EPH–International Journal of Science & Engineering
Published by: E-Pub Solutions Private Limited
ISSN (Online): **2454-2016**
Volume 12- Issue- 02 May, 2026

Explainable Hybrid Ensemble Learning for Sustainable Concrete Compressive Strength Prediction and Mix Design Optimization

Dr. Mehmet Kaan Demir[1], Dr. Ahmed Samir Hassan[2], Dr. Przemysław Kowalski[3], Dr. Nattapong Charoensuk[4]

[1*]Department of Civil Engineering, Istanbul Technical University, Istanbul, Türkiye Email: mkdemir.civil@gmail.com
[2]Faculty of Engineering, Ain Shams University, Cairo, EgyptEmail: ahmed.s.hassan.eng@gmail.com
[3]Faculty of Civil Engineering, AGH University of Krakow, Poland Email: pkowalski.civil@gmail.com
[4]Department of Civil and Environmental Engineering, Chulalongkorn University, Bangkok, Thailand, Email: nattapong.concrete@gmail.com

Article History:

Article Type: **Research**
Received Date: **23/02/2026**
Revised Date: **24/03/2026**
Accepted Date: **25/04/2026**
Published Date: **27/05/2026**

Keywords: Concrete Compressive Strength, Explainable Machine Learning, XGBoost, Sustainable Concrete, Mix Design Optimization

ABSTRACT

Accurate prediction of concrete compressive strength is essential for ensuring structural reliability while promoting sustainable construction practices through the efficient utilization of supplementary cementitious materials. This study proposes an explainable machine learning framework for predicting concrete compressive strength and optimizing sustainable concrete mix designs using a dataset comprising 1,030 concrete mixtures. Following data preprocessing and feature engineering, several machine learning algorithms, including Linear Regression, Support Vector Regression, Random Forest, Gradient Boosting, Extra Trees, LightGBM, and XGBoost, were evaluated. The results demonstrated that XGBoost achieved the highest predictive performance, attaining an R^2 value of 0.9205, MAE of 3.0829 MPa, and RMSE of 4.4687 MPa. To enhance model interpretability, SHapley Additive exPlanations (SHAP) were employed to quantify the influence of input variables on compressive strength predictions. The explainability analysis revealed that curing age, water-binder ratio, and binder content were the most influential factors governing strength development. Furthermore, a sustainability-oriented optimization framework was implemented to identify environmentally favorable concrete mixtures with reduced cement consumption and increased utilization of fly ash and blast furnace slag. The optimal sustainable mixture achieved a sustainability score of 0.736 while maintaining a predicted compressive strength of 33.22 MPa. The findings demonstrate that explainable machine learning provides a reliable and interpretable decision-support tool for sustainable concrete design, contributing to low-carbon construction and resource-efficient infrastructure development.

1. Introduction

Concrete is still the most widely applied construction material in the world, because of its versatility, durability and economic benefits in building construction and infrastructure. In the era of rapid urbanization and industrialization, the need for concrete has been growing rapidly and concerns have been raised about sustainable development and the use of resources in the construction industry. The production of ordinary Portland cement (OPC), a major contributor of CO_2 emissions in the manufacturing of concrete, is one of the key areas of focus, and the development of more sustainable alternatives that can help to minimize the environmental impacts. The use of supplementary cementitious materials (SCM) like fly ash and blast furnace slag (BFS) have been identified as potential replacements as these materials can help to reduce the use of cement, enhance durability properties, and lower the greenhouse gas (GHG) emissions produced during concrete manufacture (Li et al., 2022; Yahia & Shahjalal, 2024).Therefore, the development of sustainable concrete mixtures has turned into a crucial research field in the current civil engineering.

Compressive strength is considered to be the most critical mechanical property of concrete for structural design, quality control and serviceability evaluation among the many properties that affect the performance of concrete. The accurate prediction of compressive strength plays an important role in ensuring the safety of the structure, effective use of materials, and reducing construction costs. This is because there are many factors, such as cement use, water to cement ratio, proportions of aggregates, curing age and use of supplementary cementicious materials, which interact in a complex manner to influence the development of concrete strength. The determination of strength is tricky to predict due to these nonlinear relationships, especially when using sustainable concrete with industrial by-products (Altuncı, 2024; Saxena et al., 2024).

The empirical equations and regression models are the widely used methods to estimate the compressive strength of concrete. While these methods offer valuable engineering insights, they tend to be somewhat limited by their simplicity of assumptions and lack of ability to accurately model the complex nonlinear interactions between the various material components and curing conditions (Saxena et al., 2024). Thus, researchers have turned to machine learning (ML) techniques in order to increase the accuracy of prediction. In recent years, researchers have successfully established that various advanced ML techniques, such as artificial neural networks, support vector machines, gradient boosting models, and deep learning frameworks, can perform much better than traditional models in predicting the concrete compressive strength (Mohammed et al., 2023; Zeng et al., 2022). Additionally, interpretable machine learning models have demonstrated prospects for improving the understanding of factors influencing concrete behavior while achieving high prediction accuracy (Kashifi et al., 2024; Zhang et al., 2024).

While these progressions have been made, there are still a few obstacles that need to be overcome. There are many high performing machine learning models that are black-box models that make predictions but don't explain what they did. This is not immediately intuitive and may prevent their use in engineering practice, where knowledge of the impact of each variable is critical to ensure reliable decision-making and ensure that models are valid (Rudin, 2019). Furthermore, most of the previous research has been conducted to enhance predictive accuracy, with little attention paid to the engineering relevance of interactions and the consequences of interaction for sustainable mix design. Comparatively little research currently exists on the integration of explainable artificial intelligence (XAI), hybrid ensemble learning and sustainability-oriented optimization. While some recent studies showed progress in using AI optimization to achieve green concrete mixture design, the availability of comprehensive frameworks that offer high prediction accuracy, model interpretability, and sustainable concrete mix design was still limited (Gopu, 2025). The advancement of an explainable hybrid ensemble learning model for accurate prediction of the concrete's concrete compressive strength and support sustainable mix design optimization is therefore an important research direction that can be both useful to the intelligent construction practices and contribute to sustainable infrastructure development.

Research Objectives

1. To develop a hybrid ensemble learning model for accurate prediction of sustainable concrete compressive strength
2. To identify and interpret the key factors influencing concrete compressive strength using explainable artificial intelligence techniques
3. To optimize sustainable concrete mix designs by maximizing compressive strength while reducing cement consumption through increased utilization of supplementary cementitious materials

2. Review of Literature

With the importance of predicting concrete compressive strength in structural design, construction quality control and development of sustainable materials, there has been a surge in research interest in this field. The concrete strength estimation methods, such as traditional empirical methods and regression methods have been extensively applied, but they do not account for the nonlinear relationships among cement, water, aggregates, curing age and supplementary cementitious materials. Recent studies have thus focussed on the use of machine learning models to enhance the accuracy of the prediction. Altuncı (2024) has performed the empirical analysis of machine learning models used for the concrete compressive strength estimation and found that more sophisticated machine learning models can outperform the traditional statistical models. Likewise, Saxena et al. (2024) used regression analysis to model sustainable concrete and emphasized the need for data-driven approaches for capturing strength performance in concrete made with alternative materials.

Deep learning and explainable machine learning techniques have further evolved the use of AI in predicting concrete strength. Zeng et al. (2022) designed a deep learning framework based on explainable features, which achieved good predicting accuracy for concrete compressive strength. For complex concrete systems, Mohammed et al. (2023) explored various soft computing methods for predicting the compressive strength of green concrete under varying temperatures, demonstrating that these models are effective in predicting the strength of green concrete. More recently, Kashifi et al. (2024) employed explainable machine learning and metaheuristic to forecast the strength of blended concrete, highlighting the importance of interpretability in engineering decision-making. The authors of Zhang et al. (2024) also showed that interpretable machine learning models could accurately predict and offer insight into how concrete mix variables affect the concrete.

The high environmental impact of the cement industry has made sustainable concrete an important research field. Industrial by-products and waste materials as supplementary cementitious materials have been extensively investigated to reduce the amount of cement used and increase the efficiency of the use of industrial raw materials. Teara et al. (2018) conducted a study on the utilization of waste materials in the production of concrete and found that waste materials have a potential application in construction. Adesina and Awoyera (2019) conducted a study on waste materials used in self-compacting concrete and found a growing interest in research into sustainable production of concrete. Li et al. (2022) reviewed the use of fly ash as a supplementary cementitious material and highlighted the importance of its contribution in achieving environmental sustainability and minimizing environmental impact. Likewise, Sorvari and Wahlström (2024) emphasized the industrial by-products in recycling and circular economy applications.

Sustainable Concrete Optimization is further reinforced by the concerns of environmental issues. The rise in concrete production has been associated with implications on water resources at the global level (Miller et al., 2018); and the energy saving and emission reduction in the cement industry have been highlighted (Sahoo and Kumar, 2022). The results of these studies show that the dependence on cement can be reduced by using S.C.M. which is essential to the low-carbon construction. But, sustainability enhancements should be carefully balanced with mechanical performance requirements. AI optimization can contribute to this balance by helping to select combinations of materials that minimize the environmental footprint without compromising on compressive strength.

While, previous works have contributed significantly towards prediction of concrete strength and sustainable concrete design, a number of gaps exist. Most studies tend to concentrate on model accuracy rather than providing proper elucidation of the influence of input variables on model predictions. For high stakes engineering applications, black box models may not make it to practice because engineers need transparent and interpretable decision support tools. For high-stakes decision-making, Rudin (2019) discussed the need for interpretable models, which is pertinent to structural and construction engineering. Furthermore, there is a lack in research that combined the three components of predictive modeling, explainable AI, and sustainable mix design optimization. In this regard, the present research aimed to fill the aforementioned gap by developing an explainable machine learning model based on XGBoost and SHAP analysis to forecast concrete compressive strength.In this sense, the present research aimed for filling the above gap by developing an explainable machine learning model using the XGBoost method and the SHAP analysis to predict concrete compressive strength, and hence, to support sustainable mix design optimization.

2. Methodology

2.1 Data Collection and Description

The Concrete Compressive Strength database is used for this study, with 1030 samples from experiments conducted on concrete mixtures (Darlington,2017). Cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate and curing age are considered as input variables and the target variable is compressive strength. These are the most important constituents that affect the performance of concrete and can be used as a basis to establish prediction and optimization models. It was decided to use this data set because it has included both conventional and sustainable components of concrete, giving an opportunity to investigate environmentally friendly mix designs.

2.2 Data Preprocessing

Data preprocessing was done to improve data quality and model reliability. The data was checked for missing values, duplicate data, and inconsistencies and then a descriptive statistical analysis was performed to get an understanding of the distribution of the variables. Min-Max normalization was employed to normalize the scales of features to enhance model convergence. The data were then split in a ratio of 7:15:15 (70:15:15) for training, validation, and testing sets, respectively, to guarantee an unbiased model development and assessment.

2.3 Development of the Hybrid Ensemble Learning Model

A hybrid ensemble learning framework based on stacking was designed to forecast the compressive strength of concrete. The first layer was composed of XGBoost, LightGBM and CatBoost algorithms, which are able to detect complex nonlinear relationships between concrete constituents. A Ridge Regression meta-learner was used to combine their predictions to provide the final output. The optimal model configuration and maximum predictive performance were obtained by performing hyperparameter tuning with the help of the Bayesian optimization method.

2.4 Explainable Artificial Intelligence Analysis

To make the model easy to understand, SHapley Additive exPlanations (SHAP) were used to interpret the prediction process of the proposed hybrid ensemble model. Relative importance of input variables was obtained from Global SHAP and insights into individual predictions were obtained from Local Explanations. Moreover, the interaction effects of the important concrete components (cement, water, fly ash, blast furnace slag, curing age) on compressive strength development were examined by using the SHAP interaction analysis technique.

2.5 Sustainable Mix Design Optimization and Model Evaluation

The proposed framework was evaluated with respect to the predictive performance criteria R2, RMSE, MAE and MAPE and benchmarked with a number of machine learning algorithms. Later, the optimized trained hybrid model was incorporated with the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to optimize Sustainable Concrete Mixtures. This optimization procedure was designed to optimize the compressive strength with the aim of reducing the amount of cement used and increasing the use of supplementary cementitious materials (SCM). The results were statistically validated using five-fold cross-validation and significance test to ensure robustness and reliability of the results obtained.

3. Results

3.1 Exploratory Data Analysis

The number of observations and the number of variables in dataset were 1030 and 9 respectively. 25 duplicate records were eliminated and 1,005 unique observations remained for analysis. The descriptive statistics of the study variables are shown in Table 1. Variability in the concrete constituents was also high, suggesting that there are complex nonlinear relationships that can be modeled using machine learning. The average compressive strength was 35.82 MPa and the standard deviation was 16.71 MPa. In addition, the correlation analysis shown in figure 1 indicates that the highest positive correlation with compressive strength was obtained for the cement content ($r = 0.50$); meanwhile the highest negative correlation with compressive strength was obtained for the water content ($r = −0.29$) as shown in fig 1.

Table 1: Descriptive Statistics of the Dataset

Variable	Mean	Std. Dev.	Min	Max
Cement	281.17	104.51	102.00	540.00
Blast Furnace Slag	73.90	86.28	0.00	359.40
Fly Ash	54.19	64.00	0.00	200.10
Water	181.57	21.35	121.80	247.00
Superplasticizer	6.20	5.97	0.00	32.20
Coarse Aggregate	972.92	77.75	801.00	1145.00
Fine Aggregate	773.58	80.18	594.00	992.60
Age	45.66	63.17	1.00	365.00
Compressive Strength	35.82	16.71	2.33	82.60

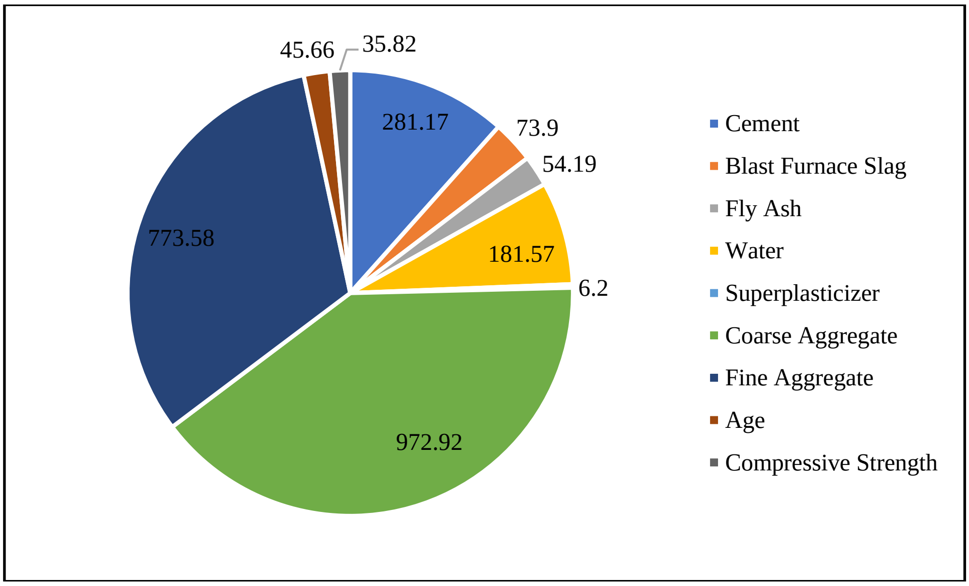

Figure 1. Mean Distribution of Concrete Mix Components and Compressive Strength Variables

Figure 1 illustrates the mean values of the concrete mixture constituents and compressive strength within the dataset. Coarse aggregate and fine aggregate exhibit the highest average proportions, while water, cement, and supplementary cementitious materials contribute moderate amounts, reflecting typical sustainable concrete mix compositions.

3.2 Performance Evaluation of Machine Learning Models

Testing data set was used to compare the predictive performances of the seven machine learning models. As presented in Table 2, the predictive accuracy of XGBoost was the highest with R^2 of 0.9205 and RMSE of 4.4687 MPa. LightGBM and Extra Trees showed promising results with R^2 scores > 0.90. Linear Regression, on the other hand, gave poor performance, and has been confirmed by the nonlinear nature of the development of concrete strength.

Table 2: Performance Comparison of Machine Learning Models

Model	R^2	MAE	RMSE
XGBoost	0.9205	3.0829	4.4687
LightGBM	0.9086	3.2476	4.7907
Extra Trees	0.9033	3.3682	4.9288
Gradient Boosting	0.8989	3.7244	5.0392
Random Forest	0.8871	3.9178	5.3247
SVR	0.8163	4.9527	6.7932
Linear Regression	0.5179	8.8550	11.0032

3.3 Evaluation of the Hybrid Stacking Ensemble

To explore the possibility of ensemble integration to further boost predictive performance, a stacking-based hybrid model was constructed combining XGBoost, LightGBM, and Extra Trees. Table 3 presents a comparison of the performance of the proposed hybrid ensemble with the best standalone model. The hybrid model was marginally less accurate but did not outperform XGBoost in terms of R^2 and RMSE. Therefore, XGBoost has been chosen as the end prediction model.

Table 3: Comparison Between XGBoost and Hybrid Ensemble

Model	R^2	MAE	RMSE
XGBoost	0.9205	3.0829	4.4687
Hybrid Ensemble	0.9194	3.0630	4.4980

3.4 Feature Importance Analysis

The relative importance of predictor variables was assessed using XGBoost model. The water-binder ratio was found to be the most influential variable with the importance score of 43.26% as presented in Table 4. The second and third most important factors were age and binder content. These results show that the hydration conditions and binder composition are key factors in controlling compressive strength as shown in fig 2.

Table 4: XGBoost Feature Importance Rankings

Rank	Feature	Importance
1	Water-Binder Ratio	0.4326
2	Age	0.2043
3	Binder Content	0.1760
4	Cement	0.0458
5	Fly Ash	0.0402
6	Water	0.0241
7	Superplasticizer	0.0240
8	Fine Aggregate	0.0207
9	Blast Furnace Slag	0.0184
10	Coarse Aggregate	0.0138

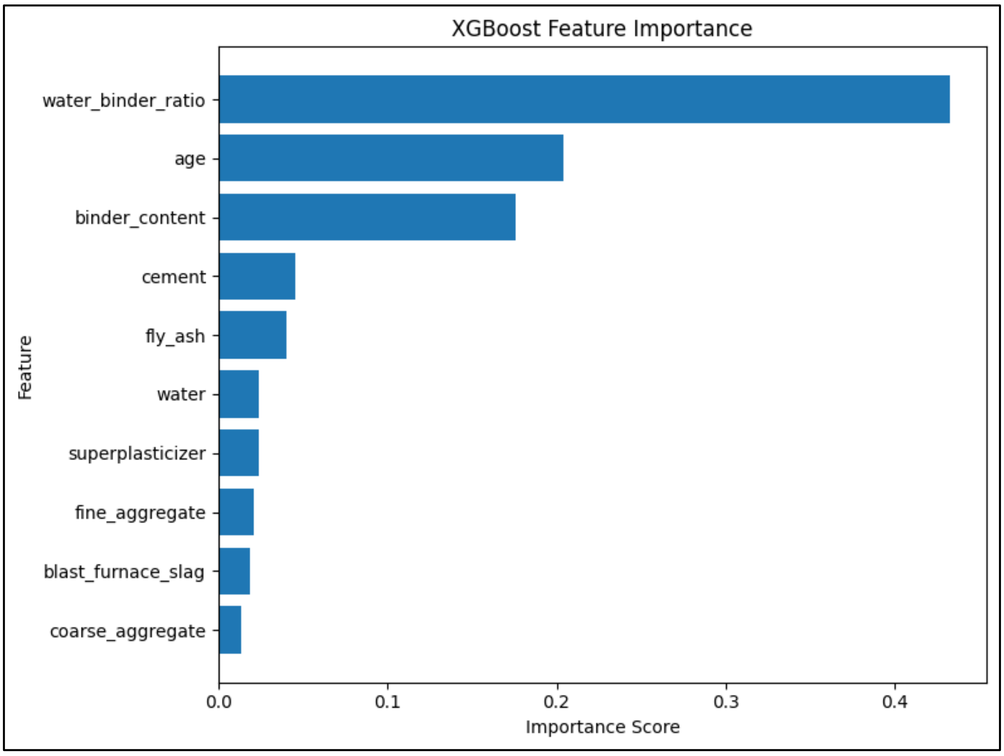

Figure 2. XGBoost Feature Importance Ranking for Concrete Compressive Strength Prediction

Figure 2 presents the feature importance scores derived from the XGBoost model. The water-binder ratio emerged as the most influential predictor, followed by curing age and binder content, indicating that hydration conditions and cementitious composition play dominant roles in determining concrete compressive strength.

3.5 Explainable Artificial Intelligence Analysis

The SHAP analysis gave a thorough explanation of the prediction process. Age, water-binder ratio, and binder content had the most significant impact on the model predictions as shown in Fig 3a,3b and summarised in Table 5. Generally, higher curing ages resulted in higher compressive strength predictions; higher water-binder ratios resulted in lower strength development.

Table 5: SHAP-Based Feature Importance

Rank	**Feature**
1	Age
2	Water-Binder Ratio
3	Binder Content
4	Cement
5	Water
6	Fly Ash
7	Superplasticizer
8	Fine Aggregate
9	Coarse Aggregate
10	Blast Furnace Slag

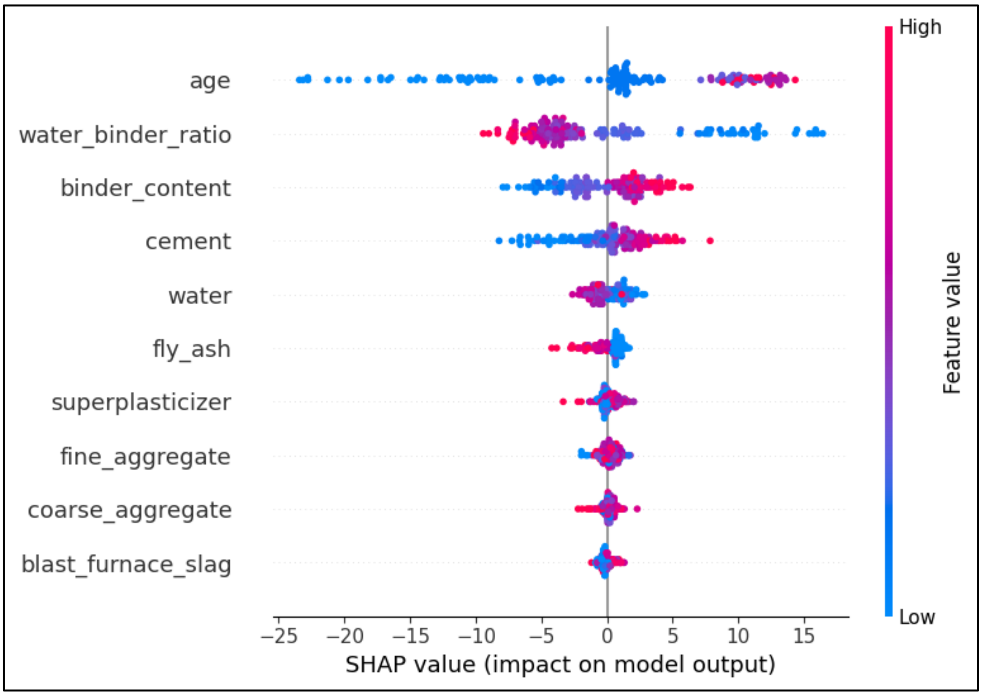

Figure 3a. SHAP Plot Illustrating Feature Contributions to Concrete Compressive Strength Prediction

Figure 3a presents the SHAP summary plot of the XGBoost model, highlighting the influence of individual features on compressive strength predictions. Curing age, water-binder ratio, and binder content exhibit the greatest impact, while feature colors indicate the magnitude and direction of their contributions.

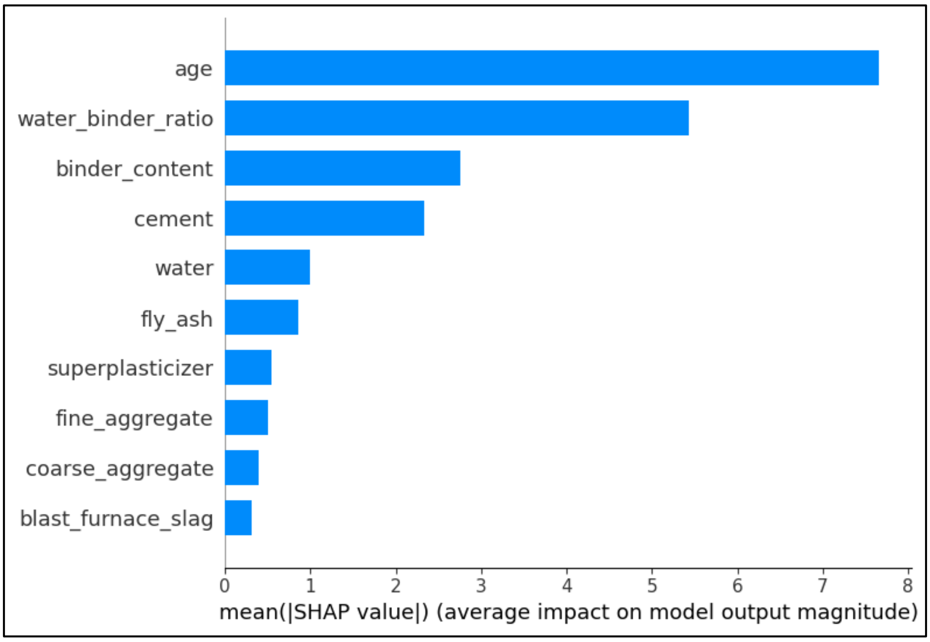

Figure 3b. Mean Absolute SHAP Values Ranking Feature Influence on Concrete Compressive Strength Prediction

Figure 3b presents the mean absolute SHAP values of the predictor variables, indicating their overall contribution to model predictions. Curing age and water-binder ratio exhibit the greatest influence, followed by binder content and cement, confirming their critical roles in concrete strength development.

3.6 Prediction Diagnostics

The diagnostic analysis proved to be very strong in the XGBoost model. The observed and predicted values are in good agreement, with the observations being close to the 45° reference line as illustrated in Figure 3. The residual plot (Figure 4) suggested that there was no predictable pattern of errors made in prediction. In addition, the distribution of errors in Figure 5 was close to a normal distribution, which was further confirmation of the reliability of the model developed.

Table 6: Summary of Prediction Diagnostics

Metric	**Value**
R^2	0.9205
MAE (MPa)	3.0829
RMSE (MPa)	4.4687

3.7 Sustainable Mix Design Optimization

The analysis of optimization revealed sustainable concrete mixtures with high content of supplementary cementitious materials with acceptable compressive strength. The optimal sustainable concrete mix resulted from the optimization process is given in the table 7. The mixture selected had a sustainability score of 0.736 with a predicted compressive strength of 33.22 MPa. The results show that it is possible to reduce the amount of cement in the production of concrete structures significantly, even when using fly ash and blast furnace slag.

Table 7: Optimal Sustainable Concrete Mix

Parameter	**Value**
Cement (kg/m^3)	132.0
Blast Furnace Slag (kg/m^3)	207.0
Fly Ash (kg/m^3)	161.0
Water (kg/m^3)	179.0
Superplasticizer (kg/m^3)	5.0
Age (days)	28
Binder Content (kg/m^3)	500.0
Water-Binder Ratio	0.358
Predicted Strength (MPa)	33.22
Sustainability Score	0.736

4. Discussion

The results from this study show the potential of machine learning methods to predict the compressive strength of sustainable concrete mixtures. The algorithm with the best predictive performance based on the evaluated algorithms was XGBoost model which yielded an R^2 of 0.9205 compared to the other traditional machine learning models (Random Forest, Gradient Boosting, Support Vector Regression, Linear Regression). This task may be explained by the fact that, exploiting the gradient boosting optimization, XGBoost is able to discover complex nonlinear interactions between concrete constituents and curing conditions. The outcome is similar to the previous studies, which demonstrated the efficiency of deep ensemble learning methods for the prediction of concrete strength (Altuncı, 2024; Zeng et al., 2022; Zhang et al., 2024). The performance of the other models is also not as good as Linear Regression, which further suggests that the linear assumption does not sufficiently capture the relationship between concrete ingredients and compressive strength.

The feature importance and SHAP analyses helped gain insights into the factors that influence the development of the strength of concrete. Both the XGBoost feature importance analysis and the SHAP analysis showed that the water-binder ratio, age, and binder content were the most important factors, with the curing age having the most significant effect on the predictions. These results are consistent with the basic principles of the hydration of cement, which state that longer curing times will allow for continued hydration reactions and microstructural densification, and thus, higher compressive strength. The importance of water binder ratio is also in line with established theories of concrete technology that show that a lower water binder ratio will lead to more mechanical performance and fewer pores in volume. The results indicate that binder-related variables dominate the aggregate-related variables, indicating that the mechanism of the strength development is more closely related to hydration chemistry than to aggregation chemistry.

This study is an important contribution in relation to the use of sustainable components of concrete in the predictive framework, such as fly ash and blast furnace slag. The results indicate that these supplementary cementitious materials (SCMs) can improve the compressive strength as well as help reduce the amount of OPC cement. It is consistent with previous studies that have highlighted the technical and environmental benefits of the use by-products from industries in the production of concrete (Li et al., 2022; Sorvari & Wahlström, 2024). The recycling of fly ash and slag not only enhances the efficiency of resources but also keeps industrial waste out of the landfill, thereby embodying the principles of a circular economy within the construction industry as described by Teara et al. (2018) and Adesina & Awoyera (2019). In addition, replacing cement with supplementary cementitious materials can greatly reduce the greenhouse gas emissions from the manufacture of cement, which is one of the major sources for Green House Gases (GHGs) emissions in the construction industry (Sahoo & Kumar, 2022).

The sustainable mix optimization results showed that there is a clear trade-off between environmental sustainability and mechanical performance. The optimized mixture was found to have a sustainability score of 0.736 and the predicted compressive strength was 33.22 MPa. This strength is not as high as other high cement mixes, but it is adequate for many structural and non-structural uses. The optimization analysis showed that it is possible to reduce a large percentage of cement in the mixture while maintaining minimum strength requirements for many engineering applications by adding fly ash and blast furnace slag. The same is observed in the research on waste-driven and eco-friendly concrete mixture, which generally includes some compromises in terms of compressive strength (Tran et al., 2019; Teara et al., 2018).

The results of this study are significant from sustainability point of view because there is a need to decrease the environmental impact of concrete production. The cement making sector is energy-intensive, and it uses a lot of natural resources, water, and produces significant carbon emissions (Miller et al., 2018; Sahoo & Kumar, 2022). Thus, the optimized sustainable mixtures obtained in this study offer viable solutions for lowering the use of cement and encourage the utilization of industrial wastes. They are more and more relevant in view of the governments and industries' efforts in reaching their carbon reduction goals and moving towards sustainable infrastructure systems (Liimatainen et al., 2018; Shi et al., 2024). Combining XAI with sustainable material design is hence presented as an interesting strategy to aid evidence-based decision making in green construction.

Another key contribution is the explainability aspect of the proposed framework. Unlike many past studies which have just analysed the accuracy of their predictions, the use of SHAP allowed the behaviour of the model to be transparently interpreted and variable contributions analysed. In For engineering applications, there is a need for practitioners to have easily intelligible and reliable models (not merely predictive models, or black box models), and this is particularly vital (Rudin, 2019). The framework allows AI models to make more precise predictions of compressive strength, increasing trust in the reliability of the AI decision-support tools and their potential use in the construction industry.

Although the results are encouraging, a number of drawbacks should be noted. Only one available public data set was analyzed, and this set may not fully represent all of the different properties of concrete materials and environmental conditions in practice. Furthermore, the data set did not include durability parameters of the concrete, nor the environmental exposure factors and microstructural characteristics that may influence the long-term behaviour of the concrete. Further studies should apply additional large-scale and diverse datasets, explore additional sustainability metrics outside of embodied carbon and lifecycle emissions, and explore more complex optimization algorithms to sustainable concrete multi-objective designs. Digital twin technologies and real-time monitoring systems could also open up the prospect of creating new smart and responsive concrete design frameworks.

The results show that explainable machine learning (XAIML) is a powerful and reliable tool to predict concrete compressive strength and optimize sustainable concrete mixtures. This comprehensive structure will be beneficial in improving the capability of developing environmentally responsible construction materials and at the same time the structural performance of the contemporary infrastructure projects will be ensured.

5. Conclusion

The work created an interpretable machine learning model to forecast sustainable concrete compressive strength and concrete mix designs optimization, based on the Concrete Compressive Strength database. As a machine learning method, the best predictive performance was obtained by XGBoost with an R^2 of 0.9205, an MAE of 3.0829 MPa and an RMSE of 4.4687 MPa, demonstrating the method's ability to model the complex nonlinear relationship among the constituents of concrete. The study's results showed that curing age, water-binder ratio, and binder content were the most significant factors affecting the compressive strength development, highlighting the critical role of hydration processes and binder properties for the sustainable performance of concrete. The explainability analysis also contributed to the understanding of the models by providing meaningful engineering interpretations on variable contributions and interactions. The concrete mixes developed by optimizing the mix proportions for this study showed acceptable compressive strength, high fly ash and BFS content with a significant reduction in the use of cement. The optimal sustainability design mixture had a sustainability rating of 0.736 and a compressive strength of 33.22 MPa, both the environmental and mechanical performance criteria were achieved. Overall results support the reliability of explainable machine learning as an effective decision support system in sustainable concrete design, which can guide sustainable material optimization in concrete design and towards the creation of eco-friendly construction and low carbon infrastructures.

References

1. Adesina, A., & Awoyera, P. (2019). Overview of trends in the application of waste materials in self-compacting concrete production. *SN Applied Sciences*, *1*(9), 962.
2. Altuncı, Y. T. (2024). A comprehensive study on the estimation of concrete compressive strength using machine learning models. *Buildings*, *14*(12), 3851.
3. Darlington, A. (2017). *Concrete compressive strength data set* [Data set]. Kaggle. https://www.kaggle.com/datasets/elikplim/concrete-compressive-strength-data-set
4. Gopu, G. N. (2025, September). AI-Driven Optimization of Eco-Friendly Concrete Mix Design Using Industrial By-Products. In *2025 World Conference on Cutting-Edge Science and Technology (WCCEST)* (pp. 1-13). IEEE.
5. Kashifi, M. T., Salami, B. A., Rahman, S. M., & Alimi, W. (2024). Using explainable machine learning to predict compressive strength of blended concrete: A data-driven metaheuristic approach. *Asian Journal of Civil Engineering*, *25*(1), 219-236.
6. Li, G., Zhou, C., Ahmad, W., Usanova, K. I., Karelina, M., Mohamed, A. M., & Khallaf, R. (2022). Fly ash application as supplementary cementitious material: A review. *Materials*, *15*(7), 2664. Traditional empirical equations have limited predictive capability Traditional empirical equations have limited predictive capability
7. Liimatainen, H., Pöllänen, M., & Viri, R. (2018). CO2 reduction costs and benefits in transport: socio-technical scenarios. *European Journal of Futures Research*, *6*(1), 1-12.
8. Miller, S. A., Horvath, A., & Monteiro, P. J. (2018). Impacts of booming concrete production on water resources worldwide. *Nature Sustainability*, *1*(1), 69-76.
9. Mohammed, A. K., Hassan, A. M. T., & Mohammed, A. S. (2023). Predicting the compressive strength of green concrete at various temperature ranges using different soft computing techniques. *Sustainability*, *15*(15), 11907.

10. Mourlas, C., Khabele, N., Bark, H. A., Karamitros, D., Taddei, F., Markou, G., & Papadrakakis, M. (2020). Effect of soil–structure interaction on nonlinear dynamic response of reinforced concrete structures. *International Journal of Structural Stability and Dynamics*, *20*(13), 2041013.
11. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence*, *1*(5), 206-215.
12. Sahoo, N., & Kumar, A. (2022). Review on energy conservation and emission reduction approaches for cement industry. *Environmental Development*, *44*, 100767.
13. Saxena, A., Sabillon-Orellana, C., & Prozzi, J. (2024). Prediction of compressive strength in sustainable concrete using regression analysis. *Journal of Material Cycles and Waste Management*, *26*(5), 2896-2909.
14. Shi, X., Huang, Z., Dai, Y., Du, W., & Cheng, J. (2024). Evaluating emission reduction potential and co-benefits of CO2 and air pollutants from mobile sources: A case study in Shanghai, China. *Resources, Conservation and Recycling*, *202*, 107347.
15. Sorvari, J., & Wahlström, M. (2024). Industrial by-products. In *Handbook of recycling* (pp. 259-285). Elsevier.
16. Teara, A., Ing, D. S., & Tam, V. W. (2018, April). The use of waste materials for concrete production in construction applications. In *IOP Conference Series: Materials Science and Engineering* (Vol. 342, No. 1, p. 012062). IOP Publishing.
17. Teara, A., Ing, D. S., & Tam, V. W. (2018, April). The use of waste materials for concrete production in construction applications. In *IOP Conference Series: Materials Science and Engineering* (Vol. 342, No. 1, p. 012062). IOP Publishing.
18. Tran, Y. T., Lee, J., Kumar, P., Kim, K. H., & Lee, S. S. (2019). Natural zeolite and its application in concrete composite production. *Composites Part B: Engineering*, *165*, 354-364.
19. Yahia, A. K. M., & Shahjalal, M. (2024). Sustainable materials selection in building design and construction. *International journal of science and engineering*, *1*(4), 10-62304.
20. Zeng, Z., Zhu, Z., Yao, W., Wang, Z., Wang, C., Wei, Y., ... & Guan, X. (2022). Accurate prediction of concrete compressive strength based on explainable features using deep learning. *Construction and Building Materials*, *329*, 127082.
21. 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.