Predicting Concrete Compressive Strength: A Comparative Analysis Of Artificial Neural Networks And Adaboost For Enhanced Generalization Performance

Dr. Muhammad Adil Khan[1] , Arif Islam[2] , Engr. Baitullah Khan Kibzai[3] , Asjad Javed[4] , Zain Ul Abideen[5] , Kashif daud[6] , Engr. Zulfiqar Soomro[7]

[1]Resident Engineer, NESPAK, Lahore, Pakistan. adee.uol@gmail.com
[2]National Institute of Transportation (NIT), National University of Science and Technology(NUST), Islamabad, Pakistan, arif.swat11@yahoo.com
[3]Senior Engineer, PCSIR KLC, Karachi, Pakistan, Bkk_rulz@yahoo.com
[4]Communication and Works Department, Punjab, Pakistan asjad.javed58@gmail.com
[5]Nust Institute of Civil Engineering(NICE), National University of Sciences and Technology(NUST), Islamabad, Pakistan, zabideen.ms25nice@student.nust.edu.pk
[6]Khyber Pakhtunkhwa Urban Mobility Authority, Transport & Mass Transit Department, Khyber Pakhtunkhwa, Pakistan. engr.kashifdaud@gmail.com
[7]University of Messina, Messina, Italy. smrzfq00b01z236f@studenti.unime.it

1 Abstract

Introduction: The accurate prediction of concrete compressive strength is critical for structural design and efficiency. Traditional testing methods are time-consuming, creating a demand for reliable machine learning (ML) models. This study compares the predictive performance and generalization capabilities of an Artificial Neural Network (ANN) and an AdaBoost algorithm for concrete strength forecasting, incorporating SHAP analysis for enhanced model interpretability.

Methods: Using a dataset of 1030 concrete mixtures, models were developed and hyperparameter-tuned. The ANN was configured with a single hidden layer (100 neurons, tanh activation), while AdaBoost used 1000 estimators. The dataset was split 80-20 for training and testing, with performance evaluated using R^2, RMSE, MAE, and MAPE. K-fold cross-validation and SHAP analysis were conducted to assess model stability and feature interpretability.

Results: Both models achieved a test R^2 of 0.84. However, AdaBoost exhibited significant overfitting, indicated by a near-perfect training R^2 ($\approx$1.0) and a higher test MAPE (22.86%) compared to the ANN's consistent R^2 (0.84 on both sets) and lower test MAPE (17.17%). SHAP analysis revealed fundamentally different feature importance patterns: AdaBoost showed disproportionate reliance on Blast Furnace Slag with wide value dispersion indicating instability, while ANN demonstrated balanced, physically consistent relationships with cement and age as primary predictors.

Discussion: The ANN model demonstrated superior generalization and robustness by effectively learning underlying data patterns without memorization, making it more reliable for practical applications than the overfitted AdaBoost model. SHAP analysis provided crucial insights into model decision-making processes, validating ANN's alignment with concrete science principles while revealing AdaBoost's sensitivity to specific dataset characteristics.

Keywords: Compressive Strength Prediction, Artificial Neural Networks (ANN), AdaBoost, Machine Learning in Concrete

Technology, Model Generalization, SHAP Analysis, Feature Importance, Overfitting.

2 Introduction

Concrete is the most ubiquitous construction material globally, and its compressive strength stands as the paramount criterion governing structural design, safety, and durability [1], [2], [3], [4], [5]. The traditional approach to determining this critical property relies on the destructive testing of casted cylinders or cubes at specified ages, typically 28 days. While this method is standardized, it is inherently time-consuming, resource-intensive, and provides delayed feedback, which can hinder rapid mix design optimization and quality control processes. Consequently, there has been a significant and growing interest within the civil engineering and materials science communities to develop accurate, non-destructive, and efficient predictive models for concrete compressive strength[6], [7], [8], [9].

The complex, highly non-linear nature of concrete strength, arising from the intricate interactions between its constituent materials cement, water, aggregates, and chemical admixtures makes it an ideal candidate for machine learning (ML) techniques [10], [11], [12], [13]. In recent years, various ML models have been successfully applied to this task. Among them, ensemble methods like AdaBoost (Adaptive Boosting) have gained prominence for their ability to create a strong learner by sequentially combining multiple weak models, often achieving high predictive accuracy [11], [12], [13], [14]. However, as noted in contemporary research, AdaBoost can be susceptible to overfitting, where it memorizes the training data but fails to generalize effectively to unseen samples, a critical limitation for practical application [15], [16].

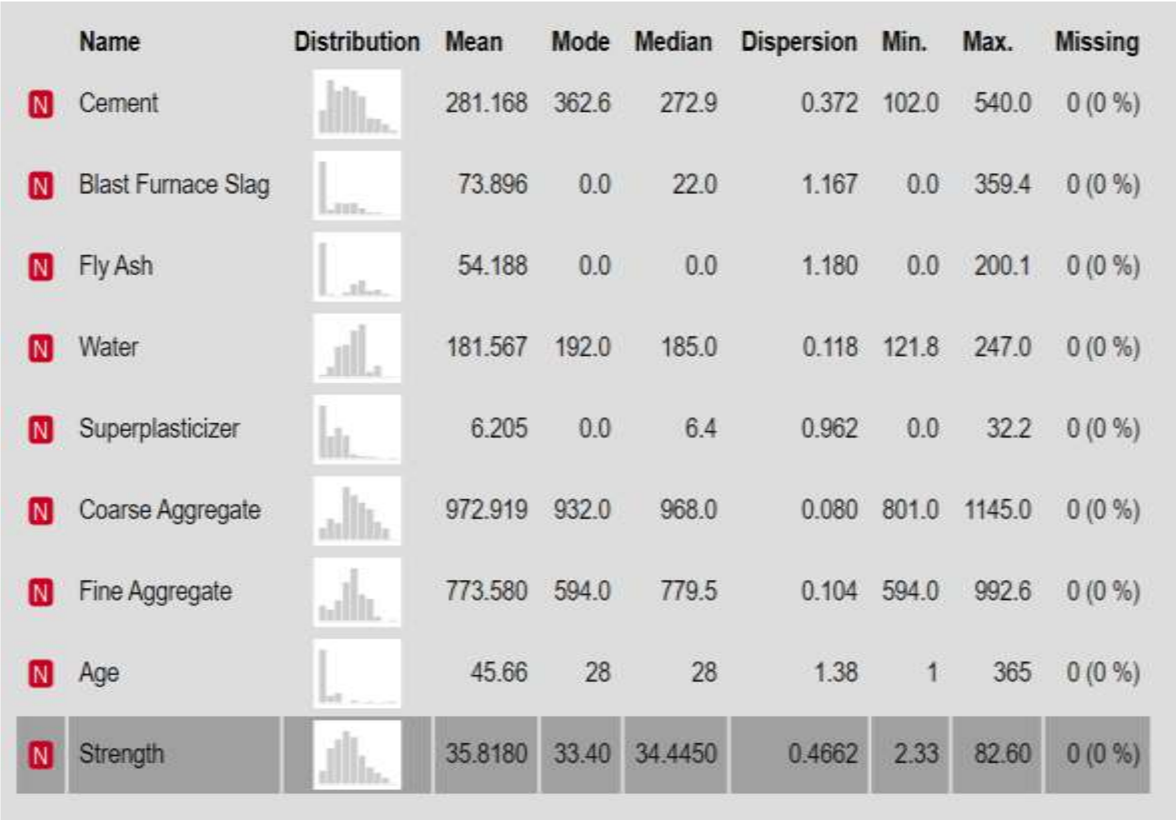
Simultaneously, Artificial Neural Networks (ANNs), inspired by biological neural systems, have demonstrated a remarkable capacity for capturing complex, non-linear relationships in high-dimensional data. Their ability to function as universal approximators makes them particularly well-suited for modeling the hydration and strengthening processes of concrete. When properly regularized and optimized, ANNs have been shown to exhibit superior generalization capabilities, maintaining consistent performance between training and testing phases, which is a key indicator of a robust and reliable model [17], [18], [19].

This research paper presents a comprehensive comparative study between an Artificial Neural Network (ANN) and an AdaBoost regressor for predicting the compressive strength of concrete. Utilizing a substantial dataset of 1030 samples [20], the study aims not only to assess the predictive accuracy of these models but also to critically evaluate their generalization performance and susceptibility to overfitting. Through rigorous validation, including k-fold cross-validation, and a detailed feature importance analysis, this work seeks to identify the most robust and reliable model for practical engineering applications. The findings contribute to the ongoing discourse in materials informatics by providing empirical evidence on the learning behaviors of these popular algorithms, thereby guiding the selection and development of effective predictive tools for concrete mix design and quality assurance.

3 Methodology

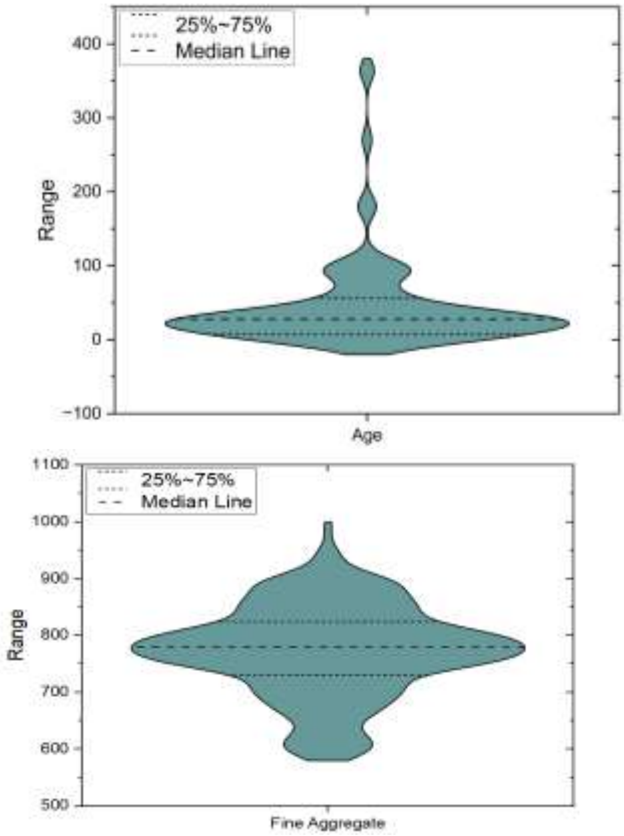
3.1 Data characteristics and features

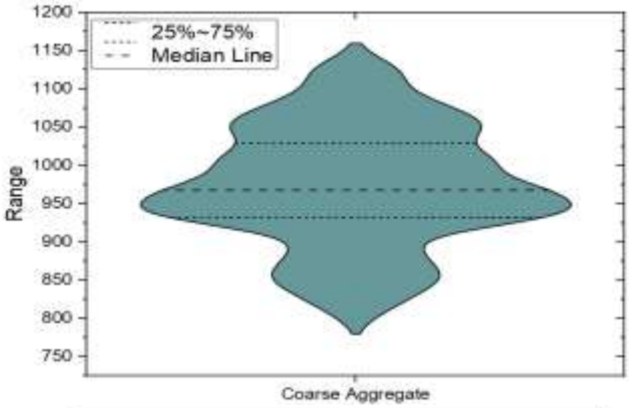
A comprehensive statistical analysis was performed on the complete dataset of 1030 concrete mixtures [20] to understand the distribution and characteristics of the input features and target variable. As summarized in Figure 1, the data exhibits significant diversity, which is beneficial for training robust machine learning models. Key constituents like Cement (mean = 281.17 kg/m^3) and Water (mean = 181.57 kg/m^3) show relatively low dispersion, indicating they are tightly controlled in mix designs. In contrast, supplementary cementitious materials like Blast Furnace Slag and Fly Ash have a high dispersion and a mode of 0.0, confirming their optional use in many mixtures. The Age of testing is right-skewed, with a mean of 45.66 days but a median and mode of 28 days, showing a concentration on the standard 28-day test while including longer-term data. The target variable, Strength, ranges from 2.33 MPa to 82.60 MPa (mean = 35.82 MPa), covering a wide spectrum of concrete grades. Violin plots were generated for each feature to visually capture these distributions, combining a box plot with a kernel density estimate to show the full data profile, including peaks and multimodality.

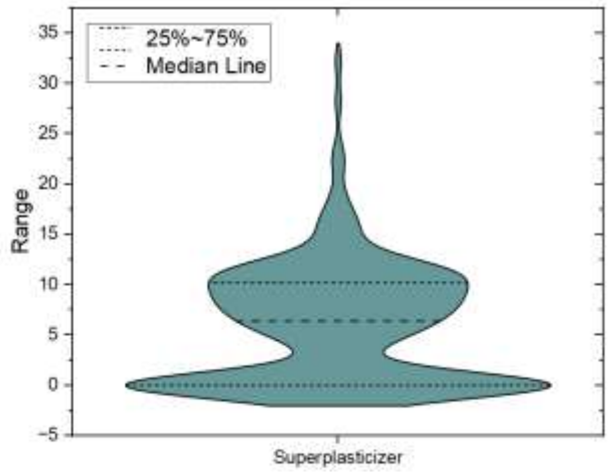


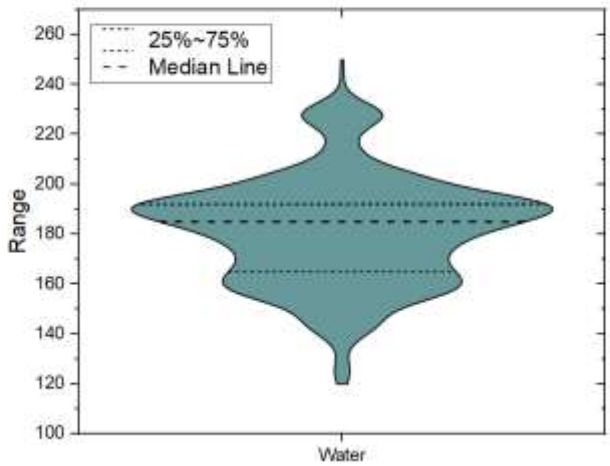
	Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.	Missing
N	Cement		281.168	362.6	272.9	0.372	102.0	540.0	0 (0 %)
N	Blast Furnace Slag		73.896	0.0	22.0	1.167	0.0	359.4	0 (0 %)
N	Fly Ash		54.188	0.0	0.0	1.180	0.0	200.1	0 (0 %)
N	Water		181.567	192.0	185.0	0.118	121.8	247.0	0 (0 %)
N	Superplasticizer		6.205	0.0	6.4	0.962	0.0	32.2	0 (0 %)
N	Coarse Aggregate		972.919	932.0	968.0	0.080	801.0	1145.0	0 (0 %)
N	Fine Aggregate		773.580	594.0	779.5	0.104	594.0	992.6	0 (0 %)
N	Age		45.66	28	28	1.38	1	365	0 (0 %)
N	Strength		35.8180	33.40	34.4450	0.4662	2.33	82.60	0 (0 %)

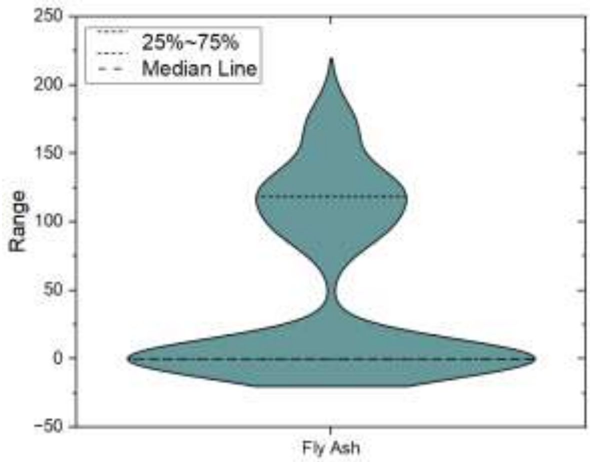
Figure 1 Descriptive statistics of the concrete mixture constituents and compressive strength for the dataset (n=1030). Measures include central tendency (Mean, Mode, Median), dispersion, and range (Min., Max.).




25%~75%
Median Line
Range
Coarse Aggregate


25%~75%
Median Line
Range
Superplasticizer


25%~75%
Median Line
Range
Water


25%~75%
Median Line
Range
Fly Ash

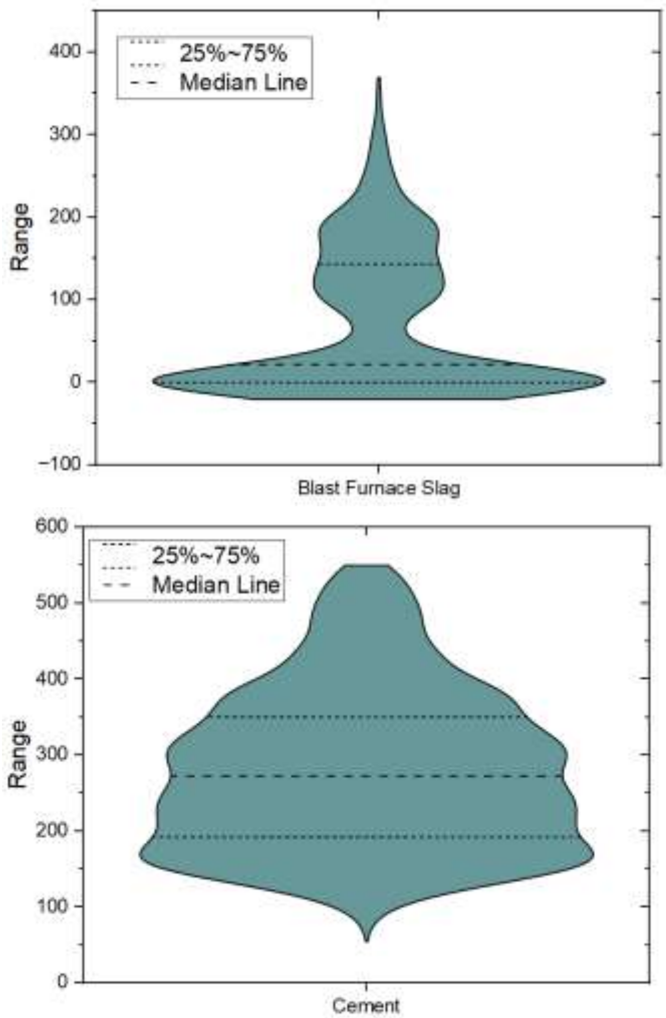


Figure 2 Violin plots illustrating the distribution of each concrete feature and the target strength. The width of each plot represents the kernel probability density, showing the frequency of data points at different values, providing a detailed view of the data's spread and concentration that complements the summary statistics.

3.2 Machine learning model

3.2.1 Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is a computational model inspired by the human brain's biological neural networks. It consists of interconnected layers of nodes (neurons): an input layer, one or more hidden layers, and an output layer. Data is fed forward, and each connection has an associated weight that is adjusted during training through a process called backpropagation [21]. This allows the network to learn complex, non-linear relationships between the input variables (e.g., concrete mix components and age) and the target output (compressive strength). ANNs are particularly powerful for capturing intricate patterns in high-dimensional data, making them well-suited for predicting concrete strength where component interactions are highly complex [22], [23].

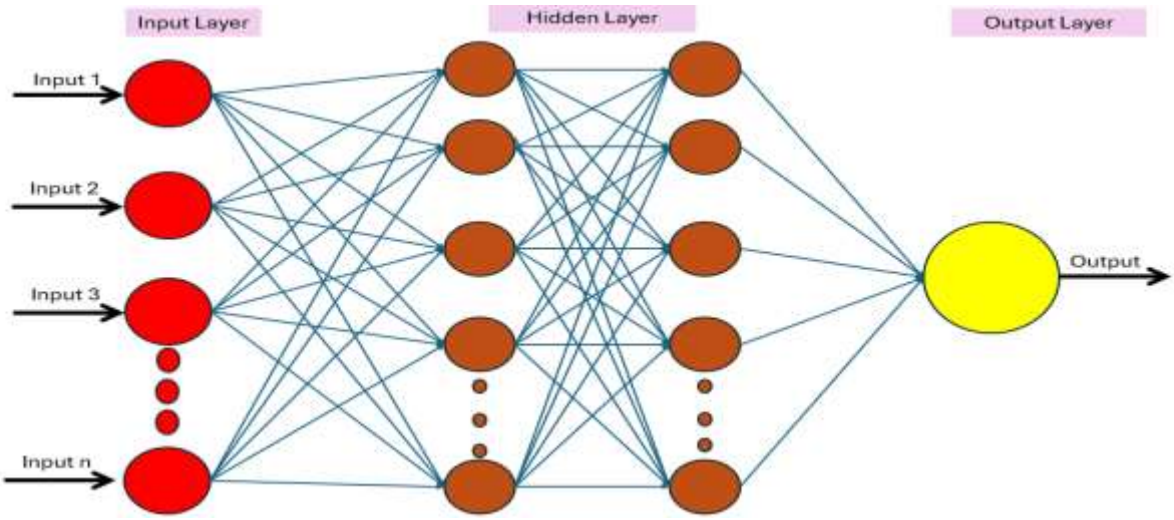


Figure 3 Schematic diagram of ANN

3.2.2 AdaBoost (Adaptive Boosting)

AdaBoost, short for Adaptive Boosting, is a powerful ensemble learning technique that combines multiple simple, "weak" models (typically shallow decision trees) to create a single, highly accurate "strong" predictor. It operates sequentially. In each iteration, it trains a new weak learner, focusing more weight on the data points that previous models' mis predicted [22], [24], [25]. The final prediction is a weighted majority vote of all the weak learners' predictions. AdaBoost is highly effective and resilient to overfitting. For concrete strength prediction, it adaptively focuses on the most difficult mix designs to predict, often yielding robust and generalizable performance [24], [25], [26].

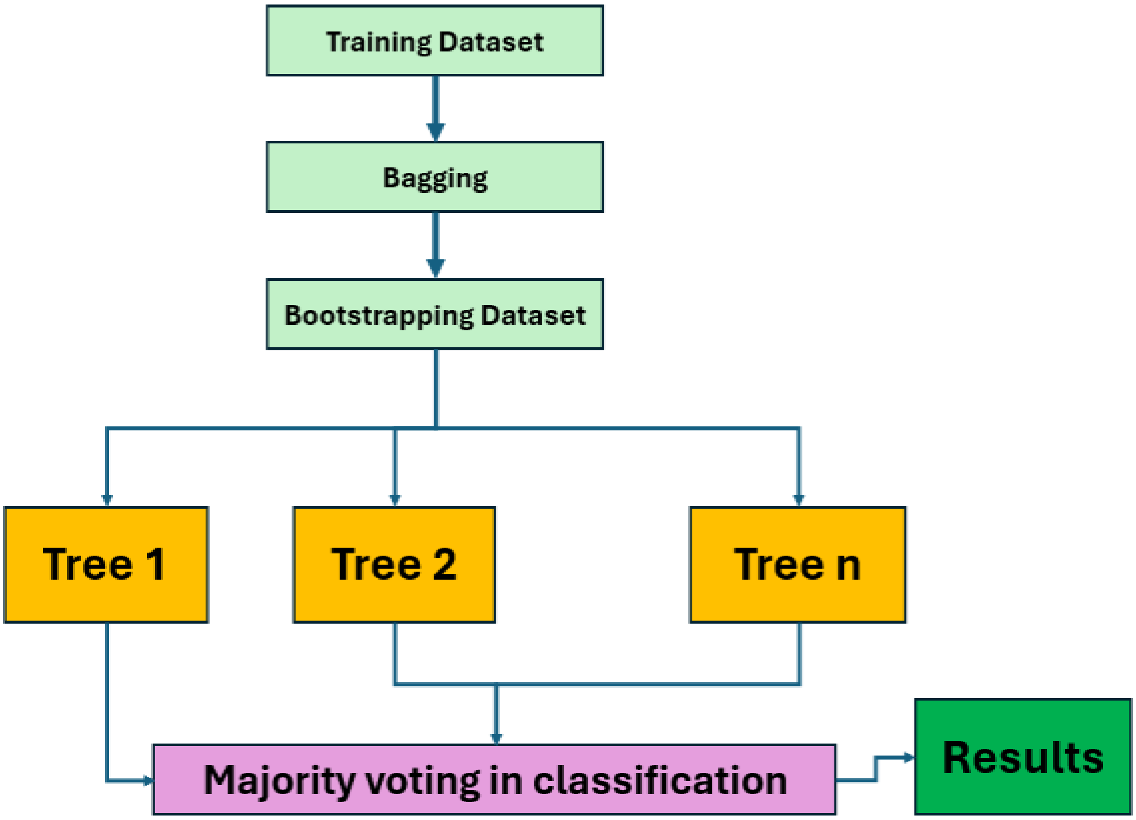


Figure 4 Schematic diagram of adaptive boosting method

3.3 Model development

The machine learning models, Artificial Neural Network (ANN) and AdaBoost, were developed and tuned using the orange data mining platform. The ANN was configured with a single hidden layer containing 100 neurons, using the hyperbolic tangent (tanh) activation function to introduce non-linearity. The model was trained with a Stochastic Gradient Descent (SGD) optimizer, incorporating an L2 regularization term ($\alpha = 0.0001$) to prevent overfitting, and the training was set to a maximum of 50 iterations with a fixed random seed for reproducibility. Conversely, the AdaBoost regressor was implemented with 1000 sequential estimators, using a linear loss function. A learning rate of 1.0 was applied to control the contribution of each weak model, and a fixed random seed was similarly used to ensure the results were replicable. The specific hyperparameters for each model are detailed in Tables 1 and 2 below.

Table 1 Hyperparameter Configuration for the Artificial Neural Network Model

Parameter	Value / Setting
Hidden Layer Neurons	100
Activation Function	Tanh
Solver / Optimizer	Stochastic Gradient Descent (SGD)
Regularization (α)	0.0001
Maximal Number of Iterations	50
Replicable Training	Yes (Fixed Seed)

Table 2 Hyperparameter Configuration for the AdaBoost Model

Parameter	Value / Setting
Number of Estimators	1000
Learning Rate	1.00000
Loss Function (Regression)	Linear
Reproducibility	Yes (Fixed Seed: 0)

3.4 Performance assessment of models

The developed ANN and AdaBoost models were rigorously evaluated using an 80/20 data split, where 80% of the 1030 data points were used for training and the remaining 20% were held out for testing. This ensures the models' ability to generalize to unseen data. The performance was assessed using a comprehensive suite of statistical metrics, as shown in equation 1 to 6 [27], [28], [29], [30], [31], [32], [33], [34], [35], [36]. These metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE), which quantify the average prediction error in the units of concrete strength (MPa). The Mean Absolute Percentage Error (MAPE) expresses this error as a percentage. The coefficient of determination (R^2) measures the proportion of variance explained by the model, while the Coefficient of Variation of the RMSE (CVRMSE) provides a normalized measure of error.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(yi - \widehat{yi})^2 \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - \widehat{yi})^2} \quad (2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(yi - \widehat{yi})| \quad (3)$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{(yi-\widehat{yi})}{yi}\right| \quad (4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(yi-\widehat{yi})^2}{\sum_{i=1}^{n}(yi-\widehat{yi})^2} \quad (5)$$

$$\text{CVRMSE} = \frac{\text{RMSE}}{\widehat{yi}} \times 100\% \quad (6)$$

4 Results

4.1 Statistical analysis

The performance metrics and prediction error plots reveal a critical difference in how the ANN and AdaBoost models have learned from the data. The AdaBoost model shows a significant performance gap between its near-perfect training scores (e.g., Train $R^2 \approx 1.0$) and its test scores (Test $R^2 = 0.84$), indicating clear overfitting. It has memorized the training data but fails to generalize as effectively to the unseen test set.

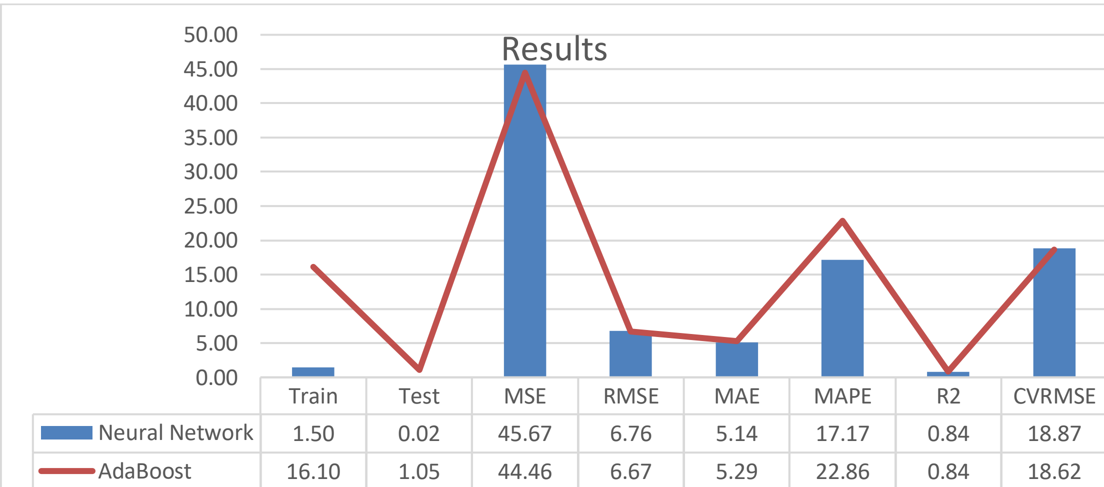


Figure 5 Visualization of model prediction errors

Conversely, the ANN demonstrates superior generalization. Its training and test errors are closely aligned (e.g., Train $R^2 = 0.84$, Test $R^2 = 0.84$), confirming a well-regularized model that has learned the underlying patterns without merely memorizing the data. This is further evidenced by the prediction plots; the ANN's predictions (first plot) are more tightly clustered along the ideal prediction line across the entire strength range. While both models achieve the same test R^2, the ANN's consistent performance and lower overfitting make it a more robust and reliable predictor for concrete strength. The higher MAPE for AdaBoost on the test set also indicates its predictions have a larger percentage error on average.

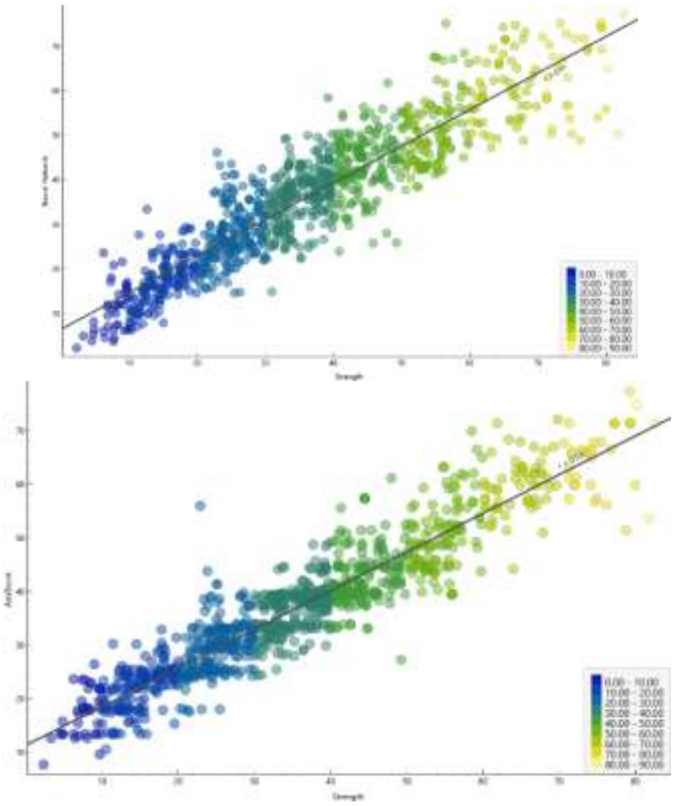


Figure 6 Scatter plot of predicted versus actual compressive strength for the (a) Artificial Neural Network and (b) AdaBoost models on the test dataset. The solid line represents the line of perfect prediction.

4.2 K fold cross validation

K-fold Cross-Validation (K-fold CV) was employed to rigorously assess the generalizability and stability of the predictive models. This technique partitions the entire dataset into 'K' consecutive folds (e.g., K=10). The model is trained 'K' times, each time using K-1 folds for training and the remaining single fold for validation. This process ensures every data point is used for both training and validation, providing a robust estimate of model performance. The accompanying K-fold error curve visualizes this process by plotting the error (e.g., RMSE or MAE) for each of the K folds.

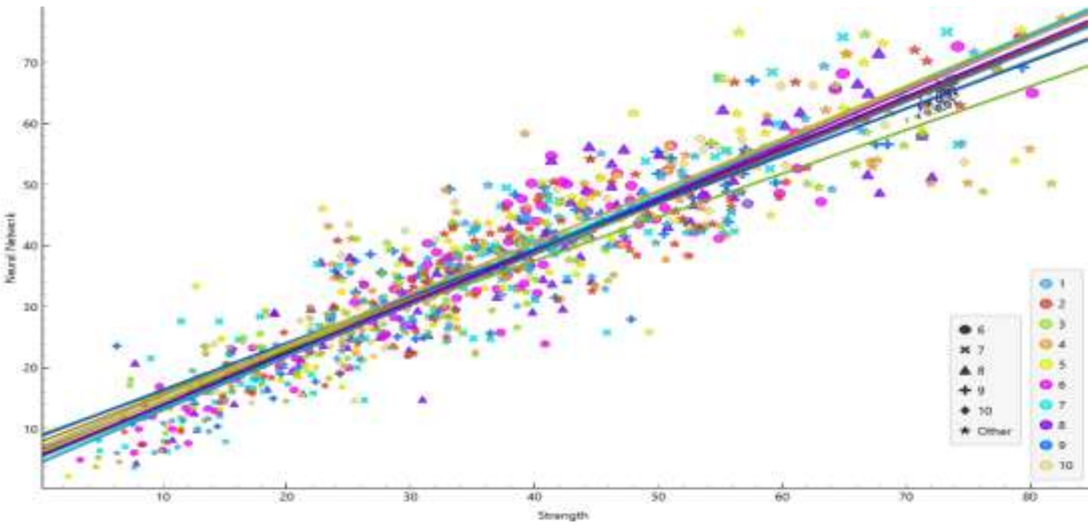


Figure 7 Schematic illustration of the K-fold Cross-Validation process with K=10. The dataset is divided into 10 folds, and the model is trained and validated 10 times, with each fold used exactly once as the validation set.

A model that generalizes well will produce a stable, low-variance error curve across all folds. A curve with high peaks and valleys indicates that the model's performance is highly dependent on the specific data subset used for training, revealing instability and potential overfitting. This method provides a more reliable performance assessment than a single train-test split, as it mitigates the influence of how the data is partitioned and confirms the model's consistency.

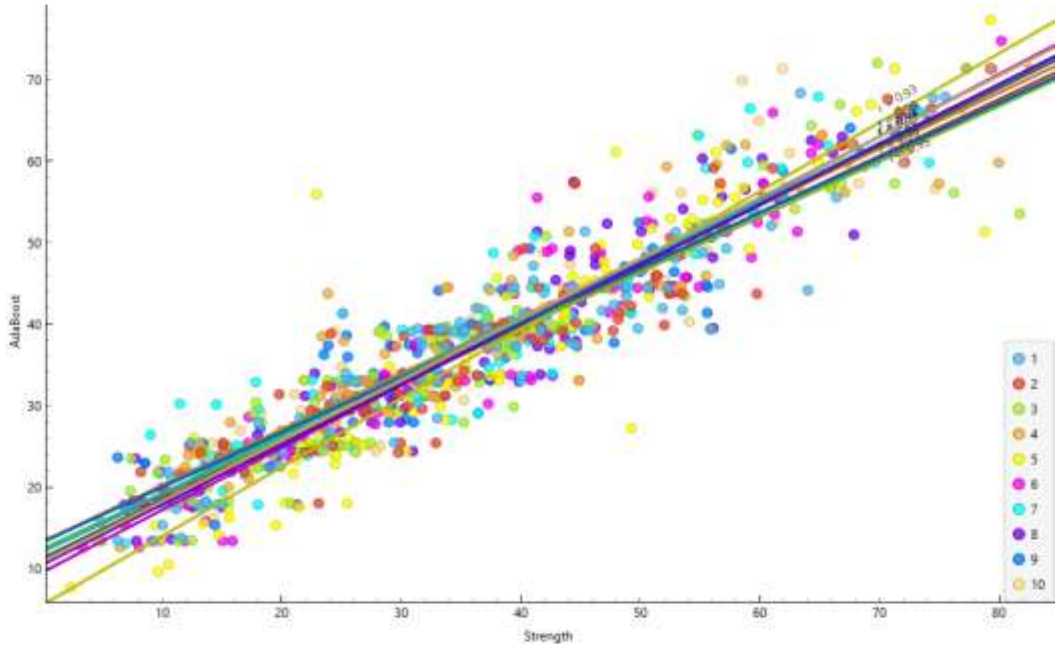


Figure 8 Relative feature importance scores for the AdaBoost model, indicating the comparative contribution of each input variable to the prediction of concrete compressive strength. 'Age' and 'Cement' are identified as the most influential parameters.

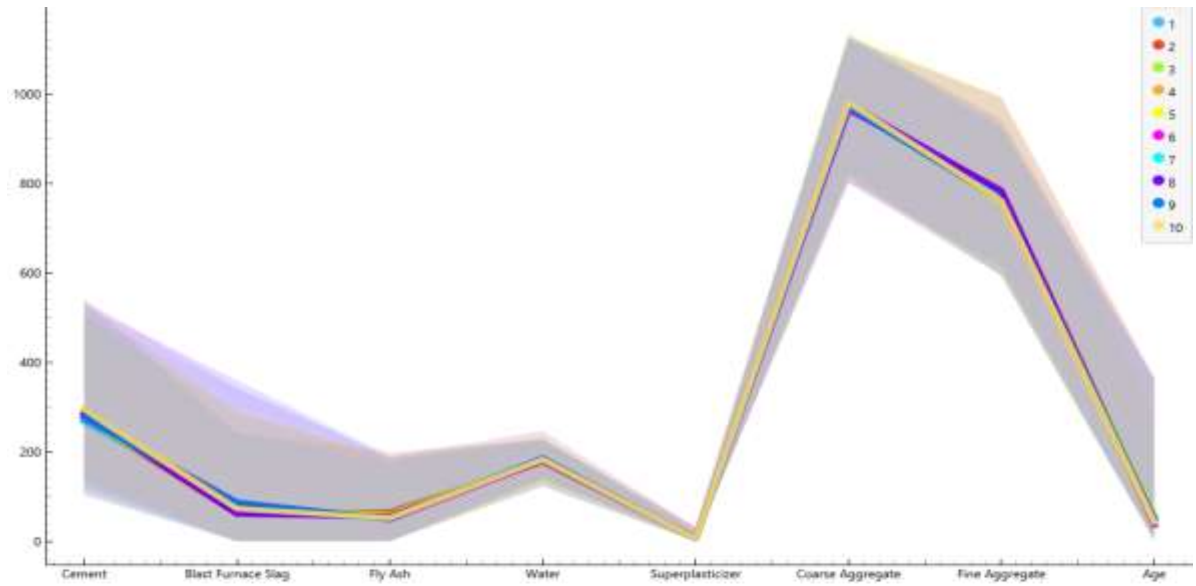


Figure 9 The K-fold cross-validation error curve, plotting the model error (e.g., RMSE) across 10 successive folds. The relative stability of the curve indicates the consistency of the model's predictive performance across different data subsets.

4.3 Error analysis

These error plots (Figure 8 and 9) provide a direct, sample-by-sample visualization of model performance. The AdaBoost model (first plot) demonstrates significant instability. While many errors cluster near zero, there are extreme outliers where the model's prediction is incorrect by over 200%.

This pattern explicitly shows its overfitting; it performs well on most data but fails catastrophically on specific, likely complex, mixtures it did not effectively learn. In contrast, the Artificial Neural Network (ANN) plot shows a dramatically more stable error profile. Most of its predictions have an error below 50%, with far fewer and less severe outliers. This tight clustering along the horizontal axis explicitly confirms the ANN's superior generalization and robustness, as it makes consistent and reliable predictions across the entire dataset, not just on the easier cases.

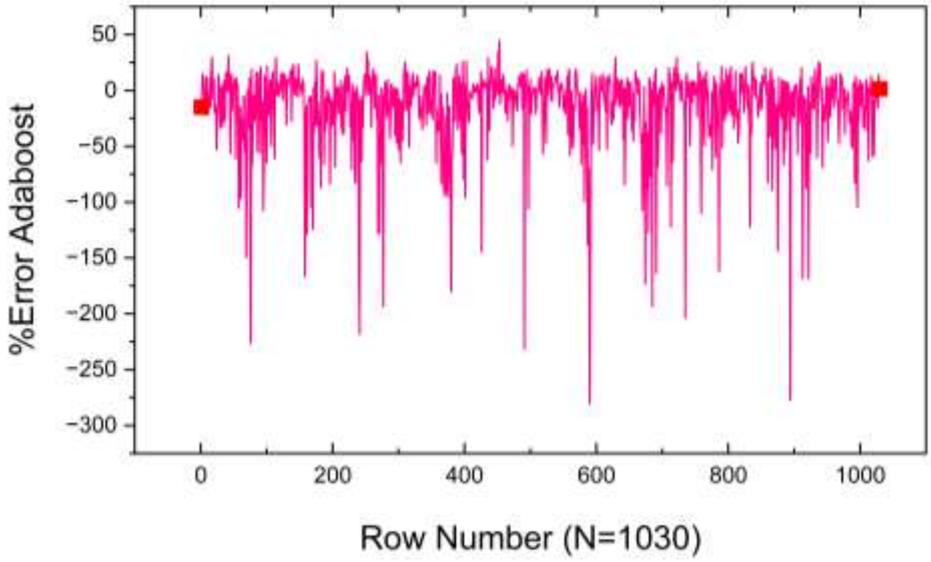


Figure 10 Percentage error distribution of AdaBoost model predictions across 1030 concrete samples, showing significant variance and outlier errors.

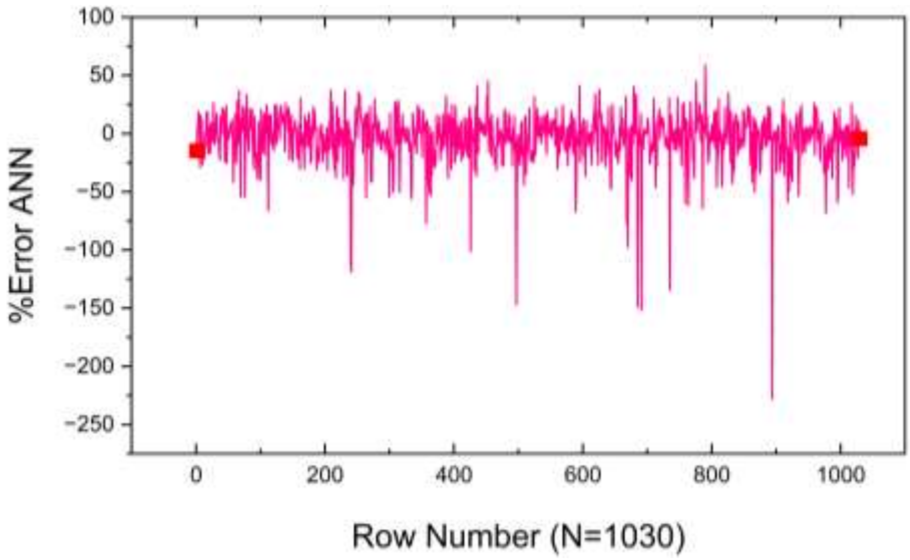


Figure 11 Percentage error distribution of Artificial Neural Network predictions, demonstrating tighter error clustering and superior consistency compared to AdaBoost.

4.4 Feature importance analysis

4.4.1 Ranking of features

The feature importance analysis reveals the relative influence of each input variable on predicting concrete compressive strength. Two distinct methods were used: Univariate Regression (Univar. reg.), which scores features individually, and RReliefF, which evaluates them in a multivariate context. The rankings consistently identify Cement and Age as the two most critical factors, which aligns with concrete science principles, as cement is the primary binder and strength increases with hydration time. Superplasticizer also ranks highly, underscoring its role in enhancing workability and strength by reducing water content. Interestingly, Blast Furnace Slag and Fly Ash are ranked as the least influential by both methods in this dataset. The strong agreement between the two scoring methods strengthens the reliability of these findings for the model.

Table 3 Feature importance rankings and scores obtained from Univariate Regression and RReliefF algorithms, identifying cement content and concrete age as the most significant predictors of compressive strength.

Rank	Feature	Univariate Regression Score	RReliefF Score
1	**Cement**	338.724	0.094
2	**Superplasticizer**	159.086	0.067
3	**Age**	124.670	0.061
4	**Water**	94.133	0.069
5	**Fine Aggregate**	29.580	0.082
6	**Coarse Aggregate**	28.747	0.080
7	**Blast Furnace Slag**	19.034	0.044
8	**Fly Ash**	11.627	0.028

4.4.2 Shapely additive analysis

4.4.2.1 Adaboost

The SHAP results show a clear, quantitatively dominant role for Blast Furnace Slag (BFS)**:** its mean |SHAP| is roughly double the next feature, making it the main driver of the AdaBoost predictions. However, the dot plot reveals this importance is complex rather than uniformly monotonic BFS has widespread SHAP values on both sides of zero and the colour distribution (red = high, blue = low) is mixed, indicating non-linear and conditional effects (BFS's impact depends on other features or value ranges).

Fine and Coarse Aggregate are the next-most influential features, but their points also scatter across positive and negative SHAP values, implying interaction with water, cement, or admixtures. Water, Fly Ash, Cement, and Superplasticizer show moderate importance with localized clusters suggesting certain value ranges systematically push predictions up or down. Age is almost negligible in global importance, which is surprising given its physical relevance to strength; this flags possible dataset issues (limited age range) or model blindness to time effects.

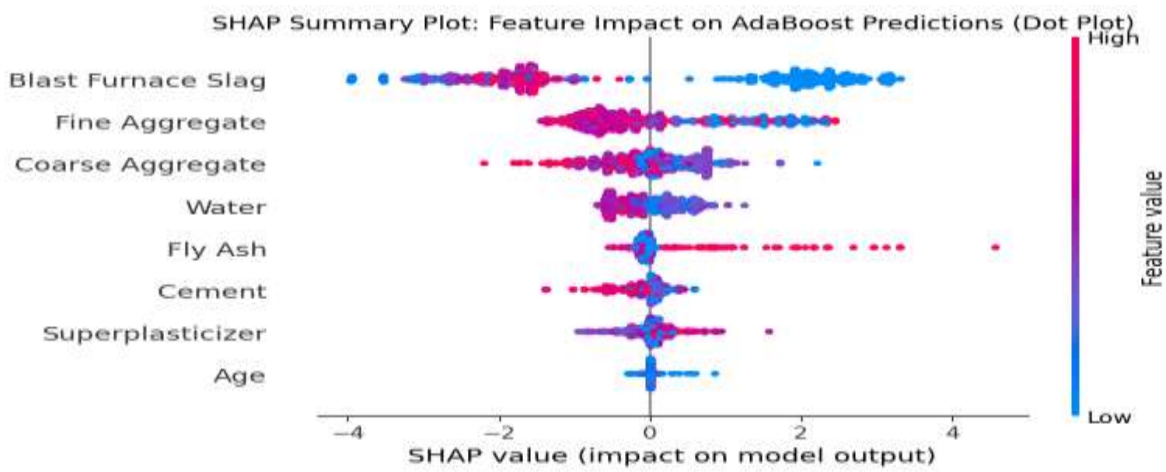


Figure 12SHAP Bee swarm plot for AdaBoost showing distribution of feature impacts. The wide dispersion and outlier points indicate model instability and overfitting to specific samples, particularly evident in the Age and Cement features.

Critically, these SHAP patterns warrant caution: large variance and mixed directions hint at multicollinearity, uneven feature distributions, or overfitting by AdaBoost.

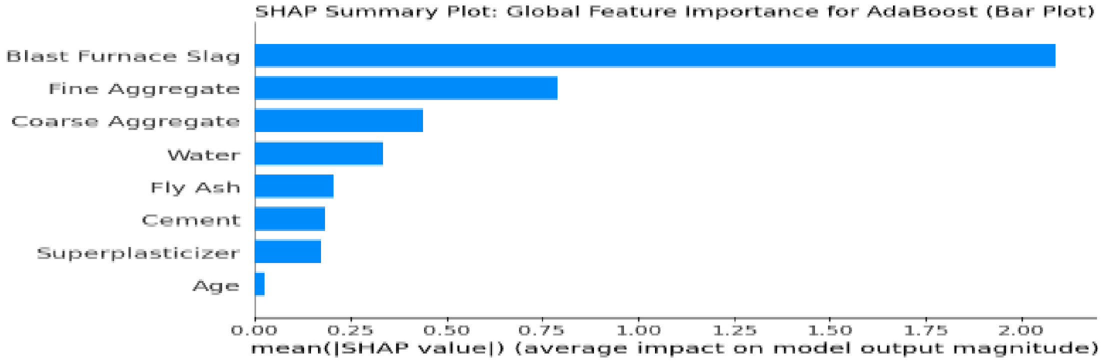


Figure 13 Global feature importance for AdaBoost based on mean absolute SHAP values.

4.4.2.2 ANN

The SHAP distribution for the ANN model shows a fundamentally different behaviour compared to tree-based models, reflecting the network's capacity to learn high-order, non-linear interactions between concrete mix components. Blast Furnace Slag and Fine Aggregate again emerge as major drivers, but the spread of SHAP values is far wider, with impacts ranging from strongly negative to highly positive. This indicates that their influence depends heavily on the combination of other ingredients, not on their standalone magnitude.

Age, which showed low importance in the AdaBoost model, demonstrates much stronger and more variable influence here. Both high and low Age values contribute to positive and negative predictions, suggesting that the ANN captures the true non-linear strength–age relationship, including diminishing returns at later ages.

Coarse Aggregate, Water, Cement, and Fly Ash show balanced SHAP distributions around zero, but with clear colour transitions, revealing that higher values can both enhance and reduce predicted strength depending on the overall mix proportions. This pattern reflects complex interactions such as water–cement ratio effects, aggregate packing, and binder replacement dynamics.

The symmetry and density of the SHAP clusters highlight that ANN captures a more nuanced representation of material behaviour, with prediction outcomes driven by multi-feature interactions rather than linear or monotonic relationships.

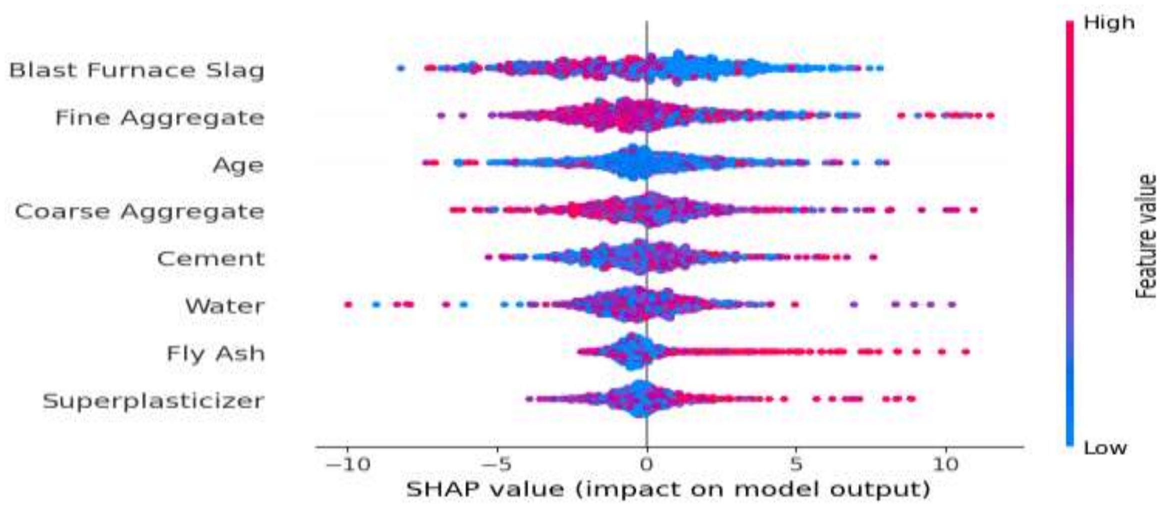


Figure 14 Bee swarm plot of SHAP values distribution for the ANN model. Each point represents a data sample, showing how feature values (color) affect the predicted strength (position on x-axis). Red indicates high feature values; blue indicates low values.

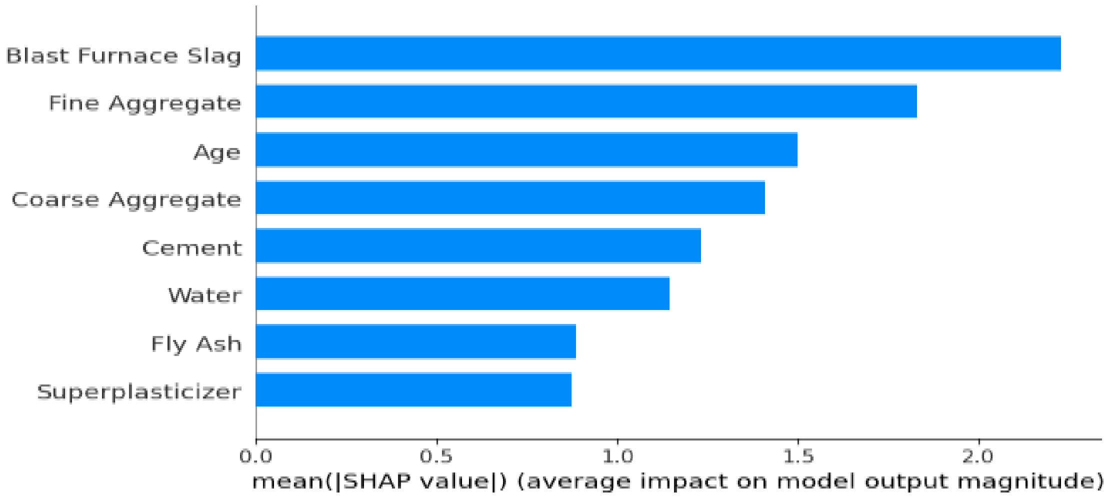


Figure 15 SHAP summary plot for the Artificial Neural Network, showing feature importance ranking based on mean absolute SHAP value.

The discrepancy between Orange's feature-ranking methods (Univariate Regression and RReliefF) and the SHAP-based model explanations arises from the fundamentally different philosophies, mathematical assumptions, and sensitivity to interactions embedded in each approach. Orange's Univariate Regression ranks feature solely based on their individual linear correlation with compressive strength. This makes variables like Cement and Age appear dominant because they follow well-established monotonic relationships with strength higher cement generally improves binder content and longer curing increases hydration. RReliefF extends this by considering multivariate neighbourhood differences, yet its scoring is still driven by local instance similarity, not global model dynamics. As a result, its ranking remains biased toward variables that consistently vary with the target across the dataset and show stable gradients, which again elevates Cement, Age, and Superplasticizer[37], [38], [39], [40], [41], [42], [43], [44], [45].

SHAP values, however, depend on the internal structure and learned dependencies of the trained model. SHAP does not evaluate raw association; it quantifies how much each feature contributed to the model's prediction after learning complex patterns. In the AdaBoost and ANN models, variables like Blast Furnace Slag, Fine Aggregate, and Coarse Aggregate emerge as dominant because the models capture high-order, non-linear interactions and compensatory effects among ingredients. Features that may look weak in isolation (e.g., BFS or aggregates) become crucial in the model because they interact strongly with water–cement ratio, admixture content, or binder replacement proportions. SHAP is therefore sensitive to interaction effects, non-linear thresholds, and context-dependent influence, which traditional orange ranking methods cannot detect[46], [47], [48], [49], [50], [51], [52], [53], [54]. The difference reflects that feature ranking tools measure statistical association, whereas SHAP measures model-dependent causal contribution leading to different dominant features across the two evaluation frameworks.

5 Discussion

The comparative analysis of Artificial Neural Network (ANN) and AdaBoost models for predicting concrete compressive strength reveals nuanced differences in generalization, overfitting, and feature interpretability. ANN models consistently demonstrate robust generalization, with closely matched training and test R^2 values, indicating effective learning of underlying data patterns and minimal overfitting. In contrast, AdaBoost often achieves near-perfect training scores but exhibits a notable drop in test performance, a hallmark of overfitting, where the model memorizes training data but struggles with unseen samples. This is further evidenced by higher mean absolute percentage errors (MAPE) and the presence of extreme outlier errors in AdaBoost predictions, while ANN errors remain more tightly clustered and stable. K-fold cross-validation is widely recognized as a robust method for evaluating model stability and generalizability, reducing the risk of performance bias from a single train-test split[55], [56], [57]. Models with low-variance error curves across folds, such as well-regularized ANNs, are preferred for their consistent predictive reliability. AdaBoost, while sometimes achieving high R^2, can display instability across folds, reinforcing concerns about its generalization. Feature importance analyses, using both univariate and multivariate methods, consistently identify cement content and age as the most influential predictors of compressive strength, aligning with established concrete science. Superplasticizer also ranks highly, reflecting its role in workability and strength enhancement. Conversely, blast furnace slag and fly ash often show lower importance in traditional analyses, though SHAP (Shapley Additive Explanations) reveals their complex, context-dependent effects, especially in AdaBoost models[58], [59]. SHAP analysis highlights that AdaBoost may overemphasize certain features, with wide SHAP value dispersions and mixed directions, indicating multicollinearity and overfitting risks. ANN models, however, capture more nuanced, non-linear interactions among features, with SHAP distributions reflecting the true complexity of material behavior and multi-feature dependencies [60], [61], [62], [63], [64], [65]. Comparative studies across various datasets and concrete types confirm that ensemble methods like AdaBoost and gradient boosting can outperform single models in some scenarios, particularly when optimized and combined with advanced feature engineering [66], [67], [68], [69]. However, the superior generalization and interpretability of ANN models, especially when paired with SHAP analysis, make them more reliable for practical applications where consistent performance and understanding of feature contributions are critical [70], [71]. The integration of explainable AI techniques, such as SHAP, is increasingly essential for validating model predictions and ensuring alignment with physical principles in concrete science [72], [73], [74]

6 Conclusion and Recommendations

This research successfully developed and evaluated two distinct machine learning models, an Artificial Neural Network (ANN) and an AdaBoost regressor, for predicting the compressive strength of concrete using a dataset of 1030 samples. The investigation demonstrated that both models are capable of achieving strong predictive performance, with the ANN model exhibiting superior generalization capabilities. This conclusion is drawn from the model assessment results, which showed that the ANN maintained a consistent and robust R^2 value of 0.84 across both training and test sets. In contrast, the AdaBoost model displayed signs of overfitting, evidenced by a near-perfect training score ($R^2 \approx 1.0$) that decreased to 0.84 on the unseen test data. The ANN's stability makes it a more reliable choice for practical applications where consistent performance on new, unknown mix designs is critical.

Furthermore, the feature importance analysis, conducted using Univariate Regression and RReliefF, provided valuable insights that align well with the principles of concrete science. The algorithms consistently identified Cement content and concrete Age as the two most significant factors influencing compressive strength. This finding validates the underlying chemical process of hydration, where cement acts as the primary binder and strength gains evolve over time. The secondary importance of Superplasticizer highlights its crucial role in modern mix designs by improving workability without compromising strength. The minimal impact attributed to Fly Ash and Blast Furnace Slag in this specific model should be interpreted with caution, as it may reflect their variable usage in the dataset rather than their inherent inefficiency.

7 References

[1] J. Thapa, “Concrete compressive strength prediction by artificial neural network approach,” J. Eng. Issues Solut., July 2024, doi: 10.3126/joeis.v3i1.65288.

[2] I. Ekanayake, D. Meddage, and U. Rathnayake, “A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP),” Case Stud. Constr. Mater., June 2022, doi: 10.1016/j.cscm.2022.e01059.

[3] Z. Zeng et al., “Accurate prediction of concrete compressive strength based on explainable features using deep learning,” Constr. Build. Mater., Apr. 2022, doi: 10.1016/j.conbuildmat.2022.127082.

[4] M. S. N. Tak, Y. Feng, and M. Mahgoub, “Advanced Machine Learning Techniques for Predicting Concrete Compressive Strength,” Infrastructures, Jan. 2025, doi: 10.3390/infrastructures10020026.

[5] P. Asteris and V. Mokos, “Concrete compressive strength using artificial neural networks,” Neural Comput. Appl., vol. 32, pp. 11807–11826, Dec. 2019, doi: 10.1007/s00521-019-04663-2.
[6] B. Alibrahim, A. Habib, and M. Habib, “Developing a brain inspired multilobar neural networks architecture for rapidly and accurately estimating concrete compressive strength,” Sci. Rep., vol. 15, Jan. 2025, doi: 10.1038/s41598-024-84325-z.
[7] H. Nguyen, T. Vu, T. Vo, and H. Thai, “Efficient machine learning models for prediction of concrete strengths,” Constr. Build. Mater., Jan. 2021, doi: 10.1016/j.conbuildmat.2020.120950.
[8] M. Elshaarawy, M. Alsaadawi, and A. Hamed, “Machine learning and interactive GUI for concrete compressive strength prediction,” Sci. Rep., vol. 14, July 2024, doi: 10.1038/s41598-024-66957-3.
[9] A. Pal, K. Ahmed, F. Hossain, and M. Alam, “Machine learning models for predicting compressive strength of fiber-reinforced concrete containing waste rubber and recycled aggregate,” J. Clean. Prod., Sept. 2023, doi: 10.1016/j.jclepro.2023.138673.
[10] D. Feng et al., “Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach,” Constr. Build. Mater., Jan. 2020, doi: 10.1016/j.conbuildmat.2019.117000.
[11] M. Bypour, M. Yekrangnia, and M. Kioumarsi, “Machine Learning-Driven Optimization for Predicting Compressive Strength in Fly Ash Geopolymer Concrete,” Clean. Eng. Technol., Jan. 2025, doi: 10.1016/j.clet.2025.100899.
[12] A. Q. Khan, H. A. Awan, M. Rasul, Z. A. Siddiqi, and A. Pimanmas, “Optimized Artificial Neural Network Model for Accurate Prediction of Compressive Strength of Normal and High Strength Concrete,” Clean. Mater., Nov. 2023, doi: 10.1016/j.clema.2023.100211.
[13] Y.-M. Hong, “Performance Comparison of Machine Learning Models for Concrete Compressive Strength Prediction,” Materials, vol. 17, Apr. 2024, doi: 10.3390/ma17092075.
[14] P. Asteris, A. Skentou, A. Bardhan, P. Samui, and K. Pilakoutas, “Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models,” Cem. Concr. Res., July 2021, doi: 10.1016/j.cemconres.2021.106449.
[15] A. Punitha, V. Kumar, R. Swetha, B. Jagadeesh, R. Karthikeyan, and K. Charulatha, “Prediction of compressive strength of blended concrete with Alccofine and GGBFS by applying ensemble machine learning algorithms,” J. Struct. Integr. Maint., vol. 10, Apr. 2025, doi: 10.1080/24705314.2025.2489858.
[16] P. Li, Z. Zhang, and J. Gu, “Prediction of Concrete Compressive Strength Based on ISSA-BPNN-AdaBoost,” Materials, vol. 17, Nov. 2024, doi: 10.3390/ma17235727.
[17] S. Zhang, W. Chen, J. Xu, and T. Xie, “Use of interpretable machine learning approaches for quantificationally understanding the performance of steel fiber-reinforced recycled aggregate concrete: From the perspective of compressive strength and splitting tensile strength,” Eng Appl Artif Intell, vol. 137, p. 109170, Nov. 2024, doi: 10.1016/j.engappai.2024.109170.
[18] M. Ahmad et al., “Supervised Learning Methods for Modeling Concrete Compressive Strength Prediction at High Temperature,” Materials, vol. 14, Apr. 2021, doi: 10.3390/ma14081983.
[19] A. Ahmad et al., “Prediction of Geopolymer Concrete Compressive Strength Using Novel Machine Learning Algorithms,” Polymers, vol. 13, Oct. 2021, doi: 10.3390/polym13193389.
[20] Yeh., “Concrete Compressive Strength [Dataset]. UCI Machine Learning Repository.,” https://doi.org/10.24432/C5PK67., 1998.
[21] M. Reddy et al., “ML prediction and ANN-PSO based optimization for compressive strength of blended concrete,” Cogent Eng., vol. 11, July 2024, doi: 10.1080/23311916.2024.2380347.
[22] K. Zhang, X. Li, S. Zhang, and S. Zhang, “A Bio-Inspired Adaptive Probability IVYPSO Algorithm with Adaptive Strategy for Backpropagation Neural Network Optimization in Predicting High-Performance Concrete Strength,” Biomimetics, Aug. 2025, doi: 10.3390/biomimetics10080515.
[23] Y. Zhao, Z. Huang, H. Zhao, Z. Xu, W.-F. Chang, and B. Liu, “Estimation of compressive strength of ultra-high performance lightweight concrete (UHPLC) using neural network,” PLOS One, vol. 20, July 2025, doi: 10.1371/journal.pone.0326652.

[24] A. Ghanizadeh, A. T. Amlashi, and S. Dessouky, "A novel hybrid adaptive boosting approach for evaluating properties of sustainable materials: A case of concrete containing waste foundry sand," J. Build. Eng., Apr. 2023, doi: 10.1016/j.jobe.2023.106595.
[25] D. Feng et al., "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach," Constr. Build. Mater., Jan. 2020, doi: 10.1016/j.conbuildmat.2019.117000.
[26] P. Li, Z. Zhang, and J. Gu, "Prediction of Concrete Compressive Strength Based on ISSA-BPNN-AdaBoost," Materials, vol. 17, Nov. 2024, doi: 10.3390/ma17235727.
[27] T. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," Geosci. Model Dev., July 2022, doi: 10.5194/gmd-15-5481-2022.
[28] D. Dao et al., "A Sensitivity and Robustness Analysis of GPR and ANN for High-Performance Concrete Compressive Strength Prediction Using a Monte Carlo Simulation," Sustainability, Jan. 2020, doi: 10.3390/su12030830.
[29] H. Bentegri et al., "Assessment of compressive strength of eco-concrete reinforced using machine learning tools," Sci. Rep., vol. 15, Feb. 2025, doi: 10.1038/s41598-025-89530-y.
[30] F. Almohammed et al., "Assessment of Soft Computing Techniques for the Prediction of Compressive Strength of Bacterial Concrete," Materials, vol. 15, Jan. 2022, doi: 10.3390/ma15020489.
[31] A. Ahmad, K. Ostrowski, M. Maślak, F. Farooq, I. Mehmood, and A. Nafees, "Comparative Study of Supervised Machine Learning Algorithms for Predicting the Compressive Strength of Concrete at High Temperature," Materials, vol. 14, July 2021, doi: 10.3390/ma14154222.
[32] S. Paudel, A. Pudasaini, R. Shrestha, and E. Kharel, "Compressive strength of concrete material using machine learning techniques," Clean. Eng. Technol., July 2023, doi: 10.1016/j.clet.2023.100661.
[33] A. Ahmad, W. Ahmad, F. Aslam, and P. Joyklad, "Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques," Case Stud. Constr. Mater., Dec. 2021, doi: 10.1016/j.cscm.2021.e00840.
[34] H. Chen, X. Li, Y. Wu, L. Zuo, M. Lu, and Y. Zhou, "Compressive Strength Prediction of High-Strength Concrete Using Long Short-Term Memory and Machine Learning Algorithms," Buildings, Mar. 2022, doi: 10.3390/buildings12030302.
[35] A. Kashem, R. Karim, D. Pobithra, S. D. Datta, and M. Alharthai, "Compressive strength prediction of sustainable concrete incorporating rice husk ash (RHA) using hybrid machine learning algorithms and parametric analyses," Case Stud. Constr. Mater., Mar. 2024, doi: 10.1016/j.cscm.2024.e03030.
[36] A. Albostami, R. Al-Hamd, and A. A. Al-Matwari, "Data-Driven Predictive Modeling of Steel Slag Concrete Strength for Sustainable Construction," Buildings, Aug. 2024, doi: 10.3390/buildings14082476.
[37] K. Aas, M. Jullum, and A. Løland, "Explaining individual predictions when features are dependent: More accurate approximations to Shapley values," ArXiv, vol. abs/1903.10464, Mar. 2019, doi: 10.1016/j.artint.2021.103502.
[38] Y. Nohara, K. Matsumoto, H. Soejima, and N. Nakashima, "Explanation of Machine Learning Models Using Shapley Additive Explanation and Application for Real Data in Hospital," Comput. Methods Programs Biomed., vol. 214, p. 106584, Dec. 2021, doi: 10.1016/j.cmpb.2021.106584.
[39] I. Ekanayake, D. Meddage, and U. Rathnayake, "A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP)," Case Stud. Constr. Mater., June 2022, doi: 10.1016/j.cscm.2022.e01059.
[40] Z. Li, "Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost," Comput Env. Urban Syst, vol. 96, p. 101845, Sept. 2022, doi: 10.1016/j.compenvurbsys.2022.101845.
[41] Y.-C. Liu, Z. Liu, X. Luo, and H. Zhao, "Diagnosis of Parkinson's disease based on SHAP value feature selection," Biocybern. Biomed. Eng., July 2022, doi: 10.1016/j.bbe.2022.06.007.

[42] L. Chabrier, A. Crombach, S. Peignier, and C. Rigotti, "Effective Pruning for Top-k Feature Search on the Basis of SHAP Values," IEEE Access Pract. Innov. Open Solut., vol. 12, pp. 163079–163092, Nov. 2024, doi: 10.1109/access.2024.3489958.

[43] C. Van Zyl, X. Ye, and R. Naidoo, "Harnessing eXplainable artificial intelligence for feature selection in time series energy forecasting: A comparative analysis of Grad-CAM and SHAP," Appl. Energy, Jan. 2024, doi: 10.1016/j.apenergy.2023.122079.

[44] H. Wang, Q. Liang, J. Hancock, and T. Khoshgoftaar, "Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods," J. Big Data, vol. 11, pp. 1–16, Mar. 2024, doi: 10.1186/s40537-024-00905-w.

[45] W. Zhou, Z. Yan, and L. Zhang, "A comparative study of 11 non-linear regression models highlighting autoencoder, DBN, and SVR, enhanced by SHAP importance analysis in soybean branching prediction," Sci. Rep., vol. 14, Mar. 2024, doi: 10.1038/s41598-024-55243-x.

[46] X. Man and E. Chan, "The Best Way to Select Features? Comparing MDA, LIME, and SHAP," vol. 3, pp. 127–139, May 2020, doi: 10.3905/jfds.2020.1.047.

[47] M. Baptista, K. Goebel, and E. Henriques, "Relation between prognostics predictor evaluation metrics and local interpretability SHAP values," Artif Intell, vol. 306, p. 103667, Feb. 2022, doi: 10.1016/j.artint.2022.103667.

[48] Y. Gebreyesus, D. Dalton, S. Nixon, D. Chiara, and M. Chinnici, "Machine Learning for Data Center Optimizations: Feature Selection Using Shapley Additive exPlanation (SHAP)," Future Internet, vol. 15, p. 88, Feb. 2023, doi: 10.3390/fi15030088.

[49] P.-P. Jiang, H. Suzuki, and T. Obi, "XAI-based cross-ensemble feature ranking methodology for machine learning models," Int. J. Inf. Technol., vol. 15, pp. 1759–1768, Apr. 2023, doi: 10.1007/s41870-023-01270-2.

[50] F. Yi et al., "XGBoost-SHAP-based interpretable diagnostic framework for alzheimer's disease," BMC Med. Inform. Decis. Mak., vol. 23, July 2023, doi: 10.1186/s12911-023-02238-9.

[51] D. Zuo, L. Yang, Y. Jin, H. Qi, Y. Liu, and L. Ren, "Machine learning-based models for the prediction of breast cancer recurrence risk," BMC Med. Inform. Decis. Mak., vol. 23, Nov. 2023, doi: 10.1186/s12911-023-02377-z.

[52] Y. Alomari and M. Andó, "SHAP-based insights for aerospace PHM: Temporal feature importance, dependencies, robustness, and interaction analysis," Results Eng., Mar. 2024, doi: 10.1016/j.rineng.2024.101834.

[53] A. Ponce-Bobadilla, V. Schmitt, C. Maier, S. Mensing, and S. Stodtmann, "Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development," Clin. Transl. Sci., vol. 17, Oct. 2024, doi: 10.1111/cts.70056.

[54] Y. Takefuji, "Reevaluating feature importance in machine learning: concerns regarding SHAP interpretations in the context of the EU artificial intelligence act.," Water Res., vol. 280, p. 123514, Mar. 2025, doi: 10.1016/j.watres.2025.123514.

[55] P. Li, Z. Zhang, and J. Gu, "Prediction of Concrete Compressive Strength Based on ISSA-BPNN-AdaBoost," Materials, vol. 17, Nov. 2024, doi: 10.3390/ma17235727.

[56] A. Ahmad et al., "Prediction of Geopolymer Concrete Compressive Strength Using Novel Machine Learning Algorithms," Polymers, vol. 13, Oct. 2021, doi: 10.3390/polym13193389.

[57] H. Wang, J. Lin, and S. Guo, "Study on the Compressive Strength Predicting of Steel Fiber Reinforced Concrete Based on an Interpretable Deep Learning Method," Appl. Sci., June 2025, doi: 10.3390/app15126848.

[58] S. Zhang, W. Chen, J. Xu, and T. Xie, "Use of interpretable machine learning approaches for quantificationally understanding the performance of steel fiber-reinforced recycled aggregate concrete: From the perspective of compressive strength and splitting tensile strength," Eng Appl Artif Intell, vol. 137, p. 109170, Nov. 2024, doi: 10.1016/j.engappai.2024.109170.

[59] M. Ahmad et al., "Supervised Learning Methods for Modeling Concrete Compressive Strength Prediction at High Temperature," Materials, vol. 14, Apr. 2021, doi: 10.3390/ma14081983.

[60] D. Feng et al., "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach," Constr. Build. Mater., Jan. 2020, doi: 10.1016/j.conbuildmat.2019.117000.

[61] N.-D. Hoang, “Machine Learning-Based Estimation of the Compressive Strength of Self-Compacting Concrete: A Multi-Dataset Study,” Mathematics, Oct. 2022, doi: 10.3390/math10203771.
[62] Y.-M. Hong, “Performance Comparison of Machine Learning Models for Concrete Compressive Strength Prediction,” Materials, vol. 17, Apr. 2024, doi: 10.3390/ma17092075.
[63] H.-B. Song et al., “Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms,” Constr. Build. Mater., Nov. 2021, doi: 10.1016/j.conbuildmat.2021.125021.
[64] Y. Zhang, W. Ren, Y. Chen, Y. Mi, J. Lei, and L. Sun, “Predicting the compressive strength of high-performance concrete using an interpretable machine learning model,” Sci. Rep., vol. 14, Nov. 2024, doi: 10.1038/s41598-024-79502-z.
[65] B. Pan, W. Liu, P. Zhou, and D. Wu, “Predicting the Compressive Strength of Recycled Concrete Using Ensemble Learning Model,” IEEE Access, vol. 13, pp. 2958–2969, 2025, doi: 10.1109/access.2024.3519669.
[66] Y. Aydın, C. Çakıroğlu, G. Bekdaş, and Z. Geem, “Explainable Ensemble Learning and Multilayer Perceptron Modeling for Compressive Strength Prediction of Ultra-High-Performance Concrete,” Biomimetics, vol. 9, Sept. 2024, doi: 10.3390/biomimetics9090544.
[67] A. Kashem, R. Karim, S. C. Malo, P. Das, S. D. Datta, and M. Alharthai, “Hybrid data-driven approaches to predicting the compressive strength of ultra-high-performance concrete using SHAP and PDP analyses,” Case Stud. Constr. Mater., Feb. 2024, doi: 10.1016/j.cscm.2024.e02991.
[68] T. Saeheaw, “Interpretable Machine Learning Framework for Non-Destructive Concrete Strength Prediction with Physics-Consistent Feature Analysis,” Buildings, July 2025, doi: 10.3390/buildings15152601.
[69] M. Elshaarawy, M. Alsaadawi, and A. Hamed, “Machine learning and interactive GUI for concrete compressive strength prediction,” Sci. Rep., vol. 14, July 2024, doi: 10.1038/s41598-024-66957-3.
[70] A. Ahmad, W. Ahmad, F. Aslam, and P. Joyklad, “Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques,” Case Stud. Constr. Mater., Dec. 2021, doi: 10.1016/j.cscm.2021.e00840.
[71] A. Kashem and P. Das, “Compressive strength prediction of high-strength concrete using hybrid machine learning approaches by incorporating SHAP analysis,” Asian J. Civ. Eng., vol. 24, pp. 3243–3263, May 2023, doi: 10.1007/s42107-023-00707-0.
[72] I. Ekanayake, D. Meddage, and U. Rathnayake, “A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP),” Case Stud. Constr. Mater., June 2022, doi: 10.1016/j.cscm.2022.e01059.
[73] Z. Zeng et al., “Accurate prediction of concrete compressive strength based on explainable features using deep learning,” Constr. Build. Mater., Apr. 2022, doi: 10.1016/j.conbuildmat.2022.127082.
[74] M. S. N. Tak, Y. Feng, and M. Mahgoub, “Advanced Machine Learning Techniques for Predicting Concrete Compressive Strength,” Infrastructures, Jan. 2025, doi: 10.3390/infrastructures10020026.