Journal of ICT, Design, Engineering and Technological Science (JITDETS)
VOL. 9, NO. 1, pp. 33-43, 2025
DOI: https://doi.org/10.33150/JITDETS-9.1.5

ORIGINAL CONTRIBUTION

Machine Learning–Based Prediction and Interpretability Analysis of Ultra-High-Performance Concrete Compressive Strength Using Random Forest

Imran Ali Channa[1*], Muhammad Khisrow Khan[2], Saad Hanif[3], Abdul Wahab[4]

[1] Department of Civil Engineering, Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan

[2] AFOHS New Malir, Karachi, Pakistan

[3] Military College of Engineering, National University of Sciences & Technology Risalpur, Pakistan

[4] Department of Civil Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Swabi, Pakistan

***Abstract*— Ultra-High-Performance Concrete (UHPC) is a considerably advanced cementitious concrete with great characteristics of strength and durability, but the compressive strength is highly dependent on the multi-faceted interplay between mixture proportions and curing conditions. These interactions are nonlinear and multivariate, making it difficult to accurately estimate the UHPC compressive strength using the previous experimental and empirical methods. In the paper, a Random Forest (RF) regression model has been constructed to estimate UHPC compressive strength based on a large-scale dataset of 810 samples and 13 predictors (material composition and curing parameters). Multiple statistical measures were strictly used to evaluate the performance of the model, such as R^2, RMSE, MAE, MAPE, and CVRMSE, as well as 10-fold cross-validation to evaluate stability and ability to generalize. The optimized RF model had a high predictive accuracy with a value of 0.96 on the testing set and small values of errors, which showed high robustness and consistency in diverse segmentations of data. Hyperparameter tuning also improved the model performance by finding a balance between model complexity and generalization. SHAP (Shapley Additive Explanations) analysis was used to enhance the transparency and interpretability of the models, to measure the contribution of the individual input feature to the compressive strength predictions. The findings demonstrated that curing age, fibre, silica fume, and dosage of superplasticizer were the most significant parameters that controlled the strength development of UHPC. The suggested modeling framework reveals the efficiency of bringing ensemble machine learning along with explainable artificial intelligence methods to provide accurate, reliable, and interpretable predictions of UHPC compressive strength, which creates a useful instrument in the process of mix design optimization and performance evaluation.**

***Index Terms*— Ultra-High-performance concrete, Random forest, Compressive strength, Machine learning, SHAP analysis, Hyperparameter tuning, Model interpretability**

Received: 22 January 2025; **Accepted:** 9 April 2025; **Published:** 23 June 2025

I. INTRODUCTION

Ultra-High-Performance Concrete (UHPC) has become a superior cementitious material of the next generation as it is characterized by exceptional mechanical performance, high durability, and increased microstructural properties in comparison with standard concrete. Usually, UHPC has a very high compressive strength, low permeability, and high resistance to environmental degradation, which is why it is used in the structural field that is highly demanding, with high-span bridges, high-rise buildings, precasts, and protective structures [1, 2, 3]. These enhanced characteristics are attained by optimum packing of the particles, minimized water-to-binder ratios, addition of supplementary cementitious materials, high-range water reducer, and fibre reinforcement. Nevertheless, design and performance prediction of UHPC is much more difficult than conventional concrete due to its complex structure and a combination of various constituents [4].

One of the most important performance indicators of UHPC is compressive strength, and it has a direct effect on structural capacity, safety, and serviceability. Compressive strength of UHPC depends on a great number of different factors, which are interrelated, and they include cement content, silica fume, slag, fly ash, quartz powder, limestone powder, nano-silica, aggregate grading, water content, fibre dosage, superplasticizer content, curing temperature, and curing age. Changes in either of these parameters may result in significant alterations in the strength development [5, 6]. Conventional experimental methods of establishing compressive strength are lengthy, resource-demanding, and, in many cases, impractical in the event of investigating a large design space of mixture proportions and curing conditions. Consequently, there has been an increasing demand for predictive models, which are reliable and can predict UHPC compressive strength with a minimum amount of experimentation and cost [6, 7].

Over the last few years, data-based and machine learning methods have received more interest in the field of concrete materials due to the possibility of modeling complex nonlinear relationships without making

*Corresponding author: Imran Ali Channa
†Email: engrimranalichanna@gmail.com

explicit assumptions about the underlying physical processes [8, 9, 10, 11, 12]. Ensemble-based approaches are also promising in predicting concrete properties among other machine learning methods. Such techniques group together several weak learners and enhance prediction accuracy, strength, and ability to generalise. The Random Forest algorithm, especially, has been well acknowledged to perform well on regression tasks with high-dimensional and nonlinear data sets. It is particularly appropriate because it can be used to model UHPC, whose effects of individual components are not independent but rather interdependent [7, 13].

Although the use of machine learning is increasingly being applied in actual studies, there are several challenges. Among these is the possibility of overfitting, particularly when models are trained with a small or highly varying set of data. The other critical problem is that many machine learning models cannot be interpreted, and this fact can reduce their applicability in engineering practice. Engineers and researchers need to have accurate predictions but understand the effect input variables have on the output. With no interpretability, it is hard to justify that the model projections have physical meaning, and to exploit the model to optimize the mix design. As such, explainable artificial intelligence methods should be incorporated into predictive modeling systems to promote the feasibility of the practical implementation of machine learning in UHPC studies [14, 15, 16, 17, 18].

Moreover, model testing and verification are important in determining the predictive tool's reliability. Using one performance measure or a single division of the data may give false results regarding the performance of the model. The overall evaluation of several statistical measures, including the coefficient of determination, root mean square error, mean absolute error, and percentage-based error measures, provides a more holistic view of its predictive power and strength. In addition, methods like K-fold cross-validation enable the stability and generalizability of models to be evaluated with varying data divisions and minimize the effect of chance on training and testing divisions [19, 20, 21].

Another important parameter of machine learning model development that will have a direct impact on predictive performance is hyperparameter tuning. Random Forest models' parameters, like the number of trees, feature sampling strategy, and tree depth, control the balance between the complexity and generalization of the model. These parameters can be productively tuned to generate substantial improvement in prediction accuracy and error reduction, even starting at a high baseline model performance. The learning of the impact of hyperparameters also gives an idea of how the algorithm behaves and will be able to guarantee that the chosen model setup is efficient and robust [22, 23, 24].

In addition to the accuracy of prediction, the interpretability of model results has also gained real significance in applications in materials engineering. SHAP (Shapley Additive Explanations) and analogous approaches (techniques based on Shapley values) provide a highly effective framework to quantify the contribution in model predictions of each input variable [25, 26, 27, 28, 29, 30, 31, 32]. SHAP analysis allows global interpretation, ranking features by their global significance, and local interpretation, explaining predictive individual features. Such techniques are not only useful in improving transparency but also in revealing the most significant mixture parameters, which are useful in material optimization and sample design [2, 33, 34, 35, 36, 37, 38, 39].

In this context, the current paper aims at the prediction of the compressive strength of UHPC based on a strong, consistent, and interpretable model implemented in a random forest with an encompassing set of mix ratios and the curing conditions. Using a big dataset of 810 data points and 13 input features, the research will be able to capture the underlying variability and complexity of UHPC behavior. The modeling paradigm focuses on strict performance analysis with several error measures, inference of hyperparameters with systematic search, stability analysis with K-fold cross-validation, and interpretation analysis with SHAP. In this combined method, the paper aims at showing the promise of machine learning as not only a predictive agent but also as a knowledge-discovery process that will further improve the insight into the UHPC strength development and will help create a more effective and informed mix design strategy.

II. METHODOLOGY

A. Data characteristics

The dataset of Ultra-High-Performance Concrete (UHPC) employed in the current research consists of 810 data points of various designs of concrete mix and curing conditions. There are 14 numerical variables in the dataset, with Compressive Strength (CS) being the target, and the rest of the variables in the dataset are input features. These input variables are the material composition and the curing environment such as Cement Content (C), Slag (S), Silica Fume (SF), Limestone Powder (LP), Quartz Powder (QP), Fly Ash (FA), Nano-Silica (NS), Aggregate Content (A), Water Content (W), Fiber Content (Fi), Superplasticizer Dosage (SP), Curing Temperature (T), and Curing Age (Age). These variables can be combined to capture the complexity of UHPC mix designs and their effect on mechanical performance.

The statistical results of the dataset presented in terms of the mean and standard deviation give a good idea of the distribution and variance of each of the features. Cement and aggregate materials are the ones that have the highest mean value, with a demonstration of their dominance in UHPC mixtures. Aggregate content exhibits a considerably large standard deviation, indicating a large variation among various mix designs. The content of water is moderate in variation, which is appropriate to the demand of strict control in the UHPC production. Substitutive cementitious materials, including silica fume, slag, fly ash, quartz powder, and limestone powder, exhibit relatively small mean values but prominent standard deviations, which indicate that people used varied optimization strategies in the dataset. Nano-silica has a very low mean and standard deviation, which implies that this compound is used in small and controlled amounts. The process of age curing has a great diversity and a large variance, which emphasizes its acuity in strength formation. The values of the compressive strength have a relatively high mean with a moderate variability, which proves the high-performance type of UHPC and the effect of mix proportions and curing conditions.

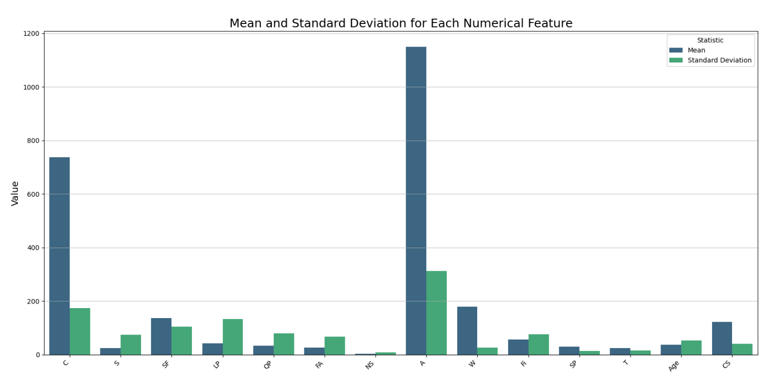
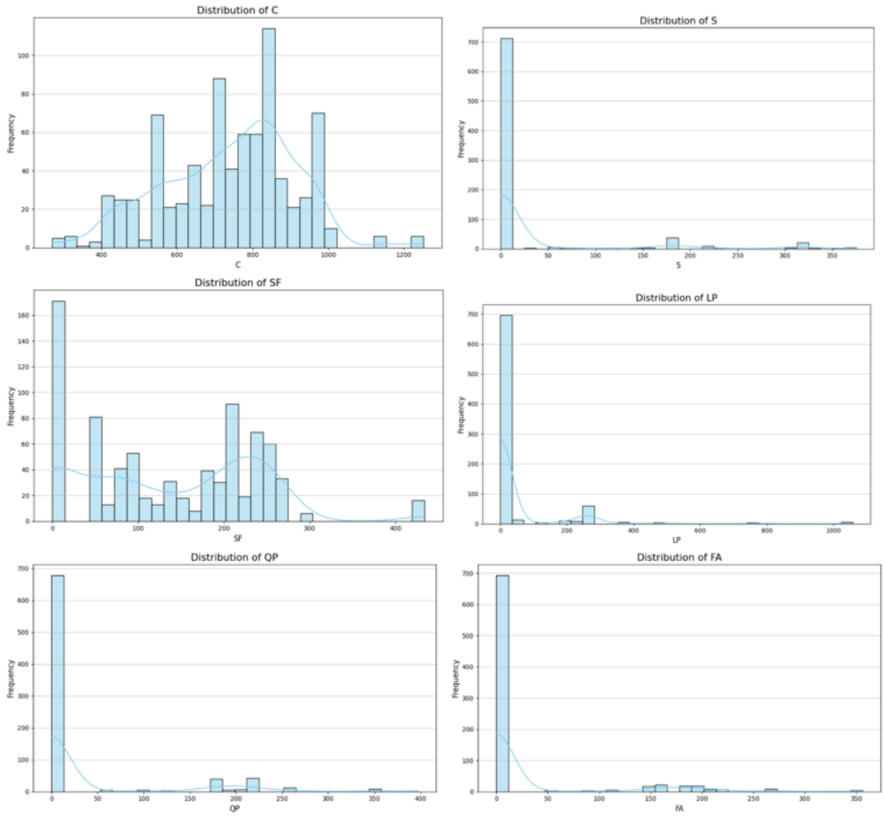
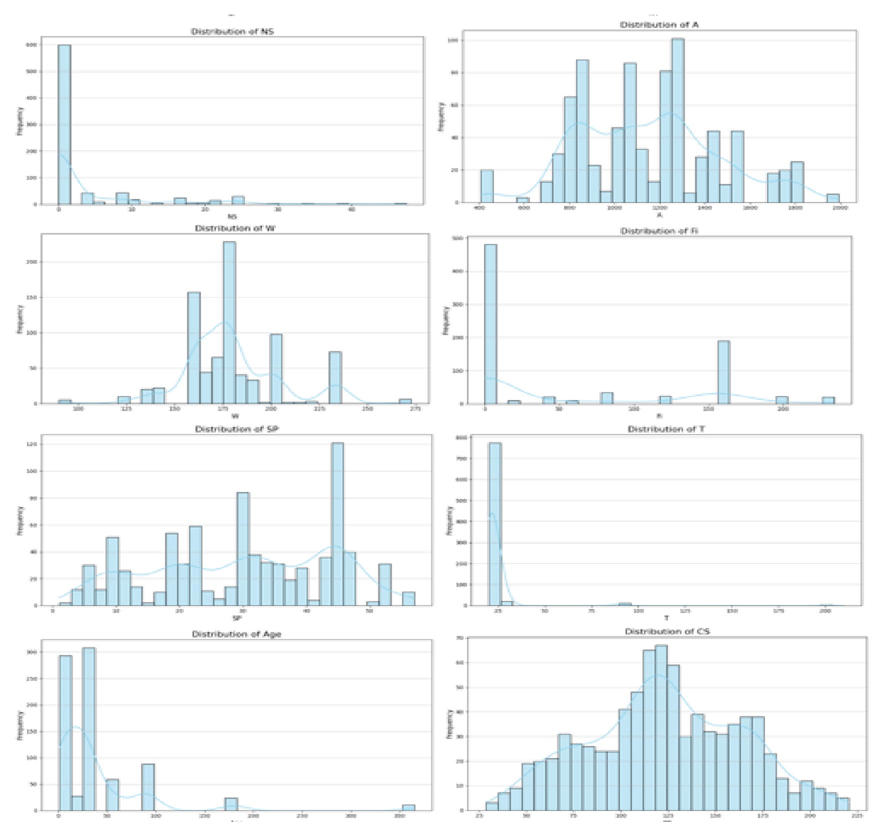


Fig. 1. Mean and standard deviation of numerical features in the UHPC dataset, illustrating the central tendency and variability of mixture proportions, curing parameters, and compressive strength



(a)



(b)

Fig. 2. Frequency distribution diagrams of individual UHPC features, illustrating distribution patterns, variability, and skewness of mixture components and curing parameters

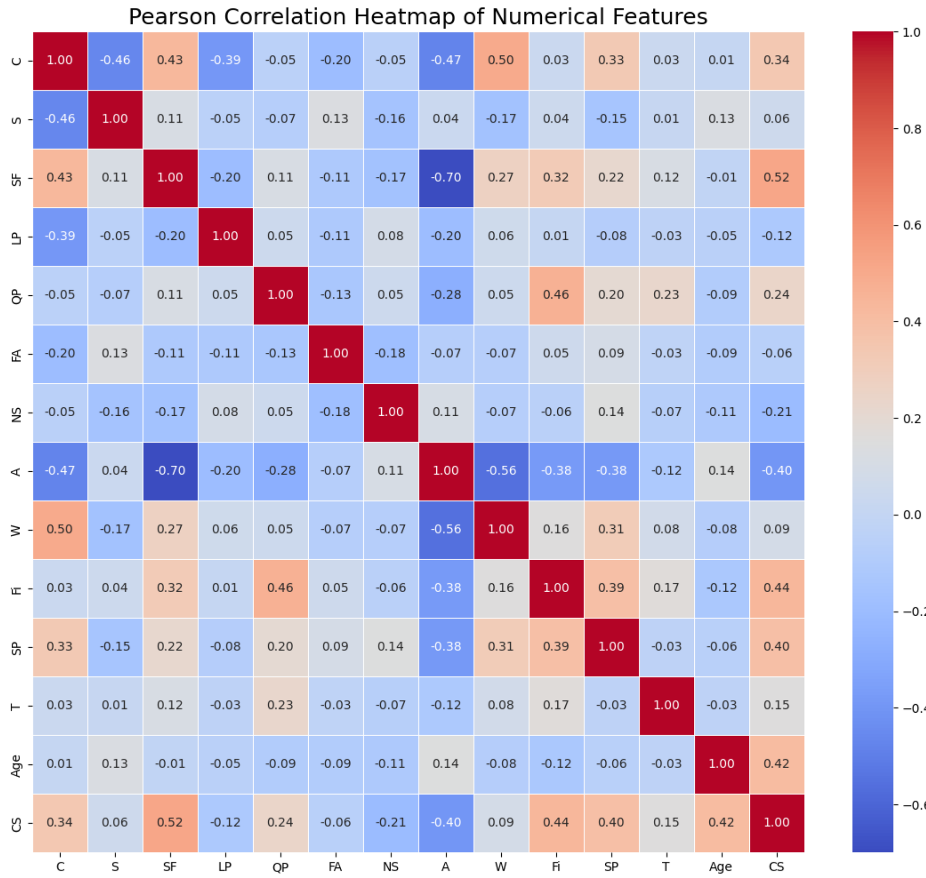


Fig. 3. Correlation heatmap of UHPC dataset features, showing the strength and direction of relationships between input variables and compressive strength

B. Model development

The Random Forest (RF) model was created to estimate the compressive strength of Ultra-High-Performance Concrete (UHPC) when 13 input variables (comprising the mixture composition and curing conditions) are used as the input variables. Before model training, the dataset with 810 samples was reviewed to be consistent and complete, and all variables were maintained in their numeric form. The compressive strength had been chosen as the output variable with the cement, slag, silica fume, limestone powder, quartz powder, fly ash, nano-silica, aggregate, water content, fibre dosage, superplasticizer content, curing temperature, and curing age as the predictive features. The dataset was also stratified and split into training and testing to guarantee an objective model analysis. Random Forest algorithm was chosen because it has a high ability to represent nonlinear relationship and interactions among multiple variables of input that are complex. The model was built as a collection of decision trees, with each tree being trained on a bootstrap of the training data and a random sample of features at each split. This method enhances generalization performance as well as lowers the overfitting in comparison to single-tree models. The number of trees and the tree depth were also key hyperparameters to be optimized to obtain a compromise between the accuracy of prediction and computational efficiency. The model performance in the testing dataset was assessed based on conventional regression measures, which validated the usefulness of the RF model in the determination of the effect of UHPC mix proportions and curing parameters on compressive strength. The findings reveal that the predictions made by the Random Forest model are reliable and robust, and thus, it is applicable in the strength estimation and optimization of the mix design of UHPC.

C. Model assessment

To evaluate the effectiveness of the Random Forest model to predict the compressive strength of Ultra-High-Performance Concrete (UHPC), a number of statistical measures, including the coefficient of determination (R2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), mean absolute percentage error (MAPE), and coefficient of variation of root Mean Square Error (CVRMSE) were examined. There are several measures that can be applied to ensure the sophisticated testing of model accuracy, reliability, and robustness. The percentage of variance of compressive strength explained by the model is determined with the use of the coefficient of determination (R2). The increment in the value of R2 implies that there is a high degree of consistency between the extent of prediction and the experiment, and this means that the model can emulate the underlying connections between the factors of the mix of the UHPC and the development of strengths. RMSE is a measure of the square root of the mean squared error between predicted and observed values, with the large errors having more weight, and therefore focusing more on prediction reliability over the entire range of the data. MAE is used to show the average magnitude of prediction errors regardless of their direction and gives a clear measure of the standard deviation of the compressive strength values that are predicted and the actual ones. MAPE represents the error of prediction in percentage, which makes the model performance intuitively represented against the actual values of strength. Lastly, CVRMSE normalizes RMSE to the mean of the observed data, and this allows comparing the performance of a model without being dependent on the magnitude of the data. Taken together, these measures make the predictive ability of the Random Forest model and its applicability to consistent and accurate prediction of UHPC compressive strengths valid.

III. RESULTS

A. RF analysis

The RF model is shown to have high predictive performance in the training, validation, and testing datasets as indicated by the measured performance indicators. The coefficient of determination (R2 = 0.99) of the model on the training set is very high, which means that the model considers almost all variability in UHPC compressive strength. The low RMSE (4.36 MPa) and MAE (2.89 MPa) also confirm that the prediction errors on the training data are insignificant, with the low MAPE value (2.51%) also demonstrating excellent relative accuracy. The CVRMSE of 3.56% indicates that the errors in prediction are narrowly dispersed as compared to the mean compressive strength, indicating that the model has been able to acquire the underlying patterns in the data. In the case of the validation set, the performance declines, which is anticipated when the model is tested on unknown data. The high level of explanatory power is still evident because the R2 value is 0.94, which means that the RF model can be generalized to the training dataset. The growth in RMSE (9.20 MPa) and MAE (6.90 MPa) represents an increased uncertainty in prediction as opposed to the training process, but all these values are acceptable to predict UHPC strength. The MAPE of 5.67% indicates that, on average, the errors in prediction are quite small in relation to the real values of compressive strength. The CVRMSE at 7.20% indicates greater variability in the errors, but still lets one think of stable predictive behavior. The stability of the RF model is also confirmed by the results of the testing set. The R2 is 0.96, which means that the model has a solid capacity to explain the variance of compressive strength of entirely unobservable data. RMSE (8.38 MPa) and MAE (5.99 MPa) are somewhat smaller than those of the validation set, implying the predictive consistency and minimum overfitting. The MAPE value of 5.37% supports the model's reliability regarding the relative error, as it shows that the deviations in predictions are negligible compared to the real values of the strengths. The CVRMSE of 6.89% justifies that the values of prediction errors are quite modest, being divided by the mean compressive strength. The comparison between the datasets shows that the RF model is well balanced, with the differences between training and testing performance being minimal. The small change in error measures between the training, validation, and testing implies that the ensemble method of learning in the case of a random forest would be an efficient way to reduce overfitting, along with maintaining the excellent predictive ability. All in all, the consistency of R2, RMSE, MAE, MAPE, and CVRMSE across all data sets is indicative of the good generalization capability of the model and its applicability in modelling the complex and nonlinear relationship information of UHPC compressive strength.

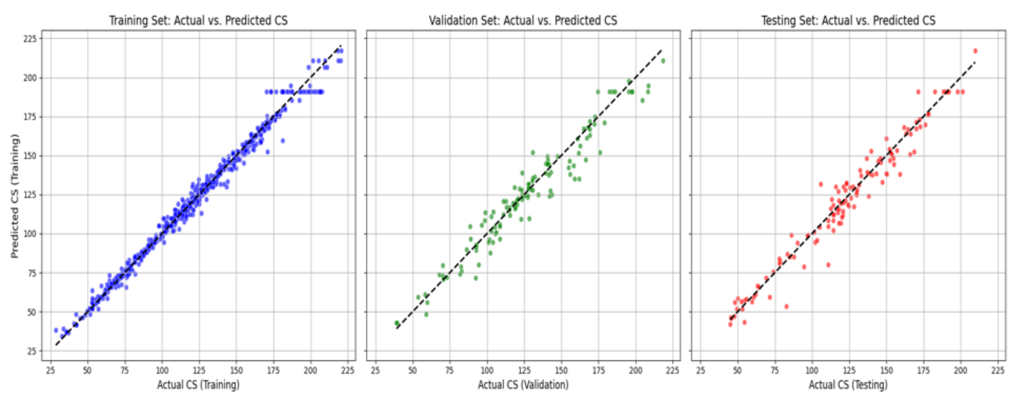


Fig. 4. Scatter plots comparing experimental and predicted compressive strength values for the training, validation, and testing datasets using the Random Forest model

B. Hyperparameter analysis

The hyperparameter optimization of the Random Forest (RF) regressor is an excellent demonstration of the boosting performance and better model stability of the Compressive Strength (CS) of the Ultra-High-Performance Concrete predictor. Systematic tuning was used to determine the optimal combination of hyperparameters, with 400 estimators, a maximum feature ratio of 0.8, and a maximum tree depth of 20. This model gave a high cross-validation R2 of 0.9506, which is good evidence that this model successfully explains the nonlinear correlation between UHPC mixture constituents, curing parameters, and compressive strength. The tuned RF model was found to predict excellently when tested on the hidden test set with a value of R2 of 0.96. The error measures, namely, RMSE= 7.96 MPa, MAE= 5.72 MPa, MAPE= 5.06%, and CVRMSE= 6.54%, support the reliability and precision of the model even more. The values show that there are quite small differences in the expected and experimental values of compressive strength, both in the absolute and relative senses. As the cross-validation and test results are consistent, the tuned model has a high generalization ability. On comparison with the untuned Random Forest model, there are slight but steady increases in all the evaluation measures. Although there was no change in the high value of the R2, the decreases in the RMSE, MAE, MAPE, and CVRMSE show that the model provides a greater degree of predictive accuracy and less variability in errors. This, alongside other improvements, shows that hyperparameter optimization, when viewed as a minor change in terms of its numerical effects, does add value to the process of hyperparameter optimization in the sense of making the model performance better and more stable on a variety of data splits. The hyperparameter effect analysis provides a good understanding of the behaviour of the RF model. The use of more estimators tended to increase the R2 score, but the gains in performance began to flatten after around 300 trees. This tendency explains why a 400 estimator was chosen as a compromise between prediction performance and computing time. The evaluation of the max-features parameter indicates that employing a subset of features at each split, especially 0.8 or square-root choice, helps to improve the model performance by decreasing correlation between trees and increasing the diversity of the ensemble. According to the max depth analysis, there is a moderate optimum depth of trees, which results in the best predictive power as it captures complex relations, and excessively deep trees lead to diminishing marginal returns and overfitting. It was also discovered that a depth of 20 was a good balance between model complexity and generalization.

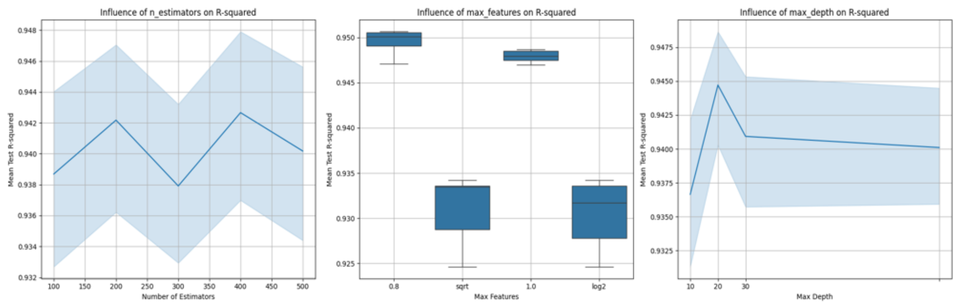


Fig. 5. Influence of Random Forest hyperparameters (n_estimators, max_features, and max_depth) on cross-validated R-squared values for compressive strength prediction

C. Residual assessment

The residual analysis gives valuable information about predictive behavior and the reliability of the Random Forest model on the training, validation, and testing sets. The plot of the difference between actual and predicted compressive strength values, known as the residuals, is then plotted against the predicted compressive strength to assess the bias of the model, distribution of error, and consistency. In the case of the training set, the residual values are concentrated around the zero line, which means that the model estimates are near the observed values. The fact that the spread of the residues is relatively small throughout the whole range of predicted compressive strength implies that the model has been able to learn the underlying relations without any systematic bias. Despite some larger residuals observed at higher levels of strength, they are few and do not have a massive impact on the overall performance of the model. The distribution of the residuals in the validation set is also symmetrically balanced about the mean of the data, which proves the model to be bias-free in its predictions on unknown data used in tuning. The dispersion of the residuals is slightly higher than in the training set, as would be anticipated by the fact that these samples were not directly learned. Nevertheless, the patterns and trends are not seen clearly, which implies that there is no systematic error in prediction but a random one. Lack of funnel-shaped patterns is an indication that there is homoscedastic behavior, and prediction errors are constant throughout the compressive strength range. The residuals of the test set also prove the strength and the ability of the Random Forest model to generalize. The residuals are distributed around zero efficiently with no visible curve or clustering, which proves that the model does not predict the compressive strength either under or over. Some isolated outliers do exist, although their effect is small, especially at extreme predicted values. The distribution of the residues in general is similar between the testing set and the validation set, which supports the model performance consistency in various data partitions. Altogether, the residual plots suggest that the Random Forest model makes consistent and impartial forecasts with erratic distribution. The absence of systematic patterns and the equal distribution of residuals between all datasets verify the appropriateness of the model in predicting UHPC compressive strength and demonstrate the credibility of the measures of reported performances.

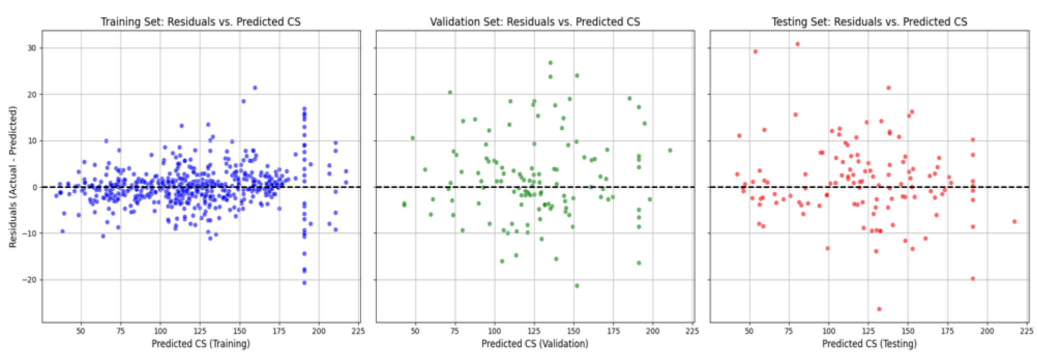


Fig. 6. Residuals versus predicted compressive strength for the training, validation, and testing datasets using the Random Forest model

D. K-fold cross-validation

K-fold cross-validation analysis gives an in-depth analysis of the cross-validation and stability of the Random Forest Regressor model applied in the prediction of the compressive strength of Ultra-High-Performance Concrete. A 10-fold cross-validation plan was used, where each group of the data was utilized as a validation set on a single occasion, with the rest as trainers. This reduces the effects of bias in data partitioning and provides a credible estimate of model performance on alternative data partitions. The findings show that there is a high predictive accuracy in all folds. The R2 values are between about 0.94 and 0.98, which proves that the model always explains such a large percentage of variance in compressive strength despite the partitioning of the data. The high level of stability of the model is reflected in the value of mean R2 equal to 0.96 with a small standard deviation of 0.01. The fact that R2 differs by such a small margin between the folds implies that the Random Forest model is not sensitive to specific sets of data and its predictive qualities are robust. In the same way, the values of the RMSE within the folds are rather low and equal (between 6.41 MPa and 8.34 MPa). The standard deviation of 0.60 MPa and a mean RMSE of 7.59 MPa suggest that the error in prediction is low, and it is consistent across various validation conditions. The values of RMSE were lower in some folds, and this indicates that the model works well, especially with certain data distributions, and slightly higher values in other folds are still in an acceptable range in the prediction of UHPC compressive strength.

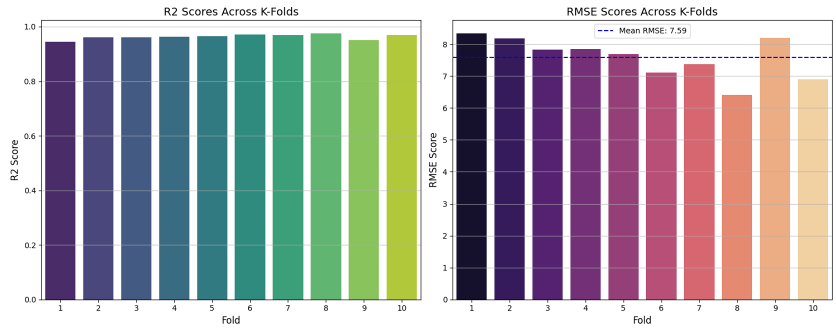


Fig. 7. K-fold cross-validation results showing R^2 and RMSE scores across 10 folds for the Random Forest model

E. Shapely additive analysis

The SHAP (Shapley Additive Explanations) analysis provides an interpretable and transparent understanding of how individual input features influence the Random Forest model's predictions of UHPC Compressive Strength (CS). By quantifying both the magnitude and direction of feature contributions, SHAP helps reveal the relative importance of mixture proportions and curing parameters and explains how changes in these variables affect predicted strength. The importance plot of SHAP, which is used to demonstrate the mean values of the SHAP, shows that the curing age has the greatest impact on the outcome by far. This shows that the overall influence of age on compressive strength prediction is the strongest, which is consistent with the basic hydration and strength development processes of UHPC. Fiber content (Fi) is the next feature that becomes the most significant and underlines the fundamental necessity to improve the load transfer, crack bridging, and performance after cracking. Such relevance is also demonstrated by Silica Fume (SF) and Superplasticizer (SP), which are considered to play a role in densifying the microstructure, increasing the workability, and strengthening of concrete. Aggregate content (A) and Cement Content (C) have moderate effects, which prove their structural and binding functions in UHPC mixtures. Conversely, the variables, water content (W), Curing Temperature (T), Limestone Powder (LP), Nano-Silica (NS), Quartz Powder (QP), Fly Ash (FA), and Slag (S) possess relatively lower SHAP values, indicating only slightly limited contribution, nevertheless a significant one within the studied range of data. The SHAP summary plot further depicts the direction of the feature effects as well as the distribution. In curing age, the values that are high (represented in red) are mainly positive SHAP values, which means they have a strong positive contribution to compressive strength, whereas the low ages have a negative contribution to compressive strength. The same positive trend can be noted with respect to fiber content, silica fume, and superplasticizer, whereby higher doses tend to raise predicted strength. There is a mixed effect with regard to aggregate content; both inadequate and excess aggregate proportions may have an adverse effect on strength. The effect of cement content is moderate and positive, and the effect of increased water content is slight and decreases compressive strength, which is in line with the effects of the water-to-binder ratios. Its other auxiliary materials depict that SHAP values are concentrated around zero, which portrays the minimal effect in directions in the range observed.

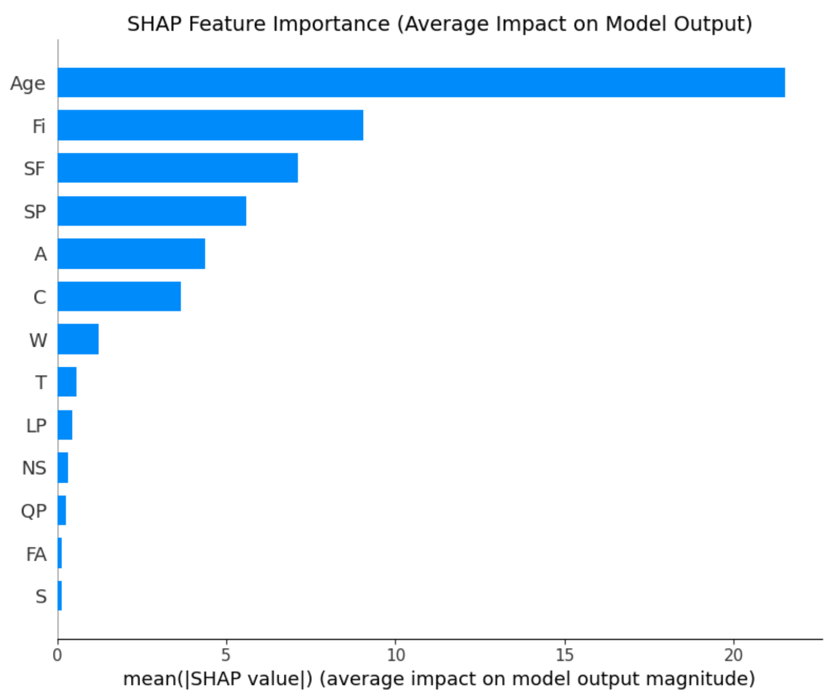


Fig. 8. SHAP feature importance plot showing the average absolute impact of each input variable on UHPC compressive strength predictions

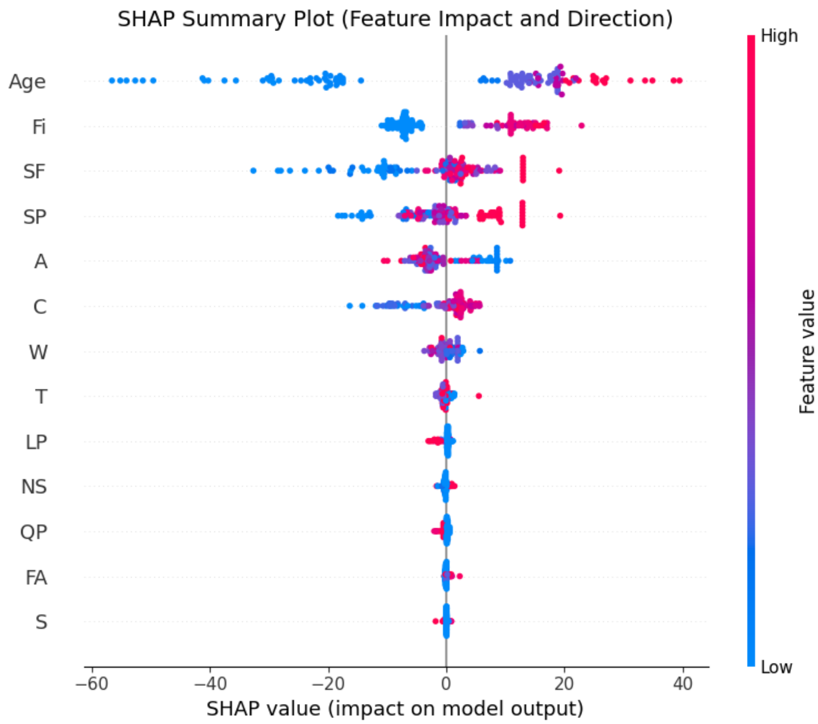


Fig. 9. SHAP summary plot illustrating the magnitude and direction of feature contributions to compressive strength, with color indicating low to high feature values

IV. DISCUSSION

The resulting Random Forest-based model has a good predictive ability in determining the compressive strength of Ultra-High-Performance Concrete (UHPC), as witnessed by high levels of quantitative performance through various assessment phases. The model obtained an R2 of 0.99 on the training set, meaning that it learned almost all the underlying data patterns, although it still had a higher generalization performance on the validation set (R2 = 0.94) and testing set (R2 = 0.96). The fact that the difference between the training and testing R2 is relatively small indicates that the overfitting was indeed kept under control, which is further proved by the fact that the Random Forest algorithm is an ensemble training method, and also the results of the K-fold cross-validation. The standard deviation of 0.01 in the average R2 of 0.96 in 10 folds indicates that the model is stable and not very sensitive to data partitioning [40, 41, 42, 43, 44]. Such consistency is especially essential when dealing with UHPC data, which can have nonlinear correlations and other complicated processes between the components of a mixture. The model can also be supported by error-based metrics in terms of reliability. The untuned Random Forest model generated an RMSE of 8.38 MPa and an MAE of 5.99 MPa on the testing dataset, and then hyperparameter tuning further decreased the values to 7.96 MPa and 5.72 MPa. In the same way, MAPE dropped to 5.06%, and CVRMSE went down to 6.54%. Though these are incremental improvements, they are significant with respect to UHPC strength prediction, in which even a modest decrease in error can be converted into more sound mix design choices and material efficiency. The comparatively small values of CVRMSE of all datasets demonstrate the fact that the errors of prediction are insignificant in comparison with the average compressive strength and reflect the strength of the suggested modelling strategy [45, 46, 47, 48, 49, 50, 51]. The hyperparameter analysis gives an additional understanding of the model behaviour of the Random Forest model. As the number of trees increased (n_estimators), the R2 improved up to an estimated 300-400 estimators at which point it levelled off. The chosen number of estimators, 400, is a good trade-off between prediction performance and the cost of computation. The max features analysis has shown that applying a subset of features (about 0.8 of all) at each split enhanced the model performance because this probably reduced the correlation between trees and increased the diversity of the ensemble. The max depth analysis revealed that shallow trees were not adequate to represent the nonlinear relationships, which were very complicated in UHPC data, and overly deep trees had diminished returns and a risk of overfitting. It was found that the optimal depth was 20 because it allowed the model to learn significant interactions without compromising generalization [22, 23, 24]. Further validation of the model is presented in the residual analysis among training, validation, and testing data sets. The residuals were evenly distributed around the zero with no characteristic trends or systematic bias, and this implies that the model does not either over-predict or under-predict compressive strength. The lack of funnel-shaped shapes in the residual plots indicates a homoscedastic behavior, i.e., the error of prediction is rather stable throughout the entire range of the predicted strength values. Although a small number of outliers were noticed at stronger strength levels, the frequency and the magnitude of the outliers were low, meaning that they do not contribute significantly to the overall model reliability. The behavior is in line with the reported RMSE and MAE values, and it confirms that the Random Forest model has constant predictions with a large strength range [52, 53, 54]. The interpretability analysis of SHAP provides a critical viewpoint regarding the way the model converts the input features to the predictions. The most significant variable was curing age, which, quantitatively, is measured by the largest mean absolute SHAP value and agrees with the observed trends of strength development in the data, with compressive strength slightly growing with greater age (that is, with strengths of over 170 180 MPa at later curing ages). Fiber content was the second most important as it showed that it has a great contribution to the enhancement of the strength and crack-bridging processes. Subsequently, silica fume and superplasticizer were mentioned, indicating the joint effect on microstructural densification and workability regulation that has a direct impact on compressive strength. The importance of cement and aggregate contents was moderate, which means that they are also necessary components, but their impact is more equal and is dependent on the interaction effects with additional supplies and the conditions of the curing. On the other hand, water, temperature, limestone powder, nano-silica, quartz powder, fly ash, and slag had relatively low SHAP values implying that over the ranges of study age, fiber content are more dominant than they are in relation to these other variables [32, 20, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65]. Critically speaking, although the model has good predictive power and interpretability, the findings also suggest that UHPC compressive strength is under the control of a few very significant variables, and others are secondary or conditional. This poses significant implications for future modeling projects, including whether a reduced-feature model can be as accurate and less complex, or indeed that the low-importance interactions among variables may become more important in out-of-sample settings. Furthermore, even though the Random Forest model is doing remarkably well, the marginal gains obtained after hyperparameter tuning indicate that the model is already at the performance level limit of the specified data set. More advancements might thus need better feature engineering, addition of microstructural or curing regime indicators, or a combination of hybrid modeling strategies instead of algorithmic optimization.

V. CONCLUSION AND RECOMMENDATION

The findings of this work prove that the modeling framework based on the method of the Random Forest (RF) is extremely efficient in predicting the compressive strength of the Ultra-High-Performance Concrete (UHPC). The model was also able to perform high predictive accuracy in training, validation, testing, and cross-validation processes, with high R2 values as high as 0.96 on unobserved data, and low error measures. The results support the claim that the RF model can be used to represent the intricate and nonlinear interactions among UHPC proportions, curing parameters, and compressive strength. The model is also robust in terms of the stable behavior of the residuals and the low variability of the K-fold cross-validation outcomes, meaning that it bears a high-quality generalization and is not prone to overfitting.

Relying on these findings, it is possible to introduce a number of recommendations regarding the necessity of future research and application in practice. First, the created Random Forest model can be successfully applied as a decision-support tool to optimize the UHPC mix design to allow practitioners to predict the compressive strength with sufficient precision and save on the experimental effort, time, and cost. Second, future research can be aimed at the further expansion of the dataset to include more curing regimes, sources of materials, and environmental conditions to make the models more applicable. Third, it might consider employing feature engineering approaches and hybrid modeling methods that involve the integration of machine learning with physical or microstructural variables to enhance the accuracy of prediction compared to the existing level. Also, comparative research on the advanced ensemble or deep learning models can provide more insights into the model efficiency and scalability. Last but not least, there should be a promotion of explainable artificial intelligence methods like SHAP into future research because this will allow transparency and confidence in UHPC performance prediction models based on data.

References

[1] M. Amran, S.-S. Huang, A. M. Onaizi, N. Makul, H. S. Abdelgader, and T. Ozbakkaloglu, “Recent trends in Ultra-High Performance Concrete (UHPC): Current status, challenges, and future prospects,” *Construction and Building Materials*, vol. 352, p. 129029, 2022. doi: https://doi.org/10.1016/j.conbuildmat.2022.129029

[2] H. Alabduljabbar, M. Khan, H. H. Awan, S. M. Eldin, R. Alyousef, and A. M. Mohamed, “Predicting ultra-high-performance concrete compressive strength using gene expression programming method,” *Case Studies in Construction Materials*, vol. 18, p. e02074, 2023. doi: https://doi.org/10.1016/j.cscm.2023.e02074

[3] J. Chen and Y. Bao, “Effect of dataset representation bias on generalizability of machine learning models in predicting flexural properties of Ultra-High-Performance Concrete (UHPC) beams,” *Engineering Structures*, vol. 326, p. 119508, 2025. doi: 10.1016/j.engstruct.2024.119508

[4] A. A. Hassoon, R. F. Hassan, and H. H. Hussein, “Optimizing UHPC Layers to Improve Punching Shear Performance in Concrete Slabs,” *Buildings*, vol. 15, no. 2, p. 209, 2025. doi: 10.3390/buildings15020209

[5] M. Katlav and F. Ergen, “Improved forecasting of the compressive strength of Ultra-High-Performance Concrete (UHPC) via the CatBoost model optimized with different algorithms,” *Structural Concrete*, vol. 26, no. 1, pp. 212-235, 2025.

[6] Y. J. Lee, S.-H. Lee, J. H. Kim, H. Jeong, S.-J. Han, and K. S. Kim, “Interlayer bond strength of 3D printed concrete members with Ultra High Performance Concrete (UHPC) mix,” *Buildings*, vol. 14, no. 7, p. 2060, 2024.

[7] D. M. Ors, M. Ramadan, A. M. F. Maree, A. H. Zaher, A. Afifi, and A. M. Ebid, “Machine learning base models to predict the punching shear capacity of posttensioned UHPC flat slabs,” *Scientific Reports*, vol. 14, no. 1, p. 3969, 2024. doi: 10.1038/s41598-024-54358-5.

[8] M. N. Khalil, Y. Farman, M. E. Kinaan, F. A. Zubair, and F. U. Rehman, “Experimental study on the effects of freeze-thaw progressions and performance of soil with non-toxic bio-enzyme,” *Journal of ICT, Design, Engineering, and Technology Sciences*, vol. 7, no. 1, p. 12–20, 2023. doi: 10.33150/JITDETS-7.1.3

[9] G. Mujtaba, N. Muhammad, Q. Shah, M. Khan, and N. K. Abbood, “Water pollution hazards and toxicity caused by textile industries effluent,” *Journal of ICT, Design, Engineering and Technological Science*, vol. 7, no. 2, pp. 22-27, 2023. doi: 10.33150/JITDETS-7.2.3

[10] S. N. K. Buledi, M. A. khan, A. Javed, Q. Shah, M. A. Khan, S. H. Shah, and U. Ali, “Flexural performance evaluation Of Glass Fiber Reinforced Polymer Bars (GFRP) in doubly reinforced beams: An experimental approach,” *Metallurgical and Materials Engineering*, vol. 31, no. 5, pp. 1747-1760, 2025. doi: 10.63278/mme.vi.1762

[11] Q. Shah, K. Byemba, G. Gali, A. Muhammad, A. Khan, and M. Khan, “To stabilize shear strength properties of an unwanted subgrade soil utilizing rock dust,” *Journal of ICT, Design, Engineering and Technological Science*, vol. 7, no. 2, pp. 1-6, 2023.

[12] K. Daud and U. Ali, “Transforming expansive clays from waste to value: Dual ash stabilization strategy for augmenting black cotton subgrade soil performance,” *International Journal of Environmental Sciences*, vol. 11, no. 23, p. 2025, 2025.

[13] S. C. Sapkota, S. Sapkota, T. Bansal, M. Haq, and M. A. Al-Osta, “Leveraging interpretable ensemble machine learning for predicting interfacial bond strength between normal-strength concrete substrate and UHPC overlays,” *Arabian Journal for Science and Engineering*, vol. 50, no. 11, pp. 8621-8645, 2025. doi: 10.1007/s13369-024-09690-3

[14] Z. Zhou, J. Chakma, M. A. Hoque, V. Chakma, and A. Ahmed, “Prediction of UHPC mechanical properties using optimized hybrid machine learning model with robust sensitivity and uncertainty analysis,” *Materials Research Express*, vol. 12, no. 8, p. 085703, 2025.

[15] S. Yin, M. Zhou, X. Zhao, and L. Ma, "Experimental investigation and analytical modeling of shear behavior at UHPC-NC interfaces in wet joints of precast assembly piers," *Advances in Structural Engineering*, vol. 28, p. 13694332251321205, 2025. doi: 10.1177/13694332251321205

[16] L. Xu, X. Yu, C. Zhu, L. Wang, and J. Yang, "Prediction of Ultra-High-Performance Concrete (UHPC) compressive strength based on convolutional neural networks," *Materials*, vol. 18, no. 12, p. 2851, 2025.

[17] W. Z. Taffese, Y. Zhu, and G. Chen, "Ensemble-learning model based ultimate moment prediction of reinforced concrete members strengthened by UHPC," *Engineering Structures*, vol. 305, p. 117705, 2024.

[18] S. C. Sapkota, A. Shrestha, M. Haq, S. Paudel, W. Tang, H. Kamyab, and D. Rocchio, "Enhancing shear strength predictions of UHPC beams through hybrid machine learning approaches," *Scientific Reports*, vol. 15, no. 1, p. 28259, 2025.

[19] M. N. Amin, B. Iftikhar, K. Khan, M. F. Javed, A. M. AbuArab, and M. F. Rehman, "Prediction model for rice husk ash concrete using AI approach: Boosting and bagging algorithms," in *Structures*. Elsevier, 2023.

[20] M. Liang, Z. Chang, Z. Wan, Y. Gan, E. Schlangen, and B. Šavija, "Interpretable ensemble-machine-learning models for predicting creep behavior of concrete," *Cement and Concrete Composites*, vol. 125, p. 104295, 2022. doi: 10.1016/j.cemconcomp.2021.104295

[21] M. S. O. Djediden, H. Reguieg, Z. M. Maaza *et al.*, "A distributed intrusion detection system based on apache spark and scikit-learn library," *Journal of Applied and Physical Sciences*, vol. 5, no. 1, pp. 30-36, 2019.

[22] P. Cihan, "Bayesian hyperparameter optimization of machine learning models for predicting biomass gasification gases," *Applied Sciences*, vol. 15, no. 3, p. 1018, 2025. doi: 10.3390/app15031018

[23] M. S. Khan, T. Peng, H. Akhlaq, and M. A. Khan, "Comparative analysis of automated machine learning for hyperparameter optimization and explainable artificial intelligence models," *IEEE Access*, vol. 13, p. 84966–84991, 2025. doi: 10.1109/access.2025.3566427

[24] R. Wang, J. Zhang, Y. Lu, S. Ren, and J. Huang, "Towards a reliable design of geopolymer concrete for green landscapes: A comparative study of tree-based and regression-based models," *Buildings*, vol. 14, no. 3, p. 615, 2024. doi: 10.3390/buildings14030615

[25] S. Demir and E. K. Sahin, "An innovative machine learning approach for slope stability prediction by combining shap interpretability and stacking ensemble learning," *Environmental Science and Pollution Research*, pp. 1-17, 2025. doi: 10.1007/s11356-025-36406-3

[26] Y. Gebreyesus, D. Dalton, S. Nixon, D. De Chiara, and M. Chinnici, "Machine learning for data center optimizations: Feature selection using Shapley additive exPlanation (SHAP)," *Future Internet*, vol. 15, no. 3, p. 88, 2023. doi: 10.3390/fi15030088.

[27] Y. Gebreyesus, D. Dalton, D. De Chiara, M. Chinnici, and A. Chinnici, "AI for automating data center operations: model explainability in the data centre context using shapley additive explanations (SHAP)," *Electronics*, vol. 13, no. 9, p. 1628, 2024.

[28] K. Islam, "Lending terms and bank capital," *Japanese Journal of Monetary and Financial Economics*, vol. 9, pp. 1-20, 2021. doi: 10.1007/s41024-024-00445-z

[29] 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 Studies in Construction Materials*, vol. 16, p. e01059, 2022. doi: 10.1016/j.cscm.2022.e01059

[30] A. Kashem and P. Das, "Compressive strength prediction of high-strength concrete using hybrid machine learning approaches by incorporating SHAP analysis," *Asian Journal of Civil Engineering*, vol. 24, no. 8, pp. 3243-3263, 2023. doi: 10.1007/s42107-023-00707-0

[31] 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 Studies in Construction Materials*, vol. 20, p. e02991, 2024.

[32] M. S. Khan, T. Peng, M. A. Khan, A. Khan, M. Ahmad, K. Aziz, M. M. Sabri Sabri, and N. Abd EL-Gawaad, "Explainable AutoML models for predicting the strength of high-performance concrete using Optuna, SHAP and ensemble learning," *Frontiers in Materials*, vol. 12, p. 1542655, 2025. doi: 10.3389/fmats.2025.1542655

[33] K. Aas, M. Jullum, and A. Løland, "Explaining individual predictions when features are dependent: More accurate approximations to Shapley values," *Artificial Intelligence*, vol. 298, p. 103502, 2021.

[34] A. Yazan and M. Andó, "SHAP-based insights for aerospace PHM: Temporal feature importance, dependencies, robustness, and interaction analysis," 2024. doi: 10.1016/j.rineng.2024.101834

[35] M. N. Amin, B. A. Salami, M. Zahid, M. Iqbal, K. Khan, A. M. Abu-Arab, A. A. Alabdullah, and F. E. Jalal, "Investigating the bond strength of FRP laminates with concrete using LIGHT GBM and SHAPASH analysis," *Polymers*, vol. 14, no. 21, p. 4717, 2022.

[36] D. Asprone, C. Menna, F. P. Bos, T. A. Salet, J. Mata-Falcón, and W. Kaufmann, "Rethinking reinforcement for digital fabrication with concrete," *Cement and Concrete Research*, vol. 112, pp. 111-121, 2018. doi: 10.1016/j.cemconres.2018.05.020

[37] M. L. Baptista, K. Goebel, and E. M. Henriques, "Relation between prognostics predictor evaluation metrics and local interpretability SHAP values," *Artificial Intelligence*, vol. 306, p. 103667, 2022.

[38] C. Cakiroglu, Y. Aydın, G. Bekdaş, and Z. W. Geem, "Interpretable predictive modelling of basalt fiber reinforced concrete splitting tensile strength using ensemble machine learning methods and SHAP approach," *Materials*, vol. 16, no. 13, p. 4578, 2023. doi: 10.3390/ma16134578

[39] D. Cardone, G. Perrone, and S. Sofia, "Numerical studies on the seismic retrofit of bridges using shape memory alloys," *Journal of materials engineering and performance*, vol. 20, no. 4, pp. 535-543, 2011.

[40] 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," *Scientific Reports*, vol. 14, no. 1, p. 5905, 2024.

[41] F. Yi, H. Yang, D. Chen, Y. Qin, H. Han, J. Cui, W. Bai, Y. Ma, R. Zhang, and H. Yu, "XGBoost-SHAP-based interpretable diagnostic framework for alzheimer's disease," *BMC medical informatics and decision making*, vol. 23, no. 1, p. 137, 2023.

[42] C. Yang, X. Guan, Q. Xu, W. Xing, X. Chen, J. Chen, and P. Jia, "How can SHAP (SHapley Additive exPlanations) interpretations improve deep learning based urban cellular automata model?" *Computers, Environment and Urban Systems*, vol. 111, p. 102133, 2024.

[43] S. R. Wani and M. Suthar, "Using machine learning approaches for predicting the compressive strength of ultra-high-performance concrete with SHAP analysis," *Asian Journal of Civil Engineering*, vol. 26, no. 1, pp. 373-388, 2025.

[44] M. Wang, R. Zhang, B. Yan, C. Song, Y. Lv, and H. Zhao, "Prediction of soil pollution risk based on machine learning and shap interpretable models in the Nansi Lake, China," *Toxics*, vol. 13, no. 4, p. 278, 2025.

[45] J. Abellan-Garcia and J. S. Guzmán-Guzmán, "Random forest-based optimization of UHPFRC under ductility requirements for seismic retrofitting applications," *Construction and Building Materials*, vol. 285, p. 122869, 2021.

[46] F. Ahmad, A. Alasskar, P. Samui, and P. G. Asteris, "Machine learning-based graphical user interface for predicting high-performance concrete compressive strength: Comparative analysis of gradient boosting machine, random forest, and deep neural network models," *Frontiers of Structural and Civil Engineering*, vol. 19, no. 7, pp. 1075-1090, 2025. doi: 10.1007/s11709-025-1201-8

[47] G. Biau, "Analysis of a random forests model," *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 1063-1095, 2012.

[48] F. Farooq, M. Nasir Amin, K. Khan, M. Rehan Sadiq, M. F. Javed, F. Aslam, and R. Alyousef, "A comparative study of random forest and genetic engineering programming for the prediction of compressive strength of High Strength Concrete (HSC)," *Applied Sciences*, vol. 10, no. 20, p. 7330, 2020.

[49] D. A. Fife and J. D'Onofrio, "Common, uncommon, and novel applications of random forest in psychological research," *Behavior Research Methods*, vol. 55, no. 5, pp. 2447-2466, 2023.

[50] T. Guhan and N. Revathy, "EMLARDE tree: Ensemble machine learning based random de-correlated extra decision tree for the forest cover type prediction," *Signal, Image and Video Processing*, vol. 18, no. 12, pp. 8525-8536, 2024.

[51] S. Han, H. Kim, and Y.-S. Lee, "Double random forest," *Machine Learning*, vol. 109, no. 8, pp. 1569-1586, 2020. doi: 10.1007/s10994-020-05889-1

[52] M. M. Kashani, J. Maddocks, and E. A. Dizaj, "Residual capacity of corroded reinforced concrete bridge components: State-of-the-art review," *Journal of Bridge Engineering*, vol. 24, no. 7, p. 03119001, 2019. doi: 10.1061/(asce)be.1943-5592.0001429

[53] M. Abed and E. Mehryaar, "A machine learning approach to predict relative residual strengths of recycled aggregate concrete after exposure to high temperatures," *Sustainability*, vol. 16, no. 5, p. 1891, 2024. doi: 10.3390/su16051891.

[54] M. Naser, "Digital twin for next gen concretes: On-demand tuning of vulnerable mixtures through Explainable and Anomalous Machine Learning," *Cement and Concrete Composites*, vol. 132, p. 104640, 2022. doi: 10.1016/j.cemconcomp.2022.104640

[55] R. Kitani and S. Iwata, "Verification of interpretability of phase-resolved partial discharge using a CNN with SHAP," *IEEE Access*, vol. 11, pp. 4752-4762, 2023.

[56] Y.-G. Lee, J.-Y. Oh, D. Kim, and G. Kim, "Shap value-based feature importance analysis for short-term load forecasting," *Journal of Electrical Engineering & Technology*, vol. 18, no. 1, pp. 579-588, 2023. doi: 10.1007/s42835-022-01161-9.

[57] Z. Li, "Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost," *Computers, Environment and Urban Systems*, vol. 96, p. 101845, 2022.

[58] W. Liao, X. Lu, Y. Huang, Z. Zheng, and Y. Lin, "Automated structural design of shear wall residential buildings using generative adversarial networks," *Automation in Construction*, vol. 132, p. 103931, 2021. doi: 10.1016/j.autcon.2021.103931

[59] K. Lin and Y. Gao, "Model interpretability of financial fraud detection by group SHAP," *Expert Systems with Applications*, vol. 210, p. 118354, 2022. doi: 10.1016/j.eswa.2022.118354

[60] Y. Liu, Z. Liu, X. Luo, and H. Zhao, "Diagnosis of Parkinson's disease based on SHAP value feature selection," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 3, pp. 856-869, 2022. doi: 10.1016/j.bbe.2022.06.007

[61] X. Man and E. Chan, "The best way to select features? comparing mda, lime, and shap," *The Journal of Financial Data Science Winter*, vol. 3, no. 1, pp. 127-139, 2021.

[62] V. Mechtcherine, R. Buswell, H. Kloft, F. P. Bos, N. Hack, R. Wolfs, J. Sanjayan, B. Nematollahi, E. Ivaniuk, and T. Neef, "Integrating reinforcement in digital fabrication with concrete: A review and classification framework," *Cement and Concrete Composites*, vol. 119, p. 103964, 2021. doi: 10.1016/j.cemconcomp.2021.103964

[63] U. Mitra, P. Sarkar, J. Mondal, and J. Kundu, "Enhancing interpretability in diabetics prediction: A comparative study of SHAP, LIME and permutation feature importance," in *2025 AI-Driven Smart Healthcare for Society 5.0*. IEEE, 2025.

[64] M. Naser, "A look into how machine learning is reshaping engineering models: The rise of analysis paralysis, optimal yet infeasible solutions, and the inevitable rashomon paradox," *Machine Learning for Computational Science and Engineering*, vol. 1, no. 1, p. 19, 2025. doi: 10.1007/s44379-025-00020-4

[65] 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," *Computer Methods and Programs in Biomedicine*, vol. 214, p. 106584, 2022. doi: 10.1016/j.cmpb.2021.106584