3390/su15086640.
26. H. V. T. Mai, T. A. Nguyen, H. B. Ly, and V. Q. Tran, "Prediction Compressive Strength of Concrete Containing GGBFS using Random Forest Model," *Adv. Civ. Eng.*, vol. 2021, no. Ml, 2021, doi: 10.1155/2021/6671448.
27. P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, 2006, doi: 10.1007/s10994-006-6226-1.
28. M. Davawala, T. Joshi, and M. Shah, "Compressive strength prediction of high-strength concrete using machine learning," *Emergent Mater.*, vol. 6, no. 1, pp. 321–335, 2023, doi: 10.1007/s42247-022-00409-4.
29. E. Mansouri, M. Manfredi, and J. W. Hu, "Environmentally Friendly Concrete Compressive Strength Prediction Using Hybrid Machine Learning," *Sustain.*, vol. 14, no. 20, 2022, doi: 10.3390/su142012990.
30. C. Qi, B. Huang, M. Wu, K. Wang, S. Yang, and G. Li, "Concrete Strength Prediction Using Different Machine Learning Processes: Effect of Slag, Fly Ash and Superplasticizer," *Materials (Basel).*, vol. 15, no. 15, 2022, doi: 10.3390/ma15155369.
31. G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, no. 2, pp. 197–227, 2016, doi: 10.1007/s11749-016-0481-7.
32. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, vol. 13-17-Augu, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.