International Journal of Mathematics and Statistics

ISSN (Online): 2208-2212
Volume 10 Issue 02 June 2024

DOI: https://doi.org/10.69980/bh15b277

STATISTICAL AND EXPLAINABLE MACHINE LEARNING MODELS FOR PREDICTING CONCRETE COMPRESSIVE STRENGTH A COMPARATIVE PREDICTIVE ANALYSIS

Dr. Julian R. Bennett[1*], Prof. Amélie C. Laurent[2]

[1]Department of Statistical Methodology and Computational Research, Westmoor University, Glasgow, United Kingdom*
[2]Institute of Applied Probability and Statistical Sciences, European Institute of Quantitative Research, Paris, France

Abstract:

This Study has created a comparative statistical and explainable machine-learning model for predicting concrete compressive strength based on the mixture-design variables. This was analyzed using experimental data of 235 concrete mixtures with five predictors, namely cement, fly ash, R Sand, M Sand and water and compressive strength at 3, 7, 28 and 90 days. The compressive strength at 28 days was considered the main prediction target. Decision Tree, Random Forest, Gradient Boosting, Support Vector Regression, and K-Nearest Neighbors models were used in combination with descriptive statistics, Pearson correlation, multicollinearity analysis, and multiple linear regression. R 2, MAE, MSE, RMSE and five-fold cross-validation were used to assess model performance. Findings indicated that nonlinear models were significantly better when compared to Multiple Linear Regression. Support Vector Regression had the best test result, R 2 = 0.9995, MAE = 0.096, MSE = 0.047, and RMSE = 0.218, and the lowest cross-validated RMSE = 0.246. SHAP analysis showed cement and fly ash to be the most significant predictors, followed by R Sand, M Sand and water. The results indicate that explainable machine learning has the potential to deliver very accurate and interpretable concrete-strength predictions, justifying the fast mix-evaluation and engineering decision-making.

Keywords: *Concrete compressive strength; Machine learning; Support Vector Regression; SHAP; Explainable artificial intelligence*

1. Introduction

One of the most significant measures that is used to assess the mechanical properties, durability and structural integrity of concrete is its compressive strength. Precise strength estimation is of paramount importance in the design of structures, quality control, material optimization, and construction safety. A variety of factors affect the strength of concrete; most importantly, the constituents of the mixture and the curing conditions. The statistical evidence has shown that there is information that can be extracted from concrete mixture proportions that can be used to estimate the compressive strength of the concrete, and the relationships are often interdependent and nonlinear [1]. The hydration and strength development also significantly depend upon different curing durations, which are greatly affected by different variations in cementitious content, aggregate composition and water availability [2]. Furthermore, high-performance and self-compacting concretes show that altering constituent proportions can yield complex strength responses that are not easily expressed in simple equations [3,4]. Traditional strength determination relies largely upon laboratory preparations of specimens, curing and destructive compression testing. These methods are reliable, but they are costly, labor-intensive, and are not feasible for rapid prediction of results, especially if multiple mix combinations and curing ages are required [5].

Machine learning has become a very effective method to model complex relationships between concrete mixture variables and compressive strength. Different algorithms can be trained directly from the experimental data to learn a nonlinear relationship without making any strong assumptions about functional relationships. Machine-learning methods have been shown to exhibit good predictive ability for concrete compressive strength, and could potentially perform better than traditional empirical models [6]. For high-strength concrete prediction, extreme learning approaches have also exhibited good performance [7]. Ensemble-based techniques are of great benefit because they fuse several decision structures and enhance generalization. Adaptive boosting has been found to be a powerful tool for predicting concrete strength in estimation and random forest is found to be an effective tool for modelling concrete of high strength and capturing complex correlations among the components of concrete [8,9]. Machine-learning-assisted numerical modelling has also been used for geopolymer concrete, and it is applicable to various concrete types [10]. The performance of the model, however, depends on the model algorithm, data characteristics, data feature structure, and dimensionality [11]. Thus, several algorithms need to be evaluated against each other, in order to find the best predictive model.

As much as they add in predictive value, the use of advanced machine-learning models is sometimes associated with a lack of interpretability. The methods of XAI enable the understanding of how individual variables affect the prediction of the model, enhancing the transparency and practical acceptance [12]. One advantage of SHAP is that it can explain both locally and globally predictive tree models [13]. In concrete-strength prediction, explainability could detect whether the variable, like cement, fly ash, sand, or water, has positive or negative effects on the predicted strength, which is helpful in making more informed decisions in concrete mixture design.

While there are many studies that utilize machine learning methods in concrete compressive strength prediction, most of the previous works are mainly concerned with predicting accuracy with only one of the three aspects such as statistical evaluation, multiple machine-learning algorithms, or explainable models. The potential of machine learning in predicting mechanical properties of concrete has been highlighted, along with the level of variation in the methods used [14]. Predictive models have also shown the benefits of comparing several algorithms for ready-mixed concrete (RMC) [15]. However, more studies are needed to integrate statistical relationships, comparative prediction, and explainable modelling, taking into account strength development of various curing ages.

Thus, the purpose of this study is to statistically analyze the relationships of the concrete mixture components to compressive strength, to build and compare several machine learning regression models and to determine the most important strength predictors by explainable artificial intelligence methods. The main purpose of the comparative framework is to identify the most reliable and meaningful way to predict concrete compressive strength, along with useful insight into the role of key variables in concrete mixtures.

2. Materials and Methods

2.1 Research Design

A quantitative predictive design has been applied in this study with experimental concrete data. The relationship between the concrete mixture components and the compressive strength was assessed using five methods: statistical analysis, comparative machine-learning modelling and explainable artificial intelligence. The analytical model was developed to evaluate the effectiveness of various regression methods in prediction and to find the most important mixture variables.

2.2 Dataset Description

The dataset had five predictor variables: Cement, Fly Ash, R Sand, M Sand and Water with 235 observations. The compressive strength was measured at 3, 7, 28 and 90 days. The compressive strength for 28 days was considered as the main predictive target, and the strength for 3 days, 7 days and 90 days were used as supplements for age-wise analysis. The data exhibited adequate variation in compressive-strength values for predictive modelling. There was, however, no such dependency found between the other two components: Cement and Fly Ash; or R Sand and M Sand, and they were taken into consideration in the statistical interpretation of the data [16].

2.3 Data Preprocessing

Missing data values, duplicate data values, inconsistent data types and possible outliers were checked in the dataset. There were no duplicate observations or missing values found. All variables were coded as normal variables. Extreme observations were evaluated and were kept if they were plausible mixture compositions. For models with scales that were sensitive to scaling (like Support Vector Regression and K-Nearest Neighbors Regression), standardization was used, but for tree-based models, the numerical scale was used as is.

2.4 Statistical Analysis

All the variables were tested for descriptive statistics, such as mean, standard deviation, minimum and maximum values. Associations between the components of the mix and compressive-strength results were analysed by Pearson correlation. Multicollinearity was checked through the Variance Inflation Factors (VIFs) and correlation coefficients. Due to the close relationship between Cement and Fly Ash and between R Sand and M Sand, there is no simultaneous interpretation of the independent effect of the two redundant variables in the traditional regression analysis. The statistical model used for prediction of 28-day compressive strength was Multiple Linear Regression.

2.5 Machine-Learning Models

The regression algorithms tested were Multiple Linear Regression, Decision Tree Regression, Random Forest Regression, Gradient Boosting Regression (also known as XGBoost Regression), Support Vector Regression and K-Nearest Neighbors Regression. The models were chosen to cover linear, tree-based, ensemble, kernel-based and instance-based learning methods. They got a comparative performance, identifying the best-suited model to predict the strength of the concrete.

2.6 Model Training and Validation

This dataset was split into 80% training and 20% testing sets. To minimize over-reliance on one partition and ensure good model reliability, five-fold cross-validation was used on the training data. The nonlinear models were cross-validated to obtain their best hyperparameters. The parameters used for the standardization were estimated only on the training data, without information leakage.

2.7 Model Evaluation

The coefficient of determination (R^2), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) were used to assess model performance. Values of R^2 were higher and error values were lower, which meant that the prediction was better. The best model was selected mainly based on the independent test set results and assisted by cross-validation results. Visual evaluation of the agreement between observed and predicted values was also performed using observed versus predicted plots.

2.8 Explainable Machine Learning

The best model was used for explainable machine-learning analysis. The contribution of Cement, Fly Ash, R Sand, M Sand and Water for prediction of the Compressive strength was quantified using Global feature importance and SHapley Additive exPlanations (SHAP). Both the magnitude and direction of predictor effects were explored using SHAP summary plots. Some mixture variables were perfectly dependent, so their SHAP values were read as a contribution that is caused by them as a whole, rather than independently.

3. Results

3.1 Descriptive Characteristics of the Dataset

The analysis was based on 235 concrete mixtures having all data required for five mixture variables and four different curing ages for their compressive strengths. The averages of the cement and fly ash were presented in Table 1, 168.59 and 126.41, respectively. The average of R Sand was 451.63, M Sand was 349.37 and the average of water was 137.60. The compressive strength was found to increase with the curing age, with a mean of 12.36 at 3 days to 16.79 at 7 days, 26.64 at 28 days and 36.17 at 90 days. The key prediction outcome investigated was the 28-day strength, which varied from 3.47 to 46.11 with a significant standard deviation (SD = 10.40) and offered a sufficient range for predictive modelling.

Table 1. Descriptive Statistics of Concrete Mixture Variables and Compressive Strength

Variable	Mean	SD	Minimum	Maximum
Cement	168.590	70.725	48.780	295.000
Fly Ash	126.410	70.725	0.000	246.220
R Sand	451.628	299.856	0.000	801.000
M Sand	349.372	299.856	0.000	801.000
Water	137.599	14.053	114.000	164.044
3-day strength	12.363	6.093	1.526	26.395
7-day strength	16.786	6.702	3.933	30.616
28-day strength	26.638	10.396	3.471	46.112
90-day strength	36.166	8.866	12.714	50.667

3.2 Relationships Among Concrete Variables

The Pearson correlation analysis showed a high correlation between some of the mixture variables and compressive-strength responses. It can be seen from Figure 1 that the cement was positively correlated with the 28-day compressive strength while a similar negative correlation was observed for fly ash. There was also a negative correlation between water and 28-day strength ($r = -0.760$). The correlation between R Sand and 28-day outcome ($r = 0.105$) and M Sand and 28-day outcome ($r = -0.105$) was comparatively weak. Importantly, cement and fly ash were perfectly negatively correlated ($r = -1.000$), and R Sand and M Sand were also perfectly negatively correlated ($r = -1.000$), thus indicating a definite compositional relationship. The values of the strength for each curing period were highly correlated, especially 7-day strength and 28-day strength ($r = 0.978$) and 28-day strength and 90-day strength ($r = 0.972$).

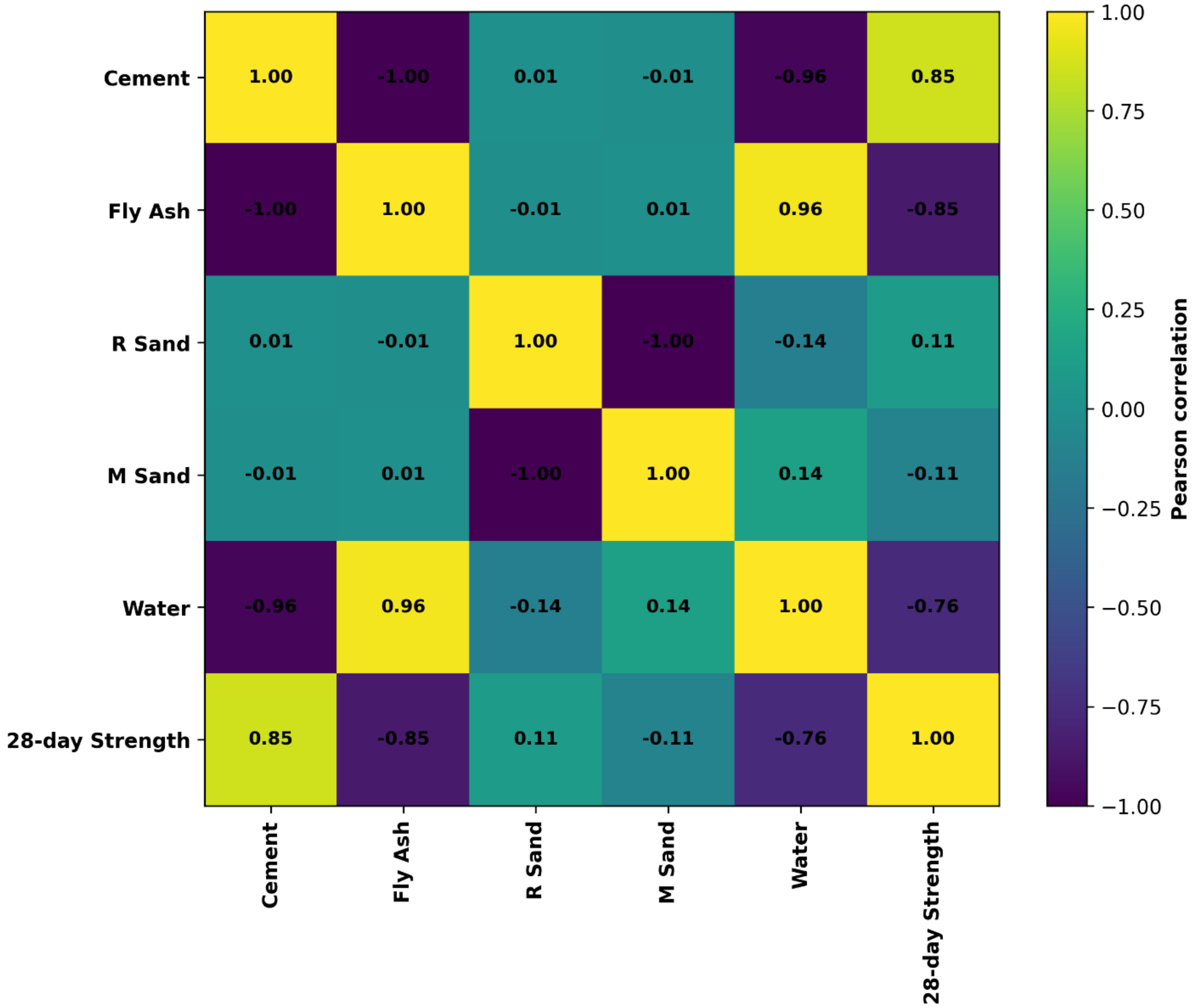

Figure 1. Correlation Matrix of Concrete Mixture Variables and Compressive Strength

3.3 Statistical Regression Results

Due to the exact correlation between Cement and Fly Ash, and between R Sand and M Sand, the conventional regression model was estimated with Cement, R Sand and Water to prevent the condition of perfect singularity. The reduced predictor set was statistically significant with $R^2 = 0.836$, adjusted R2 = 0.834, F = 393.3, and $p < 0.001$, which explained about 83.6% of the variance in 28-day compressive strength. All three of the above-mentioned variables were statistically significant as can be seen in Table 2. The coefficient for cement was positive, $\beta = 0.307$ with p-value < 0.001, and for R Sand, the coefficient was also small but positive, $\beta = 0.009$ with p-value < 0.001. The conditional regression coefficient of water was positive ($\beta = 0.950$, $p < 0.001$) but needs to be interpreted with caution due to the high level of correlation between water and cement. The VIF values were still high for both cement and water, which indicates that there was still multicollinearity.

Table 2. Statistical Regression and Multicollinearity Results for 28-Day Compressive Strength

Predictor	Coefficient	Std. Error	t-value	p-value	VIF
Cement	0.307	0.016	19.129	<0.001	16.793
R Sand	0.009	0.001	8.927	<0.001	1.304
Water	0.950	0.081	11.655	<0.001	17.123

Model $R^2 = 0.836$; Adjusted $R^2 = 0.834$; overall model $p < 0.001$.

3.4 Machine-Learning Model Performance

Significant differences between the traditional linear model and nonlinear machine-learning algorithms were noted. Comparative R 2 and RMSE values were plotted, as shown in Figure 2, which indicates that Support Vector Regression (SVR) had the highest predictive power followed closely by KNN and GB. High predictive accuracy was also obtained for Random Forest and Decision Tree models. The Multiple Linear Regression was, on the other hand, clearly less accurate in terms of R^2 and had significantly more error in the prediction of 28-day strength, which suggests that the pure linear model for the relationship between mixture composition and strength was inadequate.

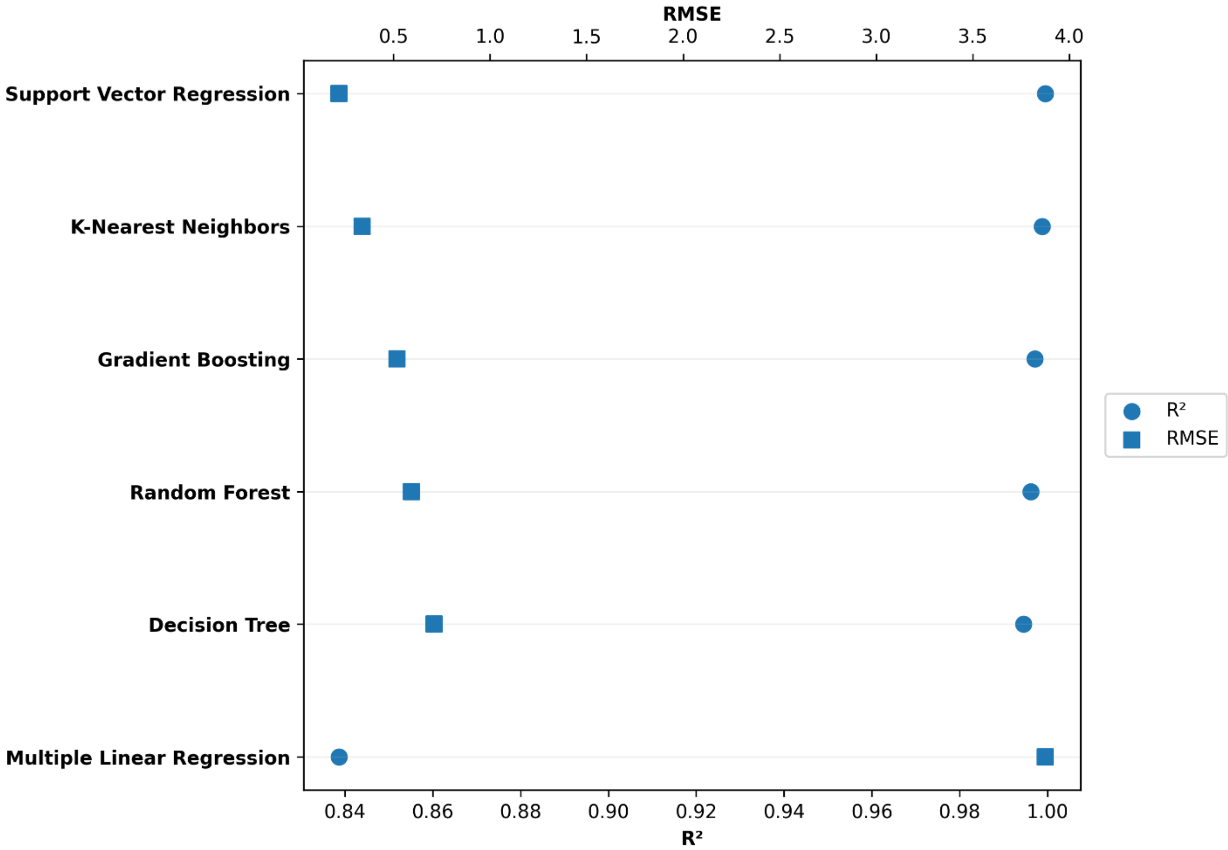

Figure 2. Comparative R^2 and RMSE Performance of the Predictive Models

3.5 Comparative Model Evaluation

Test-set detailed results are shown in Table 3. The best-performing algorithm was found to be Support Vector Regression with $R^2 = 0.9995$, MAE = 0.096, MSE = 0.047 and RMSE = 0.218. K-Nearest Neighbors achieved the second-highest performance ($R^2 = 0.9988$; RMSE = 0.338), followed by Gradient Boosting ($R^2 = 0.9971$; RMSE = 0.518). Random Forest and Decision Tree could also describe above 99% of the variation in the test set. Multiple Linear Regression yielded $R^2 = 0.8387$ and RMSE = 3.876, which shows that there is a significant difference in the predictive power of the nonlinear algorithms. Five-fold cross-validation also confirmed the stability of SVR with the mean cross-validated RMSE of 0.246.

Table 3. Comparative Machine-Learning Performance for 28-Day Compressive Strength

Model	R^2	MAE	MSE	RMSE	5-Fold CV RMSE
Support Vector Regression	0.9995	0.096	0.047	0.218	0.246
K-Nearest Neighbors	0.9988	0.257	0.114	0.338	0.671
Gradient Boosting	0.9971	0.336	0.269	0.518	0.637
Random Forest	0.9962	0.406	0.352	0.593	0.938
Decision Tree	0.9946	0.591	0.506	0.711	1.361
Multiple Linear Regression	0.8387	3.195	15.024	3.876	4.357

3.6 Observed Versus Predicted Compressive Strength

This good performance of the SVR model was followed by a statistical comparison of the observed and predicted values of the 28-day compressive strength. The predicted observations in both low and high strength ranges were tightly grouped near the 1:1 reference line as shown in Figure 3. A very slight difference between the observed and the calculated values was noticed, and the MAE and RMSE of the model were very low. The good matching between the measured values and the predicted values confirmed thus the capability of SVR to model the nonlinear relationships existing in the concrete mixture data.

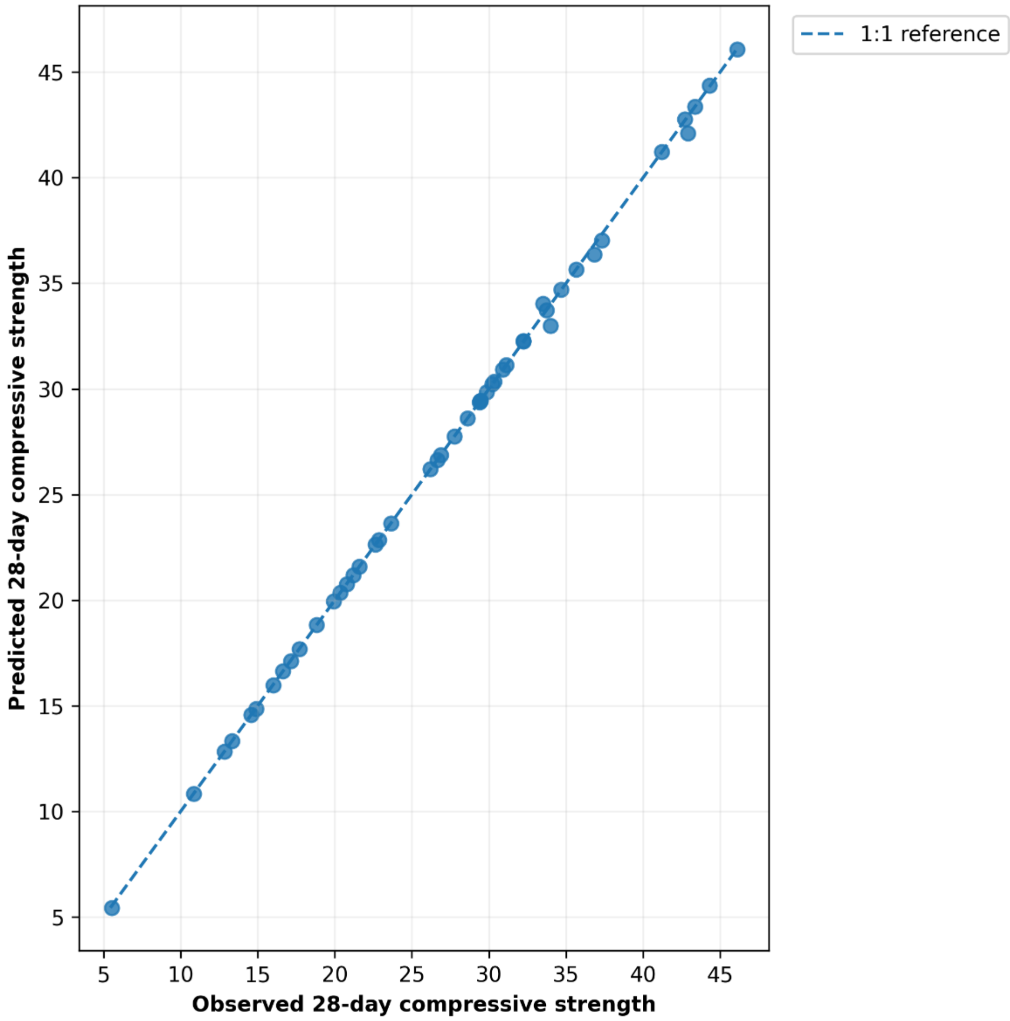

Figure 3. Observed Versus Predicted 28-Day Compressive Strength Using SVR

3.7 Explainability and Feature Importance

SHAP analysis of the most successful SVR model revealed that cementitious composition was the area that provided the most predictive information. Table 4 showed that Cement and Fly Ash contributed about 26.86% of the total mean absolute SHAP importance each. R Sand and M Sand contributed around 18.39%, and Water contributed 9.50. The direction of higher cement values was more in favor of higher predicted strength and the higher fly-ash values demonstrated the opposite directional trend in this dataset. Equally, the positive predictive direction and negative direction are generally observed in R Sand and M Sand respectively. Water had a total negative contribution. The similarity in the values of equal importance in the pairs of Cement-fly ash and R Sand-M Sand indicates that these pairs are complementary to each other and as such, should not be seen as independent effects.

Table 4. SHAP-Based Feature Importance for the Best-Performing SVR Model

Feature	Mean |SHAP|	Relative Importance (%)	Dominant Direction
Fly Ash	3.011	26.86	Negative
Cement	3.011	26.86	Positive
M Sand	2.062	18.39	Negative
R Sand	2.062	18.39	Positive
Water	1.065	9.50	Negative

3.8 SHAP-Based Interpretation

Figure4 shows that the prediction of 28-day strength was influenced by the magnitude and direction of the features and is the SHAP summary distribution of the SVR model. The high values of cement were clustered around the positive SHAP values, suggesting that the cement contributed to an increase in the predicted compressive strength, whereas high values of fly-ash tended to move the predictions towards the lower end. The opposite influence patterns were observed between R Sand and M Sand that were consistent with their complementary formulation. The SHAP values were mostly negative for higher water levels and mostly positive for lower water levels. Overall, the explainability analysis showed that the cementitious composition was the most significant predictor, while the sand composition and water were also significant. The SHAP contribution of the complementary feature pairs was also found to be approximately equal in both directions, further confirming the results found in the correlation analysis and the importance of considering these features together as an interaction, instead of individually as separate effects.

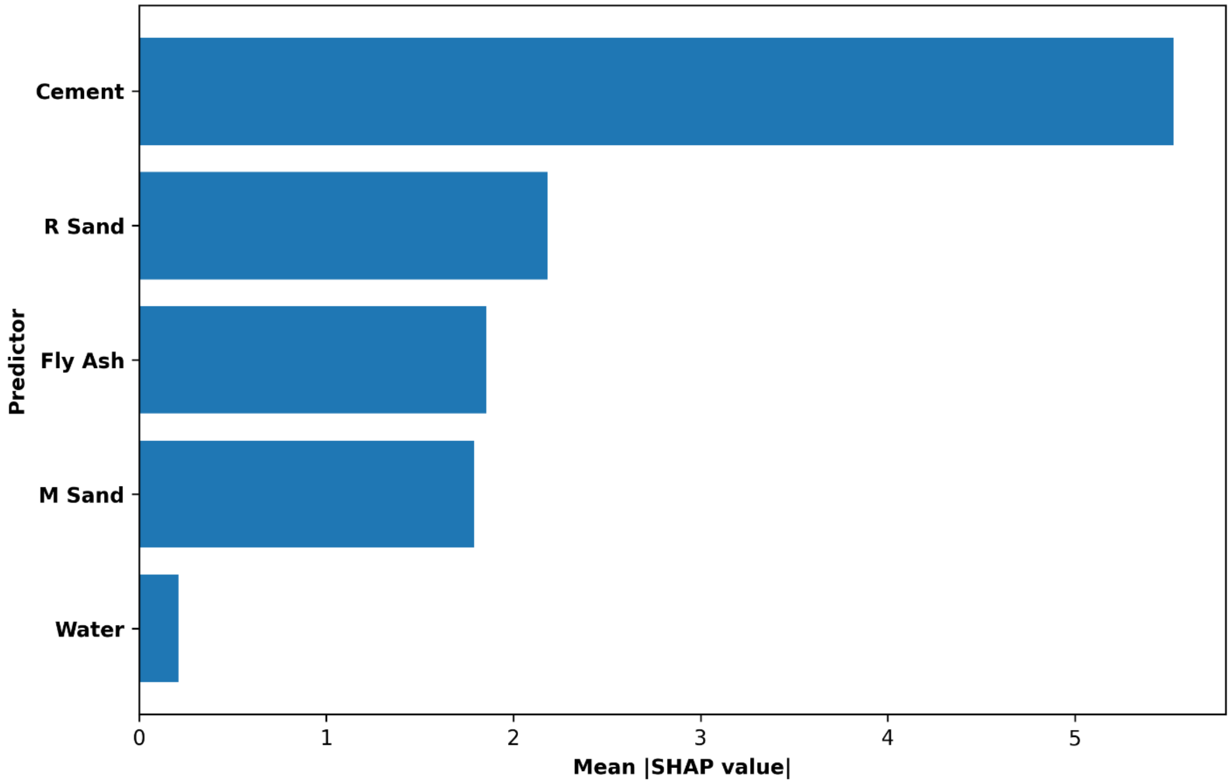

Figure 4. SHAP Summary Plot Showing Predictor Effects on 28-Day Compressive Strength

4. Discussion

This research showed that concrete compressive strength can be successfully predicted with high accuracy by applying nonlinear machine-learning models to the mixture-design variables. Data analysis indicated that the chosen predictors could account for a significant fraction of the variance in the 28-day CS values, and the machine learning models delivered significant performance for the prediction of CS. Support Vector Regression showed the best overall performance with the highest R^2 and lowest prediction errors of all the models examined. The results corroborate the existing evidence of the suitability of the use of computational intelligence to capture complex relations between concrete composition and strength development [17].

The results showed that cementitious composition had the greatest impact in providing predictive information. The relationship with 28-day compressive strength was quite good for cement and inverse in case of fly ash in the present data sets. Both the correlation analysis and the direct relationships showed water negatively correlated with strength, and sand variables less strongly correlated. This is in line with the knowledge that the strength of concrete is determined by the combined effect of the binder content, water availability, aggregate composition, and curing behavior of the mix rather than by individual factors [18]. The high correlation between some of the predictors also warrants careful interpretation of the role of individual predictors.

The nonlinear algorithms turned out to be very successful and showed that the relation between mixture composition and compressive strength was not well captured by a simple linear form as in Multiple Linear Regression. The same benefits with advanced machine-learning algorithms have been reported in the case of ultra-high-performance concrete steel-fiber-reinforced concrete, and preplaced aggregate concrete [19–21]. The high performance of SVR, KNN, Gradient Boosting and Random Forest in the present study demonstrates their capacity to model nonlinear interactions and local response patterns not captured by conventional regression.

SHAP analysis was used to gain further insight into how the predictive model was working on the inside. The mean absolute SHAP values were found to be the highest for cement and fly ash followed by R Sand, M Sand and water. The observed patterns indicated that model explainability can also detect both the strength and the direction of the predictor influence, which can reduce the ‘black-box' problem of advanced algorithms. Explainable feature analysis has also been applied in deep-learning concrete strength prediction and interpretable ensemble learning frameworks [22,23]. At a higher level, explainable AI enables clear understanding of intricate predictive models and fosters trust in data-driven choices [24].

The results align with recent research indicating that both explainable and ensemble models can deliver accurate strength prediction along with interpretability. The use of explainable boosting models has proven effective in the field of concrete compressive-strength estimation, and the integration of AutoML with SHAP has yielded enhanced feature interpretation and automated model selection [25,26]. The same can be said of concrete-strength modelling using XAI, which has been reported to clarify the contribution of mixture variables to prediction results [27]. Moreover, recent interpretable machine-learning research on HPCs underscores the need for predictive accuracy and model transparency [28].

The results show the potential for machine learning models to assist optimization of concrete mixtures, estimation of concrete strength in real time and assist engineering decision-making. Predictions that are accurate help minimize need for repeat destructive testing and the time needed to assess alternate mix designs. Interactive predictive systems have also shown the potential of embedding machine learning in engineering-oriented decision-support interfaces [29]. These techniques can help to optimize efficiency and allow engineers to analyze anticipated strength responses prior to experimental validation.

A number of constraints need to be acknowledged. There were 235 observations included in the dataset, which reduces external generalizability. Multicollinearity was found to be significant due to the perfect complementary nature of Cement–Fly Ash and R Sand–M Sand. There were also a few mixture variables, but no further factors, like aggregate grading, admixtures, temperature, detailed curing conditions. In conclusion, while the predictive accuracy was high, larger and more heterogeneous data sets should be used to validate the findings for broader application.

References:

[1]. Young BA, Hall A, Pilon L, Gupta P, Sant G. Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods. Cement and concrete research. 2019;115:379–88.

[2]. Jayaprakash G, Muthuraj MP. Prediction of compressive strength of various SCC mixes using relevance vector machine. Comput Mater Continua. 2018;54(1):83–102.

[3]. Han Q, Gui C, Xu J, Lacidogna G. A generalized method to predict the compressive strength of high-performance concrete by improved random forest algorithm. Construction and Building Materials. 2019;226:734–42.

[4]. Zhang J, Ma G, Huang Y, Aslani F, Nener B. Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression. Construction and Building Materials. 2019;210:713–9.

[5]. DeRousseau MA, Laftchiev E, Kasprzyk JR, Rajagopalan B, Srubar III WV. A comparison of machine learning methods for predicting the compressive strength of field-placed concrete. Construction and Building Materials. 2019;228:116661.

[6]. Chopra P, Sharma RK, Kumar M, Chopra T. Comparison of Machine Learning Techniques for the Prediction of Compressive Strength of Concrete. Rossi PP, editor. Advances in Civil Engineering. 2018 Jan;2018(1):5481705. doi:10.1155/2018/5481705

[7]. Al-Shamiri AK, Kim JH, Yuan TF, Yoon YS. Modeling the compressive strength of high-strength concrete: An extreme learning approach. Construction and Building Materials. 2019;208:204–19.

[8]. Feng DC, Liu ZT, Wang XD, Chen Y, Chang JQ, Wei DF, et al. Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. Construction and Building Materials. 2020;230:117000.

[9]. Farooq F, Nasir Amin M, Khan K, Rehan Sadiq M, Javed MF, Aslam F, et al. A comparative study of random forest and genetic engineering programming for the prediction of compressive strength of high strength concrete (HSC). Applied Sciences. 2020;10(20):7330.

[10]. Huynh AT, Nguyen QD, Xuan QL, Magee B, Chung T, Tran KT, et al. A machine learning-assisted numerical predictor for compressive strength of geopolymer concrete based on experimental data and sensitivity analysis. Applied Sciences. 2020;10(21):7726.

[11]. Wan Z, Xu Y, Šavija B. On the use of machine learning models for prediction of compressive strength of concrete: influence of dimensionality reduction on the model performance. Materials. 2021;14(4):713.

[12]. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information fusion. 2020;58:82–115.

[13]. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. Nature machine intelligence. 2020;2(1):56–67.

[14]. Chaabene WB, Flah M, Nehdi ML. Machine learning prediction of mechanical properties of concrete: Critical review. Construction and Building Materials. 2020;260:119889.

[15]. Xu J, Zhou L, He G, Ji X, Dai Y, Dang Y. Comprehensive machine learning-based model for predicting compressive strength of ready-mix concrete. Materials. 2021;14(5):1068.

[16]. Oyebisi S. Raw dataset for evaluating the effects of M sand and fly ash on concrete strength using experimental and machine learning approaches [Internet]. 2024 Jul 26. doi:10.5281/zenodo.13000484

[17]. Nunez I, Marani A, Flah M, Nehdi ML. Estimating compressive strength of modern concrete mixtures using computational intelligence: A systematic review. Construction and Building Materials. 2021;310:125279.

[18]. Nguyen MH, Ly HB. Development of machine learning methods to predict the compressive strength of fiber-reinforced self-compacting concrete and sensitivity analysis. Construction and Building Materials. 2023;367:130339.

[19]. Nguyen NH, Abellán-García J, Lee S, Garcia-Castano E, Vo TP. Efficient estimating compressive strength of ultra-high performance concrete using XGBoost model. Journal of Building Engineering. 2022;52:104302.

[20]. Pakzad SS, Roshan N, Ghalehnovi M. Comparison of various machine learning algorithms used for compressive strength prediction of steel fiber-reinforced concrete. Scientific Reports. 2023;13(1):3646.

[21]. Qureshi HJ, Alyami M, Nawaz R, Hakeem IY, Aslam F, Iftikhar B, et al. Prediction of compressive strength of two-stage (preplaced aggregate) concrete using gene expression programming and random forest. Case Studies in Construction Materials. 2023;19:e02581.
[22]. Zeng Z, Zhu Z, Yao W, Wang Z, Wang C, Wei Y, et al. Accurate prediction of concrete compressive strength based on explainable features using deep learning. Construction and Building Materials. 2022;329:127082.
[23]. Jia JF, Chen XZ, Bai YL, Li YL, Wang ZH. An interpretable ensemble learning method to predict the compressive strength of concrete. In: Structures [Internet]. Elsevier; 2022 [cited 2026 Sep 12]. p. 201–13. Available from: https://www.sciencedirect.com/science/article/pii/S2352012422009559
[24]. Li XH, Cao CC, Shi Y, Bai W, Gao H, Qiu L, et al. A survey of data-driven and knowledge-aware explainable AI. IEEE Transactions on Knowledge and Data Engineering. 2020;34(1):29–49.
[25]. Liu G, Sun B. Concrete compressive strength prediction using an explainable boosting machine model. Case Studies in Construction Materials. 2023;18:e01845.
[26]. Sun B, Cui W, Liu G, Zhou B, Zhao W. A hybrid strategy of AutoML and SHAP for automated and explainable concrete strength prediction. Case Studies in Construction Materials. 2023;19:e02405.
[27]. Elhishi S, Elashry AM, El-Metwally S. Unboxing machine learning models for concrete strength prediction using XAI. Scientific Reports. 2023;13(1):19892.
[28]. Zhang Y, Ren W, Chen Y, Mi Y, Lei J, Sun L. Predicting the compressive strength of high-performance concrete using an interpretable machine learning model. Scientific Reports. 2024;14(1):28346.
[29]. Elshaarawy MK, Alsaadawi MM, Hamed AK. Machine learning and interactive GUI for concrete compressive strength prediction. Scientific Reports. 2024;14(1):16694.