International Journal of Researches on Civil Engineering with Artificial Intelligence

www.ceai.reapress.com

Int. J. Res. Civ. Eng. AI. Vol. 2, No. 4 (2025) 227–244.

Paper Type: Original Article

Nonlinear Modeling of the Tensile Strength of Silica Fume-Modified Concrete Using Deep Neural Networks

Hassan Alijani[1,*], Masoud Ghasemi[2]

[1] Department of Civil Engineering, Shahid Beheshti University, Tehran, Iran; hassanalijani@gmail.com.

[2] Department of Civil Engineering, Faculty of Engineering, Universiti Malaya, 50603 Kuala Lumpur, Malaysia; ghasemi@um.edu.my.

Citation:

Received: 22 April 2025 *Revised: 11 June 2025* *Accepted: 16 September 2025*	Alijani, H., & Ghasemi, M. (2025). Nonlinear modeling of the tensile strength of silica fume-modified concrete using deep neural networks. *International journal of researches on civil engineering with artificial intelligence,* 2(4), 227-244.

Abstract

The splitting tensile strength of concrete plays a crucial role in maintaining structural integrity, as tensile stresses induced by loading and environmental variations commonly lead to cracking. This study investigates the effectiveness of advanced ensemble Machine Learning (ML) models, including LightGBM, Gradient Boosting Regression Tree (GBRT), Extreme Gradient Boosting (XGBoost), and AdaBoost, in accurately predicting the splitting tensile strength of silica-fume-modified concrete. Using a comprehensive database divided into training (80%) and testing (20%) sets, the performance of the models was evaluated using the R^2, Root Mean Square Error (RMSE), and MAE metrics. The results indicate that the GBRT and XGBoost models achieved superior predictive accuracy, with R^2 scores reaching 0.999 during the training stage and high accuracy during the testing stage (XGBoost: R^2 = 0.965, RMSE = 0.337; GBRT: R^2 = 0.955, RMSE = 0.381), outperforming the LightGBM and AdaBoost models. This study introduces GBRT and XGBoost as reliable and efficient alternatives to conventional experimental methods, offering significant savings in time and cost. In addition, SHapley Additive exPlanations (SHAP) analysis was performed to identify the key input features and clarify their effects on splitting tensile strength, providing valuable insights into the predictive behavior of silica-fume-modified concrete. The SHAP analysis revealed that the water-to-binder ratio and curing duration were the most influential factors affecting the splitting tensile strength of Silica-Fume (SF) concrete.

Keywords: Nonlinear modeling, Tensile strength, Silica-fume-modified concrete, Deep neural networks.

1 | Introduction

Concrete is one of the most widely used engineering materials in construction and civil infrastructure, and a large proportion of modern structures are designed based on its mechanical properties and durability. Despite its numerous advantages, concrete is inherently weak under tensile stresses, and this characteristic can

Corresponding Author: shariatimehdi@yahoo.com
https://doi.org/10.48314/ijrceai.v2i4.67

contribute to the initiation and propagation of cracks, loss of structural integrity, and ultimately deterioration in the performance and durability of concrete members. Therefore, accurate understanding and prediction of concrete tensile strength, in addition to compressive strength, are of particular importance in the design and assessment of concrete structures. This importance becomes even greater in concrete modified with Supplementary Cementitious Materials (SCMs), owing to the increased complexity of the material structure and behavior.

One of the widely used materials for modifying concrete properties is Silica Fume (SF). Due to its extremely fine particle size and high pozzolanic reactivity, SF can improve the microstructure of cement paste, increase the density of the cementitious matrix, and reduce concrete permeability. Studies on silica-fume-containing concrete have also shown that this material can affect a range of mechanical, permeability, and thermal properties of concrete [1]. Furthermore, investigations of the microstructure of concrete and cement paste containing SF have demonstrated that the pozzolanic reactions and filler effect of this material can lead to the formation of a denser structure [2]. From a sustainability perspective, replacing a portion of cement with SCMs such as SF can contribute to reducing clinker consumption and the environmental impacts associated with cement production, although the cost of some mix designs may increase [3].

Despite these advantages, the performance of silica-fume-containing concrete does not follow a simple linear relationship between mixture components and mechanical properties. Parameters such as the water-to-binder ratio, cement content, SF percentage, concrete age, aggregate type, and curing conditions can simultaneously affect concrete strength. Therefore, a change in one parameter may also alter the effects of other variables, making the final behavior of concrete a complex and nonlinear problem. Studies employing Machine Learning (ML) methods for concrete containing SCMs have also demonstrated that data-driven approaches can model such complex relationships with reasonable accuracy [4–6].

Among the mechanical properties of concrete, tensile strength is particularly important because many phenomena associated with cracking and failure are directly related to the capacity of concrete to withstand tensile stresses. Although compressive strength is generally considered the primary indicator of concrete strength, tensile strength, particularly splitting tensile strength, plays an important role in evaluating concrete behavior under cracking and loading-induced stresses. This issue is even more significant in silica-fume-containing concrete because changes in the microstructure and composition of the cementitious matrix can affect stress transfer and crack formation. Shah et al. [7], by simultaneously investigating the compressive strength and splitting tensile strength of silica-fume-containing concrete, demonstrated that ML methods can be used to predict these properties. However, compared with studies focusing on compressive strength, the number of studies specifically addressing the prediction of the tensile strength of silica-fume-modified concrete remains relatively limited.

In conventional approaches, concrete strength is primarily determined through laboratory experiments. Although these tests are essential and reliable from an engineering perspective, conducting a large number of experiments to investigate different concrete mixtures can require considerable time, cost, and materials. In addition, the use of empirical relationships and traditional regression models to predict concrete behavior may face limitations when the number of influential variables increases and the relationships among them become nonlinear. Under such conditions, ML methods, as data-driven approaches, can provide the ability to extract hidden relationships between input and output variables. Various studies have demonstrated that the application of ML in concrete engineering can facilitate the prediction of mechanical properties and reduce the need for some repetitive experimental tests [7], [8], [9].

In recent years, Artificial Neural Networks (ANNs) have been widely used to predict the mechanical properties of concrete. ANN models are capable of learning nonlinear relationships between mix-design parameters and concrete strength, and various studies have confirmed their applicability to conventional concrete, green concrete, and concrete modified with different materials [10–12]. However, increasing data complexity and the relationships among variables have encouraged the use of more advanced neural network architectures. Among these, Deep Neural Networks (DNNs), through the use of multiple processing layers, enable the learning of more complex relationships and the extraction of higher-level patterns from input data. Recent studies on concrete strength prediction have shown that appropriate tuning of network architecture and hyperparameters can significantly improve the performance of DNN models [13–15].

Alongside the development of neural networks, the use of ensemble learning algorithms and optimization methods has also become an important research direction in this field. Models such as Gradient Boosting, Extreme Gradient Boosting (XGBoost), and other ensemble methods have demonstrated highly satisfactory performance in some concrete strength prediction problems [16], [17]. In particular, the study by Al-Abdaly et al. [10] on the splitting tensile strength of silica-fume-modified concrete showed that Gradient Boosting Regression Tree (GBRT) and XGBoost models can achieve high prediction accuracy for this property, while the use of SHapley Additive exPlanations (SHAP) enables the relative importance of input variables to be identified [16]. This finding indicates that ML is not only a tool for prediction but can also be used to improve the understanding of the factors influencing concrete behavior.

More recent studies have further directed the development of intelligent models toward novel architectures, hyperparameter optimization, and rapid prediction systems. Naved et al. [15] investigated the potential for improving the performance of DNN models in predicting the compressive strength of concrete through DNNs and hyperparameter optimization. Dong et al. [16] also developed a fully connected neural network based on PyTorch for predicting the compressive strength of concrete. On the other hand, studies published in 2025 indicate that new neural network architectures and real-time intelligent systems are also emerging in the field of concrete strength prediction [18], [19]. This trend demonstrates the evolution of research from simple ANN models toward deeper, more optimized, and more scalable models.

Despite significant advances in the application of artificial intelligence to predict concrete properties, a substantial portion of the existing literature has focused on compressive strength. Studies on silica-fume-containing concrete have also, in many cases, addressed compressive strength or a combination of mechanical properties [20–22], whereas the prediction of tensile strength, particularly splitting tensile strength, has received comparatively less attention. On the other hand, the direct study by Al-Abdaly et al. [10] highlights the importance of using ML models to predict the splitting tensile strength of silica-fume-containing concrete. Nevertheless, the potential of deeper models to capture nonlinear relationships between mix-design parameters and tensile strength still requires further investigation.

Therefore, the main objective of the present study is to develop an intelligent approach for predicting the tensile strength of silica-fume-modified concrete using DNNs. In this approach, parameters influencing the mix design and concrete properties are considered as input variables, while tensile strength is treated as the model output. The objective is to develop a model with suitable accuracy and generalization capability by exploiting the ability of DNNs to learn complex and nonlinear relationships. Such a model can not only reduce dependence on numerous and costly experimental tests but also enable the rapid evaluation of different concrete mix designs.

Ultimately, the development of such a data-driven model can contribute to a better understanding of the relationship between mix-design parameters and tensile strength while providing a basis for optimizing the design of silica-fume-containing concrete. Given the growing application of artificial intelligence methods in civil engineering and the increasing use of deep models for predicting concrete properties [23], [24], the application of DNNs is expected to provide an effective step toward developing rapid, accurate, and cost-effective methods for evaluating the mechanical behavior of silica-fume-modified concrete.

2 | Literature Review

Concrete is one of the most widely used construction materials due to its favorable mechanical properties and its applicability in a wide range of structural members, playing a fundamental role in the construction industry. Among its various properties, tensile strength is an important parameter for evaluating concrete behavior, as the weakness of concrete in tension can contribute to crack initiation and propagation, brittle failure, and deterioration of structural performance. On the other hand, the use of pozzolanic materials and SCMs, such as SF, to improve the mechanical properties and durability of concrete has received considerable attention in recent years. However, the relationship between mixture components, curing conditions, and the ultimate strength of concrete, particularly tensile strength, is nonlinear and complex, and its prediction using experimental methods and conventional empirical relationships is subject to certain limitations. Accordingly, artificial intelligence and ML methods have been employed as effective tools for extracting complex relationships between mixture variables and the mechanical properties of concrete.

In recent years, numerous studies have demonstrated the capability of ANNs and ML algorithms for predicting concrete strength. Wu [1] developed a Radial Basis Function Artificial Neural Network (RBF-ANN) model for predicting the compressive strength of concrete and showed that appropriate selection of the network architecture can maintain high accuracy while reducing the computational complexity of the model. Lin and Wu [2] also proposed an ANN model for predicting concrete compressive strength and emphasized the ability of neural networks to model nonlinear relationships between mix-design parameters and concrete strength. Furthermore, Kovacevic et al. [3] investigated the performance of various ML methods for self-compacting concrete containing rubber and demonstrated that neural network-based models can provide acceptable accuracy in estimating the strength of special concretes.

With the development of ML methods, the application of more advanced models to concrete containing cement-replacement materials has also expanded. Nafees et al. [4] employed Multilayer Perceptron Neural Networks (MLPNN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Genetic Expression Programming (GEP) methods to model the mechanical properties of silica-fume-based green concrete and demonstrated the capability of artificial intelligence methods to predict the mechanical behavior of this type of concrete. In the study by Maherian et al. [5], ML was also used to estimate the compressive strength of nano-silica-modified concrete, and the results confirmed the high potential of data-driven models for establishing relationships between mixture characteristics and concrete strength. Nagaraju et al. [6], by investigating high-strength concrete and ternary mixtures containing different amounts of silica, demonstrated that ML methods can be used to predict concrete strength and evaluate the effects of variations in the quantities of cementitious materials.

Following this trend, increasing attention has been devoted to deep learning methods and the optimization of neural network architectures. Al-Gburi and Yusuf [25] employed DNNs to model the mechanical properties of silica-fume-modified concrete and investigated the capability of DNNs to extract complex nonlinear relationships between input variables and the mechanical properties of concrete. Similarly, Zhang et al. [8] demonstrated that deep learning models can provide satisfactory performance in predicting concrete compressive strength. Tin Bui et al. [9] employed the Whale Optimization Algorithm to optimize neural networks and improve the prediction accuracy of concrete compressive strength. These studies indicate that the use of a neural network alone is not sufficient; rather, appropriate architecture selection, hyperparameter tuning, and optimization of the training process can have a significant impact on model performance.

One of the studies closely related to the present research is that of Shah et al. [7], who predicted the compressive strength and splitting tensile strength of silica-fume-containing concrete using the M5P model tree algorithm. The results of this study demonstrated the importance of data-driven methods for predicting the tensile strength of silica-fume-containing concrete and showed that ML algorithms can serve as suitable alternatives to some costly and time-consuming experimental tests. In the same context, Al-Abdaly et al. [10] investigated advanced ensemble ML models, including LightGBM, GBRT, XGBoost, and AdaBoost, for predicting the splitting tensile strength of silica-fume-modified concrete in 2024. The results showed that the GBRT and XGBoost models achieved very satisfactory performance. In addition to their high predictive accuracy, the application of SHAP analysis enabled the identification of the factors influencing tensile strength. In that study, the water-to-binder ratio and curing time were identified as the most influential variables affecting tensile strength [10]. This finding highlights the importance of analyzing nonlinear relationships between mix-design parameters and tensile strength in the present study.

Ansari et al. [11] also employed ML models, including ANN, Random Forest, and AdaBoost, to predict the residual compressive strength of silica-fume-modified concrete subjected to thermal loading in 2024. The results demonstrated that hyperparameter tuning and the selection of appropriate evaluation metrics, such as R^2, Root Mean Square Error (RMSE), and MAE, play an important role in improving model accuracy. Abdulrahman et al. [12] modeled the compressive strength of silica-fume-modified self-compacting concrete using 240 laboratory datasets. The results showed that the M5P model outperformed linear and multivariable regression models, while curing age was identified as one of the most influential variables affecting strength. These findings once again demonstrate that concrete strength is a nonlinear function of a set of mix-design parameters and curing conditions.

In more recent studies, the use of hybrid models and optimization algorithms has also received increasing attention. Babiker et al. [13] employed ensemble ML methods to optimize the strength of concrete containing multiple cementitious compositions. Alhakim et al. [14] used the GBRT algorithm combined with GridSearchCV to predict the strength of environmentally friendly concrete and demonstrated the importance of hyperparameter tuning. Naved et al. [15], in 2024, investigated the performance of DNN models for predicting concrete compressive strength using DNNs and hyperparameter optimization, showing that optimization of the network architecture can reduce prediction error. Dong et al. [16] also developed a fully connected neural network based on PyTorch for predicting concrete compressive strength and demonstrated the potential of modern deep learning frameworks for applications in concrete engineering.

On the other hand, recent studies have moved toward more innovative modeling approaches. Alibrahim et al. [17] proposed a brain-inspired multi-lobe neural network architecture for estimating concrete compressive strength in 2025. The results showed that the proposed architecture could reduce RMSE and MAE errors compared with conventional models and improve the generalization capability of the model. Marchawka et al. [18] also developed a real-time early-strength prediction system for concrete using a DNN and data obtained from hydration-process monitoring. This study demonstrates that the application of artificial intelligence in concrete engineering is expanding from laboratory-based prediction toward intelligent and real-time systems.

Numerous studies have also been conducted in the field of sustainable concrete. Amar et al. [19] predicted the strength of concrete containing waste materials using an ANN and demonstrated the applicability of ANN models to concrete incorporating cement-replacement materials. Mater et al. [20] also proposed an ANN- and Python-based model for predicting the strength of green concrete. In addition, Peng and Onwualu [21] investigated the application of several ML algorithms to analyze the mechanical performance of fly-ash-based geopolymer concrete. These studies indicate that artificial intelligence methods are not limited to conventional concrete and can also be extended to novel and sustainable cementitious materials.

Recent review studies have further emphasized the expanding application of artificial intelligence in concrete engineering. Terlomon et al. [22] conducted a review of artificial intelligence techniques for predicting the compressive strength of self-compacting concrete and reported that deep models outperform shallow neural networks under many conditions. Schlot et al. [23], through a review of studies on silica-fume-containing concrete, examined the importance of this material in improving the mechanical, permeability, and thermal properties of concrete. Finally, Ben Chaabene et al. [24], in a review of the application of ML to the prediction of concrete mechanical properties, emphasized that the development of data-driven models can reduce dependence on experimental testing and enable faster prediction of concrete properties.

Based on the review of previous studies, it can be concluded that although the application of ANN, ensemble algorithms, metaheuristic methods, and DNNs to concrete strength prediction has been extensively investigated, a substantial proportion of existing research has focused on compressive strength, while studies specifically devoted to the nonlinear modeling of the tensile strength of silica-fume-modified concrete remain relatively limited. Furthermore, different ML methods have been employed in existing studies with varying architectures and input variables, and there is still a need to investigate deeper and more optimized models to achieve higher predictive accuracy. Therefore, the present study focuses on the application of DNNs to predict the tensile strength of silica-fume-modified concrete and aims to model the complex relationship between mix-design parameters and tensile strength in a data-driven manner. This approach can reduce the need for extensive experimental testing while providing a basis for faster strength prediction and the optimization of concrete mix designs for engineering applications.

3 | Research Methodology

This study presents its research methodology in three main stages. In the first stage, raw data obtained from various studies were processed, and a refined database containing 158 samples was established. Each sample consists of seven input variables: Cement content, SF content, water-to-binder ratio (W/B), Fine Aggregate (FA) content, Coarse Aggregate (CA) content, curing age, and one output variable, namely splitting tensile strength.

In the second stage, the database was divided into two subsets: 80% of the data were used for model training and 20% for model testing. At this stage, four ensemble ML algorithms, namely AdaBoost, XGBoost, GBRT, and LightGBM, were employed to train the models. The performance of these algorithms was evaluated using three statistical metrics: The coefficient of determination (R^2), RMSE, and Mean Absolute Error (MAE). To ensure the reliability of the models, a five-fold cross-validation (5-fold cross-validation) technique was also applied. Finally, the algorithms demonstrating the highest accuracy and reliability were advanced to the third stage.

In the third stage, the splitting tensile strength was predicted using the most effective ML models, and the results were analyzed using Taylor diagrams. The overall interpretation of the model predictions was performed using the SHAP method. Finally, the performance of the best-performing ML model was compared with a previously developed model to determine the extent of improvement in prediction accuracy and efficiency, thereby highlighting the advances achieved in predictive modeling for the intended application.

3.1 | Data Used

The dataset consists of 158 experimental observations covering a wide range of values for each predictor variable. The dataset exhibits considerable variability in its features. The characteristics of the dataset are as follows:

- *Cement content ranges from 197 to 800 kg/m³.*
- *SF content ranges from 0 to 250 kg/m³.*
- *The water-to-binder ratio (W/B) ranges from 0.14 to 0.83.*
- *FA content ranges from 535 to 1315 kg/m³.*
- *CA content ranges from 0 to 1248 kg/m³.*
- *Superplasticizer (SP) content ranges from 0 to 80 kg/m³.*
- *Curing age ranges from 1 to 91 days.*
- *Splitting tensile strength, which is the target variable of the model, ranges from 0.51 to 10.00 MPa.*

This statistical analysis provides a comprehensive overview of the variability and distribution of the factors affecting the splitting tensile strength of concrete. Furthermore, the prediction model developed using these data improves the ability to accurately predict the tensile strength of SF-modified concrete and contributes to a more effective assessment of construction material performance.

Table 1. Statistical summary of the input predictors for SF concrete and its tensile strength.

Parameter	Cement (kg/m³)	SF (kg/m³)	W/B	FA (kg/m³)	CA (kg/m³)	SP (kg/m³)	Age (Days)	Tensile Strength (MPa)
Min	197	0.00	0.14	535.00	0.00	0.00	1.00	0.51
Max	800	250.00	0.83	1315.00	1248.00	80.00	91.00	10.00
Average	458.13	54.17	0.38	815.96	893.79	13.16	32.08	4.23

Fig. 1 presents the statistical distribution of the components affecting the tensile strength of SF-modified concrete. The data are presented using a series of box plots. Examination of these plots shows that variables such as cement content, SF, water-to-binder ratio (W/B), FA, CA, SP, and curing age exhibit the expected variations, along with several outliers. This indicates the presence of diverse mixture compositions and different curing conditions within the dataset.

In particular, cement content and FA exhibit relatively few outliers, indicating that the values of these parameters are relatively stable and concentrated around their respective medians. This may indicate greater consistency in the batching of the concrete mix designs. In contrast, the water-to-binder ratio (W/B) and SP exhibit greater variability, and their outliers may represent mixture proportions associated with different practical applications and projects.

SF also contains several outliers, indicating the diverse applications and quantities of this material in different concrete mixtures, although its overall distribution remains concentrated within the interquartile range. The CA content is similar to cement and FA in terms of its distribution and contains relatively few outliers, which may indicate greater standardization in aggregate quantities and gradation.

The curing age also exhibits a relatively broad range with some degree of deviation. Most specimens were tested within a similar time interval, whereas a number of samples correspond to longer or shorter curing periods.

Finally, the tensile strength of concrete, which represents the primary target variable of the model, exhibits considerable variability. This variation reflects significant differences in the mechanical properties of the concrete, which are likely attributable to variations in mixture proportions and curing durations. In addition, the mean tensile strength may be influenced by the presence of outliers, which could result from specific mixture compositions or experimental anomalies.

Therefore, the importance of accurate sampling and experimental procedures in ensuring data quality and the reliability of predictive models is emphasized.

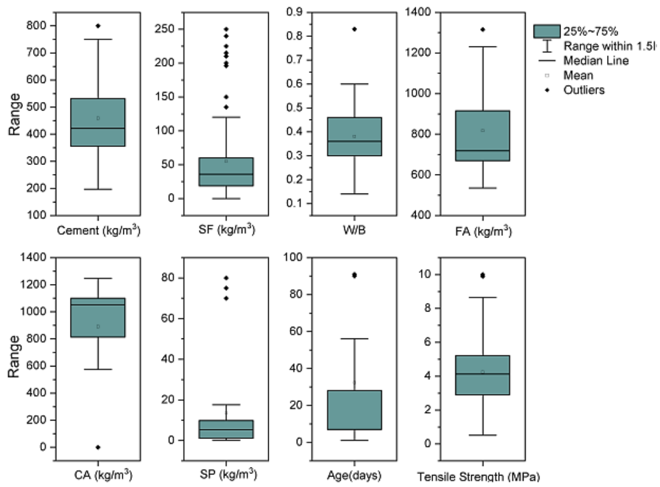


Fig. 1. Box plot of the distribution of mix-design components and tensile strength of SF concrete.

Fig. 2 presents the Pearson Correlation Coefficients (PCC), which indicate the strength and direction of the relationships between the various mix-design factors and the tensile strength of SF-containing concrete. The PCC values range from −1 to +1, where +1 represents a perfect positive correlation, −1 represents a perfect negative correlation, and 0 indicates no correlation.

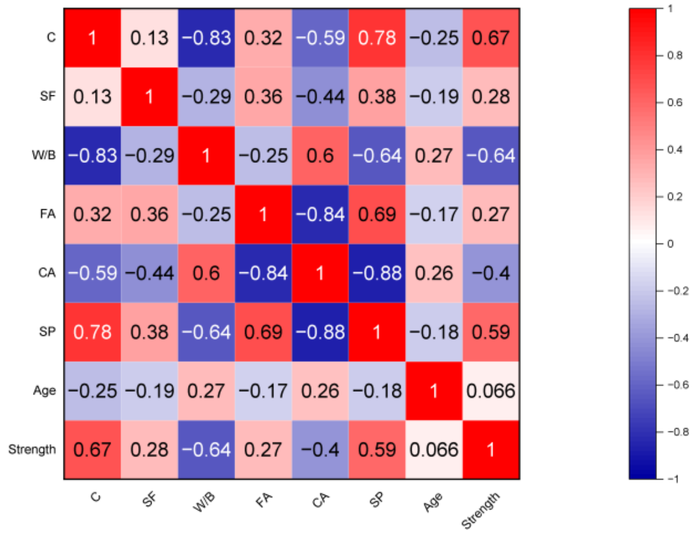
According to the results, the cement Content (C) exhibits a strong positive correlation with tensile strength, indicating that an increase in cement content is generally associated with an increase in tensile strength. In contrast, the water-to-binder ratio (W/B) shows a significant negative correlation with tensile strength, indicating that an increase in this ratio may lead to a reduction in strength.

The correlation heatmap also clearly illustrates the relationships among the different mix-design parameters. For example, a strong negative correlation is observed between the water-to-binder ratio (W/B) and cement Content (C), which may indicate that mixtures with lower water-to-binder ratios generally contain higher cement contents. Similarly, FA and SP exhibit a negative correlation, suggesting that an increase in SP content is generally associated with a reduction in the amount of FA.

A negative correlation is also observed between CA and SP, which may be attributed to compensatory adjustments in mix designs to achieve the desired workability and strength properties. Conversely, the positive correlation between cement Content (C) and SP suggests that mixtures with higher cement contents may require greater amounts of SP to maintain adequate workability.

These correlation relationships provide valuable insights into the complex interactions influencing the properties of SF concrete and are important for optimizing concrete mix formulations to enhance tensile strength and overall performance.

According to *Fig. 2*, the correlation coefficients between all variables are lower than 0.95, indicating that the relationships among the variables are not entirely linear. This may pose challenges for conventional empirical fitting methods, which often rely on assumptions of linear or relatively simple dependencies. Nevertheless, artificial intelligence-based ML algorithms have demonstrated strong performance in modeling and analyzing nonlinear interactions among variables, particularly in accurately predicting material performance [21].



	C	SF	W/B	FA	CA	SP	Age	Strength
C	1	0.13	−0.83	0.32	−0.59	0.78	−0.25	0.67
SF	0.13	1	−0.29	0.36	−0.44	0.38	−0.19	0.28
W/B	−0.83	−0.29	1	−0.25	0.6	−0.64	0.27	−0.64
FA	0.32	0.36	−0.25	1	−0.84	0.69	−0.17	0.27
CA	−0.59	−0.44	0.6	−0.84	1	−0.88	0.26	−0.4
SP	0.78	0.38	−0.64	0.69	−0.88	1	−0.18	0.59
Age	−0.25	−0.19	0.27	−0.17	0.26	−0.18	1	0.066
Strength	0.67	0.28	−0.64	0.27	−0.4	0.59	0.066	1

Fig. 2. Heatmap of the correlations between the predictor variables and the target variable.

4 | Model Results

4.1 | Statistical Evaluation of the Models

The statistical evaluation of the prediction models is presented in *Table 2*. As observed, the GBRT and XGBoost models demonstrated superior performance compared with the other models. These two models achieved very high and nearly perfect R^2 scores of approximately 0.999 during the training stage and maintained impressive R^2 values above 0.955 during the testing stage, indicating a very high level of prediction accuracy in both stages. Although the LightGBM and AdaBoost models performed slightly less effectively, they still demonstrated considerable predictive capability. LightGBM achieved R^2 values of 0.897 for training and 0.799 for testing, while AdaBoost achieved R^2 values of 0.902 and 0.845, respectively, indicating satisfactory performance.

The error metrics, including RMSE and MAE, also clearly reflect the prediction accuracy of the models. Among these metrics, GBRT and XGBoost exhibited the lowest errors across the different evaluation stages. Their RMSE values during the training stage were 0.057 and 0.059, respectively, while the corresponding values for the overall dataset decreased to 0.179 and 0.161. These results demonstrate the strong capability of these models to minimize prediction errors. In contrast, LightGBM and AdaBoost produced higher RMSE and MAE values; however, they still demonstrated reasonable predictive performance, with LightGBM achieving an MAE of 0.416 and AdaBoost recording an RMSE of approximately 0.605.

Overall, the results reveal a clear performance hierarchy, which can be expressed as

$$\text{GBRT} \approx \text{XGBoost} > \text{LightGBM} > \text{AdaBoost}.$$

This ranking is consistently observed across the different evaluation metrics, including R^2, RMSE, and MAE, as well as across the training, testing, and overall datasets.

Although the very high R^2 scores obtained by the GBRT and XGBoost models may initially suggest the possibility of overfitting to the training data, their strong performance and ability to maintain high accuracy during the testing stage indicate good generalization capability when applied to unseen data. Al-Abdaly et al. [10] also emphasized that, to avoid overfitting, the difference in R^2 between the training and testing stages should be less than 0.05. This criterion is well satisfied by the GBRT and XGBoost models.

Fig. 3 visually illustrates the prediction accuracy of each developed model across the different evaluation stages and, based on the R^2 and RMSE metrics, further highlights the superior performance of the GBRT and XGBoost models.

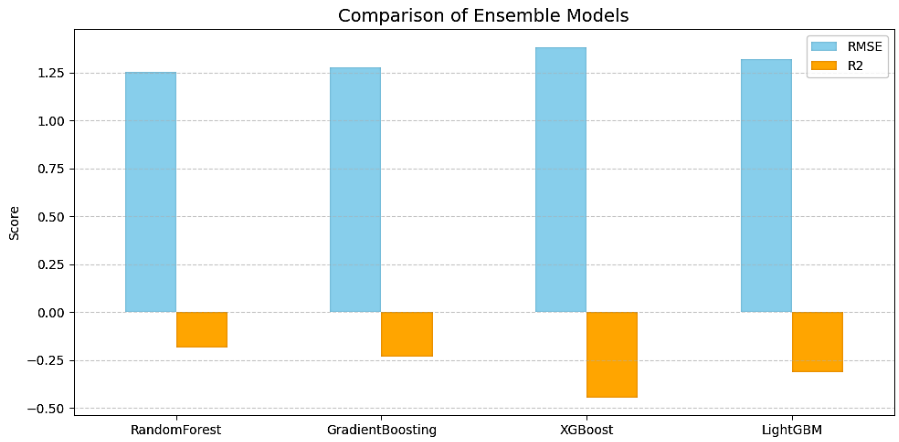


Fig. 3. Visual comparison of RMSE and R^2 metrics for the different proposed models.

Table 2. Statistical performance metrics of the LightGBM, GBRT, XGBoost, and AdaBoost models.

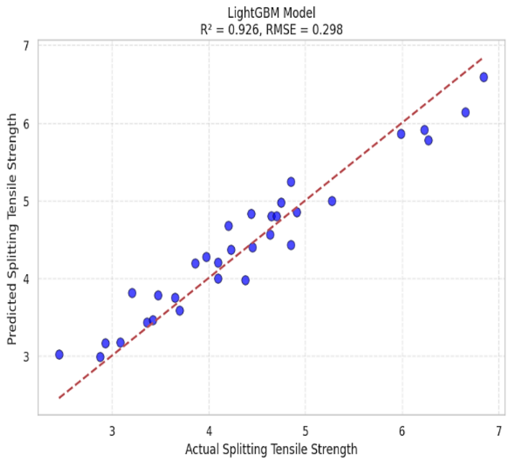
Phase	Criteria	LightGBM	GBRT	XGBoost	AdaBoost
Training	R^2	0.897	0.999	0.999	0.902
	RMSE	0.589	0.057	0.059	0.576
	MAE	0.357	0.041	0.043	0.483
Testing	R^2	0.799	0.955	0.965	0.485
	RMSE	0.807	0.381	0.337	0.707
	MAE	0.645	0.301	0.267	0.575
All	R^2	0.883	0.991	0.993	0.890
	RMSE	0.639	0.179	0.161	0.605
	MAE	0.416	0.094	0.089	0.502

4.2 | Cross Plots

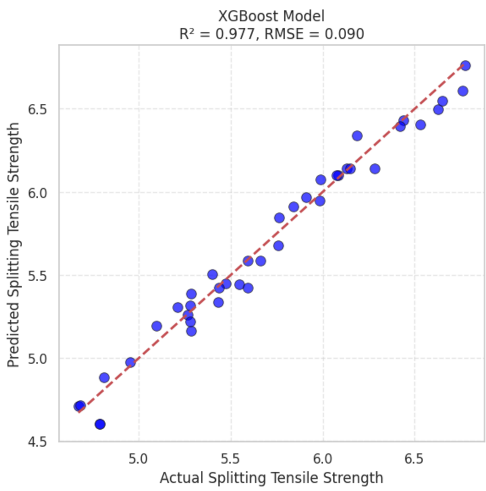
In the scatter plots presented in *Fig. 4*, the values predicted by the models are plotted against the actual experimental data. The data points are distributed around a diagonal line representing a one-to-one relationship between the predicted and actual values, commonly referred to as the 45-degree line. This line originates from the intersection of the coordinate axes, and the accuracy and reliability of the models can be assessed based on the proximity of the data points to this line. The closer the data points are to the 45-degree line, the higher the prediction accuracy of the model.

An examination of the scatter plots in *Figs. 4.a–4.d* indicates that all four developed models provide acceptable accuracy and reliability for predicting the splitting tensile strength, with a strong correlation between the predicted values and the actual experimental data. A closer examination of *Figs. 4.a* and *4.b* shows that the XGBoost and GBRT models, particularly in predicting splitting tensile strength, outperform the other models.

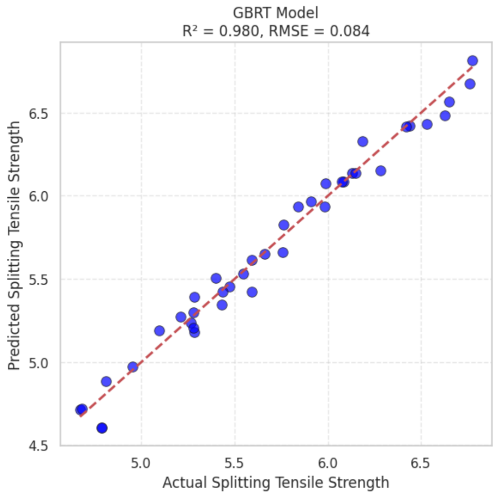
Similarly, *Figs. 4.c* and *4.d* indicate that the AdaBoost model provides accuracy comparable to that of the LightGBM model and demonstrates satisfactory predictive performance.



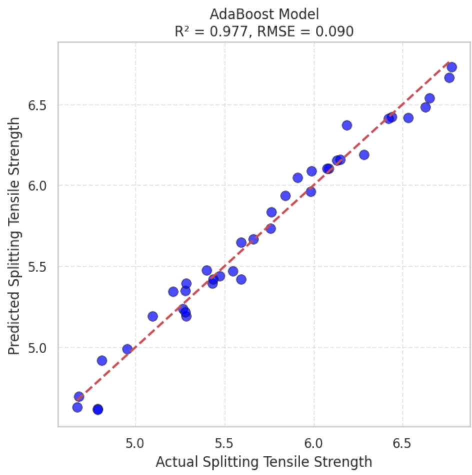
a.



b.



c.



d.

Fig. 4. Comparative scatter plots for the splitting tensile strength prediction models.

4.3 | Residual Distribution Histograms of the Models

The histogram analysis presented in *Fig. 5*, comparing the residual distributions of the four distinct ML models—XGBoost, GBRT, AdaBoost, and LightGBM—highlights several key aspects of their predictive performance. The XGBoost and GBRT models exhibit an excellent fit to the training data, with their residuals being densely concentrated around zero. This indicates a high level of prediction accuracy and a strong fit between the models and the observed data.

In contrast, the AdaBoost and LightGBM models exhibit broader residual distributions, indicating relatively weaker model fitting. When generalization capability is considered, based on model performance on the testing dataset, XGBoost and GBRT continue to maintain high accuracy and demonstrate strong capability in predicting new and previously unseen data. This robust performance contrasts with that of AdaBoost and LightGBM, whose more dispersed residuals during testing may indicate reduced prediction accuracy and potential concerns regarding their generalization capability.

Overall, the analysis indicates that XGBoost and GBRT are the most suitable models for the intended application, as their concentrated residual distributions and stable performance during both the training and testing stages provide strong evidence of their ability to achieve an appropriate balance between data fitting and generalization to new observations. In contrast, AdaBoost and LightGBM, due to their more dispersed residual distributions, exhibit greater variability in their predictions and a higher likelihood of prediction errors.

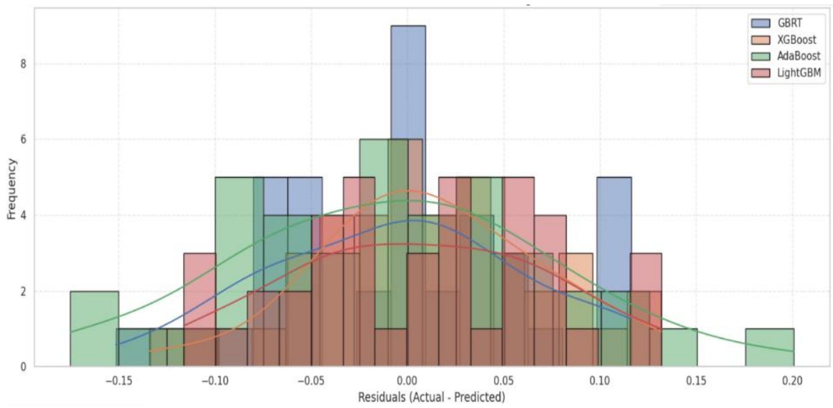


Fig. 5. Error distribution plots for splitting tensile strength prediction by different models: (a) Training, and (b) Testing.

4.4 | Cumulative Frequency Plot

Fig. 6 illustrates the relationship between the absolute prediction errors and the cumulative frequency percentage for all models investigated in this study. It can be clearly observed that a minimum proportion of the predictions made by the LightGBM model falls within the absolute error range of 0 to 0.2. The figure also indicates that more than 90% of the predictions made by the XGBoost and GBRT models have absolute errors close to 0.2.

Furthermore, the curves corresponding to these two models reach a cumulative frequency of 100% more rapidly, indicating that all data points are covered within a relatively small error range. This suggests that a substantial proportion of the predictions generated by these models have absolute errors of less than 0.8. Therefore, the cumulative frequency plot clearly demonstrates the superior predictive performance of the GBRT and XGBoost models compared with the other models investigated in this study.

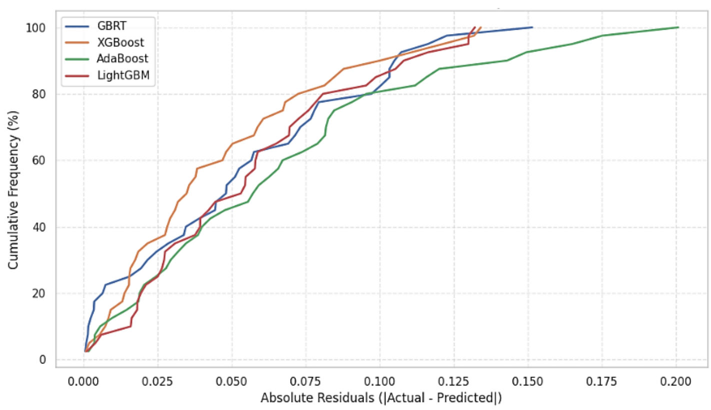


Fig. 6. Comparison of the cumulative frequency error of the intelligent models used in the study.

4.5 | Taylor Diagram

Fig. 7 presents a commonly used method for comparing multiple models through a two-dimensional diagram incorporating three key statistical measures: Standard Deviation (SD), RMSE, and correlation coefficient (R). In this diagram, each model is represented by a distinct point and color. The star symbol located on the horizontal axis represents the statistical values of the experimental data (SD = 1.87, RMSE = 0, and R = 1) and serves as the reference point.

The proximity of each model to this reference point indicates better model performance. The blue arcs represent lines of constant RMSE that originate from the reference point. The black arcs represent lines of constant SD, which originate from the center of the diagram and intersect the axes. The additional black radial lines extending outward from the center represent constant values of the correlation coefficient (R).

The position of each model relative to the reference point indicates its accuracy and effectiveness in predicting the splitting tensile strength. Based on the Taylor diagram, LightGBM and XGBoost were identified as the least effective and most effective models, respectively, for predicting this parameter.

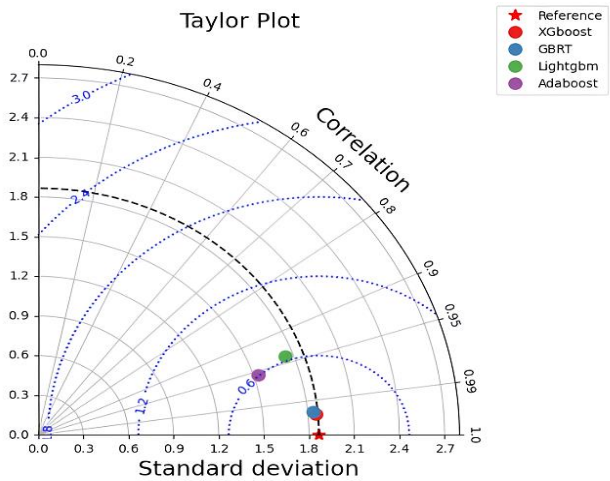


Fig. 7. Taylor diagram for the model predictions of splitting tensile strength.

4.6 | Feature Importance Analysis

The XGBoost model was identified as the best-performing model for predicting the splitting tensile strength. Understanding the influence of each input variable on splitting tensile strength is highly valuable for optimizing the mix proportions of SF concrete and, consequently, enhancing its tensile strength.

In this study, the SHAP approach was employed to comprehensively evaluate feature importance. By calculating SHAP values, this method provides a detailed assessment of the contribution of each feature to the model output. The SHAP approach accounts for the interactions among input features and systematically determines the SHAP value associated with each feature. This enables a detailed evaluation of the influence of individual features on the model predictions and facilitates the analysis of their respective contributions to the model's decision-making process.

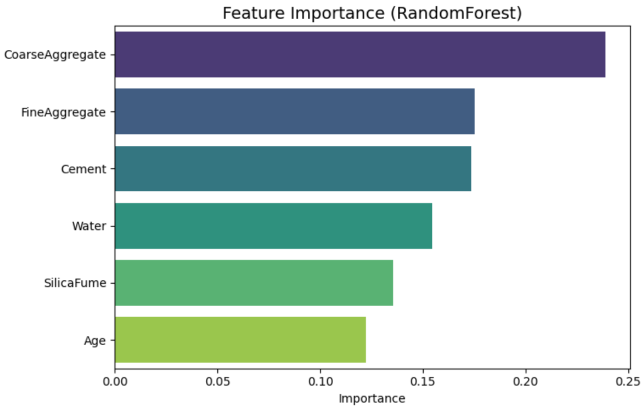
The following equation presents the method used to calculate the SHAP values:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} \left(f(S \cup \{i\}) - f(S)\right)$$

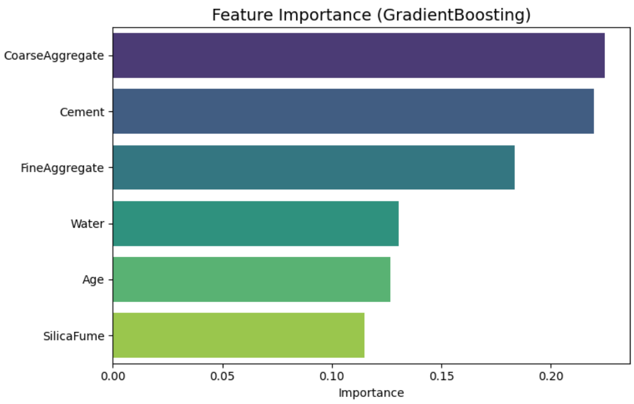
Suppose that N represents the set containing all features, while S is a subset of N that does not include a particular feature i. The size of subset S, denoted by $|S|$, represents the number of features contained in S, while $|N|$ denotes the total number of features.

The function $f(S)$ represents the output of the prediction model when only the features included in S are considered. Similarly, $f(S \cup \{i\})$ represents the model output when feature i is added to subset S. The SHAP value for feature i, denoted by ϕ_i, quantifies the average contribution of that feature across all possible combinations of the other features.

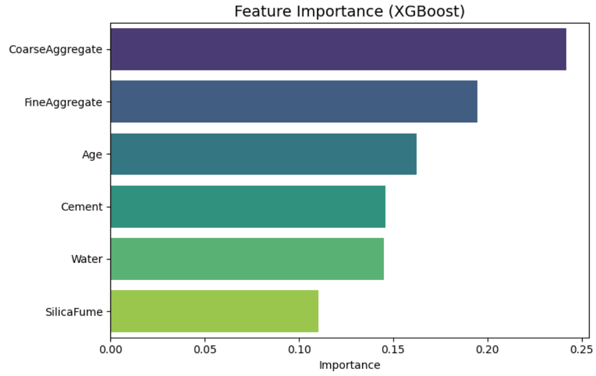
In the SHAP summary plot (*Fig. 8*), the water-to-binder ratio (W/B) and age (Age) are identified as the most important factors in predicting the splitting tensile strength of concrete, exerting a substantial influence on the model output. Their importance is clearly reflected by their SHAP values. The other variables, including cement, SF, CA, SP, and FA, are also ranked according to their influence on the model output and play important roles in the prediction process.



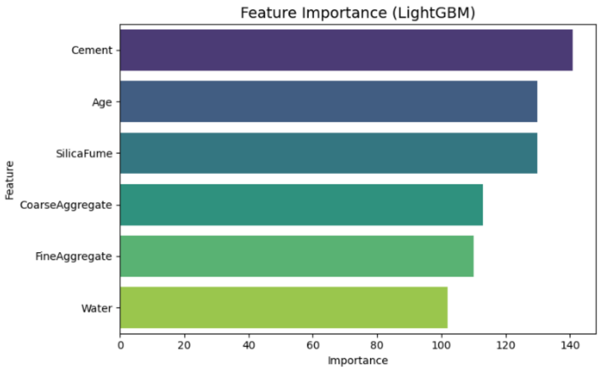
a.



b.



c.



d.

Fig. 8. Overview of the influence of each feature on splitting tensile strength.

Fig. 9 illustrates the positive and negative effects of the different input components on splitting tensile strength. Accordingly, features such as age, cement content, SF, FA, and CA have a positive effect on tensile strength. In contrast, the water-to-binder ratio (W/B) and SP have a negative effect on tensile strength.

These findings are consistent with the results reported in previous studies, particularly those of Nafees et al. [4].

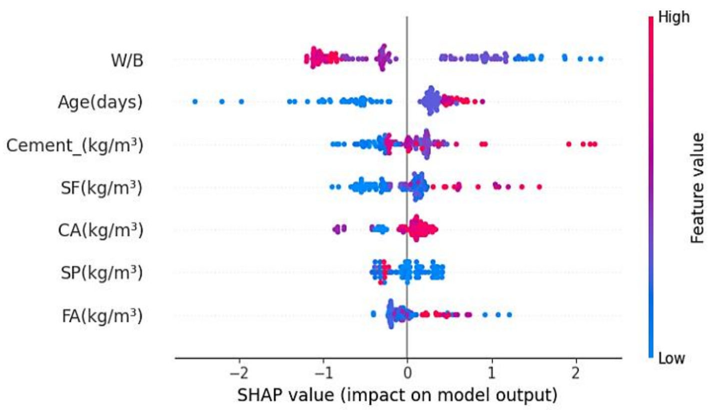


Fig. 9. Negative and positive effects of the features on splitting tensile strength.

4.7 | Comparative Analysis of the XGBoost Model

Previous studies have employed various ML techniques to predict the splitting tensile strength of SF concrete. In the present study, the results obtained using the XGBoost model were compared with other established models, including MLPNN, ANFIS, and GEP, which were employed in the study by Nafees et al. [4]. The XGBoost approach demonstrated higher accuracy in predicting the splitting tensile strength of SF concrete, while the MLPNN, ANFIS, and GEP models also achieved satisfactory performance, with R^2 values of 0.90, 0.92, and 0.93, respectively, as presented in *Table 3*.

Table 3. Comparison of statistical performance metrics for the prediction models.

Statistical Parameter	This Study (XGBoost)	Nafees et al. – MLPNN [4]	Nafees et al. – ANFIS [4]	Nafees et al. – GEP [4]
R^2	0.993	0.90	0.92	0.93
RMSE	0.16	0.51	0.40	0.31
MAE	0.089	0.41	0.26	0.31

The superior performance of the XGBoost model in predicting splitting tensile strength may be attributed to its robustness against noisy data and outliers. The model is designed to progressively focus on samples that are more difficult to predict, thereby improving its ability to handle complex and heterogeneous datasets. Compared with MLPNN and ANFIS models, XGBoost may exhibit greater resistance to overfitting when applied to noisy datasets.

Furthermore, by combining multiple weaker learners, XGBoost is capable of effectively modeling complex nonlinear relationships between input and output variables. This characteristic can provide an advantage over MLPNN and ANFIS models under certain conditions. XGBoost also offers high computational efficiency, whereas MLPNN and ANFIS models may require more complex architectures and longer training times. This computational advantage makes XGBoost a suitable option for applications with limited computational resources.

Another factor that may contribute to the differences in predictive performance is the variation in the input variables used in the respective models. Differences in dataset composition, feature selection, and experimental conditions can therefore partly explain the superior performance of XGBoost compared with the other models. Overall, the ability of XGBoost to handle noisy data, model complex nonlinear interactions, and provide high computational efficiency makes it a powerful tool for predicting the splitting tensile strength of concrete.

5 | Conclusion

This study provides a comprehensive investigation of the effectiveness of various ML models, including GBRT, XGBoost, LightGBM, and AdaBoost, for predicting the splitting tensile strength of SF concrete. The main findings of the study can be summarized as follows:

I. XGBoost and GBRT achieved the highest prediction accuracy for the splitting tensile strength of SF concrete. Their R^2 values exceeded 0.955, while they also recorded the lowest RMSE and MAE values across the evaluation stages. These models exhibited more concentrated residual distributions and closer agreement with the experimental data than LightGBM and AdaBoost, establishing a clear performance hierarchy among the investigated models. Although LightGBM and AdaBoost exhibited somewhat lower prediction accuracy, their testing R^2 values of 0.799 and 0.845, respectively, still indicate acceptable predictive performance under certain conditions.

II. The XGBoost model outperformed the models used in previous studies, including MLPNN, ANFIS, and GEP, achieving an impressive overall R^2 value of 0.993. This superior accuracy may be attributed to the robustness of XGBoost against noisy and outlier data, its ability to model complex nonlinear relationships, and its high computational efficiency.

III. SHAP-based feature importance analysis demonstrated that the water-to-binder ratio (W/B) and concrete age were the most influential factors affecting the splitting tensile strength of SF concrete. These findings confirm the effectiveness of ensemble ML models in accurately predicting this important mechanical property and provide valuable practical insights for optimizing concrete mix designs.

Overall, this study focused on predicting the splitting tensile strength of silica-fume-modified concrete using advanced ensemble ML models, including XGBoost, GBRT, LightGBM, and AdaBoost. The results demonstrated that XGBoost and GBRT achieved superior performance compared with the other models, characterized by higher prediction accuracy, more consistent residual distributions, and better agreement with the experimental data.

The outstanding performance of XGBoost can be attributed to several factors, including:

- *Its ability to handle noisy and outlier data.*
- *Its capability to model complex nonlinear relationships.*
- *Its high computational efficiency.*
- *Its relative robustness against overfitting.*

Compared with conventional models such as MLPNN, ANFIS, and GEP, the ensemble models employed in this study not only provided higher prediction accuracy but also enabled feature interpretation through the SHAP methodology. The SHAP analysis revealed that the water-to-binder ratio (W/B) and concrete age are among the most influential parameters affecting splitting tensile strength.

In conclusion, this study demonstrates the strong effectiveness of ensemble ML models in construction materials engineering and provides a basis for the broader application of artificial intelligence in concrete mix-design optimization. The use of such models can potentially reduce laboratory testing costs, accelerate the mix-design process, and improve concrete quality and performance in civil engineering projects.

References

[1] Wu, N. J. (2021). Predicting the compressive strength of concrete using an RBF-ANN model. *Applied sciences*, *11*(14), 1-10. https://doi.org/10.3390/app11146382

[2] Lin, C. J., & Wu, N. J. (2021). An ANN model for predicting the compressive strength of concrete. *Applied sciences*, *11*(9), 1-13. https://doi.org/10.3390/app11093798

[3] Kovačević, M., Lozančić, S., Nyarko, E. K., & Hadzima-Nyarko, M. (2021). Modeling of compressive strength of self-compacting rubberized concrete using machine learning. *Materials*, *14*(15), 1-25. https://doi.org/10.3390/ma14154346

[4] Nafees, A., Javed, M. F., Khan, S., Nazir, K., Farooq, F., Aslam, F., … ., & Vatin, N. I. (2021). Predictive modeling of mechanical properties of silica fume-based green concrete using artificial intelligence approaches: MLPNN, ANFIS, and GEP. *Materials*, *14*(24), 1-28. https://doi.org/10.3390/ma14247531

[5] Maherian, M. F., Baran, S., Bicakci, S. N., Toreyin, B. U., & Atahan, H. N. (2023). Machine learning-based compressive strength estimation in nano silica-modified concrete. *Construction and building materials*, *408*, 133684. https://doi.org/10.1016/j.conbuildmat.2023.133684

[6] Nagaraju, T. V., Mantena, S., Azab, M., Alisha, S. S., El Hachem, C., Adamu, M., & Rama Murthy, P. S. (2023). Prediction of high strength ternary blended concrete containing different silica proportions using machine learning approaches. *Results in engineering*, *17*, 100973. https://doi.org/10.1016/j.rineng.2023.100973

[7] Shah, H. A., Nehdi, M. L., Khan, M. I., Akmal, U., Alabduljabbar, H., Mohamed, A., & Sheraz, M. (2022). Predicting compressive and splitting tensile strengths of silica fume concrete using M5P model tree algorithm. *Materials*, *15*(15), 1-17. https://doi.org/10.3390/ma15155436

[8] Zhang, Y., Ren, W., Chen, Y., Mi, Y., Lei, J., & Sun, L. (2024). Predicting the compressive strength of high-performance concrete using an interpretable machine learning model. *Scientific reports*, *14*(1), 28346. https://doi.org/10.1038/s41598-024-79502-z

[9] Tien Bui, D., Abdullahi, M. M., Ghareh, S., Moayedi, H., & Nguyen, H. (2021). Fine-tuning of neural computing using whale optimization algorithm for predicting compressive strength of concrete. *Engineering with computers*, *37*(1), 701–712. https://doi.org/10.1007/s00366-019-00850-w

[10] Al-Abdaly, N. M., Seno, M. E., Thwaini, M. A., Imran, H., Ostrowski, K. A., & Furtak, K. (2024). Advanced ensemble machine-learning models for predicting splitting tensile strength in silica fume-modified concrete. *Buildings*, *14*(12), 4054. https://doi.org/10.3390/buildings14124054

[11] Ansari, S. S., Ansari, M. A., Saqib, M., Ghazi, M. S., & Ibrahim, S. M. (2024). Impact of thermal loads on silica fume-modified lightweight concrete: Machine learning approach to assess compressive strength evolution. *Materials today: Proceedings*. Elsevier. https://doi.org/10.1016/j.matpr.2024.04.054

[12] Abdulrahman, P. I., Jaf, D. K. I., Malla, S. K., Mohammed, A. S., Kurda, R., Asteris, P. G., & Sihag, P. (2025). Predictive modeling of compressive strength in silica fume-modified self-compacted concrete: A soft computing approach. *Structural concrete*, *26*(3), 2340–2365. https://doi.org/10.1002/suco.202400931

[13] Babiker, A., Abbas, Y. M., Khan, M. I., & Abdel-Magid, T. (2024). Optimizing compressive strength of quaternary-blended cement concrete through ensemble-instance-based machine learning. *Materials today communications*, *39*, 109150. https://doi.org/10.1016/j.mtcomm.2024.109150

[14] Alhakeem, Z. M., Jebur, Y. M., Henedy, S. N., Imran, H., Bernardo, L. F. A., & Hussein, H. M. (2022). Prediction of ecofriendly concrete compressive strength using gradient boosting regression tree combined with GridSearchCV hyperparameter-optimization techniques. *Materials*, *15*(21), 7432. https://doi.org/10.3390/ma15217432

[15] Naved, M., Asim, M., & Ahmad, T. (2024). Prediction of concrete compressive strength using deep neural networks based on hyperparameter optimization. *Cogent engineering*, *11*(1), 2297491. https://doi.org/10.1080/23311916.2023.2297491

[16] Dong, X., Liu, Y., & Dai, J. (2024). Application of fully connected neural network-based PyTorch in concrete compressive strength prediction. *Advances in civil engineering*, *2024*(1), 8048645. https://doi.org/10.1155/2024/8048645

[17] Alibrahim, B., Habib, A., & Habib, M. (2025). Developing a brain inspired multilobar neural networks architecture for rapidly and accurately estimating concrete compressive strength. *Scientific reports*, *15*(1), 1989. https://doi.org/10.1038/s41598-024-84325-z

[18] Marchewka, A., Ziolkowski, P., & García Galán, S. (2025). Real-time prediction of early concrete compressive strength using AI and hydration monitoring. *Scientific reports*, *15*(1), 15463. https://doi.org/10.1038/s41598-025-97060-w

[19] Amar, M., Benzerzour, M., Zentar, R., & Abriak, N. E. (2022). Prediction of the compressive strength of waste-based concretes using artificial neural network. *Materials*, *15*(20), 7045. https://doi.org/10.3390/ma15207045

[20] Mater, Y., Kamel, M., Karam, A., & Bakhoum, E. (2022). ANN-Python prediction model for the compressive strength of green concrete. *Construction innovation*, *23*(2), 340–359. https://doi.org/10.1108/CI-08-2021-0145

[21] Peng, Y., & Unluer, C. (2022). Analyzing the mechanical performance of fly ash-based geopolymer concrete with different machine learning techniques. *Construction and building materials*, *316*, 125785. https://doi.org/10.1016/j.conbuildmat.2021.125785

[22] Terlumun, S., Onyia, M. E., & Okafor, F. O. (2024). Predicting the compressive strength of self-compacting concrete using artificial intelligence techniques: A review. *AI in civil engineering*, *3*(1), 14. https://doi.org/10.1007/s43503-024-00029-3

[23] Shelote, K. M., Bala, A., & Gupta, S. (2023). An overview of mechanical, permeability, and thermal properties of silica fume concrete using bibliographic survey and building information modelling. *Construction and building materials*, *385*, 131489. https://doi.org/10.1016/j.conbuildmat.2023.131489

[24] Ben Chaabene, W., Flah, M., & Nehdi, M. L. (2020). Machine learning prediction of mechanical properties of concrete: Critical review. *Construction and building materials*, *260*, 119889. https://doi.org/10.1016/j.conbuildmat.2020.119889

[25] Al-Gburi, M., & Yusuf, S. A. (2022). Investigation of the effect of mineral additives on concrete strength using ANN. *Asian journal of civil engineering*, *23*(3), 405–414. https://doi.org/10.1007/s42107-022-00431-1